Promotor Prof. dr. Maarten Van Dyck
Vakgroep Wijsbegeerte en Moraalwetenschap
Copromotor Prof. dr. James Conant
Department of Philosophy
University of Chicago

FACULTEIT LETTEREN
EN WIJSBEGEERTE

Logical categories, signs, and elucidation in Frege

Wim Vanrie

Proefschrift voorgedragen tot het behalen van de graad van Doctor in de de Wijsbegeerte

2021

UNIVERSITEIT
GENT

Acknowledgments

This dissertation is deeply indebted to numerous conversations with my co-supervisor, Jim Conant. He, more than anyone else, has taught me how to find beauty in those great works of Frege. My gratitude to him, however, extends far beyond such intellectual affairs. Ever since I decided not to pursue a PhD at the University of Chicago, Jim nevertheless continued to support me in a variety of ways. His welcoming me as a visiting graduate student to the FAGI centre in Leipzig, for instance, brought me professional and personal experiences that resulted in friendships and fond memories that I will carry with me for the rest of my life.

Throughout the years, Maarten Van Dyck has evolved from being my undergraduate teacher to being my supervisor, to writing papers with me and teaching classes alongside me, and – most importantly – to paying for dinner at some of Ghent's finest culinary establishments. Yes, I still fondly remember the look on his face when he picked up the first of many checks to follow at our four-person teambuilding dinner at *Meat & Great*, a restaurant that – contrary to what Maarten's *Gestalt* at the time would suggest – has since gone out of business. But I digress. Although the topic of this dissertation is somewhat removed from Maarten's own expertise – so that his direct influence on these pages has been limited – this is easily offset by his influence on my overarching development, which has been immeasurable. Not only has Maarten had my back during moments of crisis and moments of joy, we have also – quite simply – had a lot of fun together. I sincerely hope that we will be able to continue our collaborations on all these fronts.

I wish to thank my jury members Gertrudis Van de Vijver, Maria van der Schaar, Martin Gustafsson, Michael Kremer, and Jean-Philippe Narboux for, well, being willing to be my jury members. I can think of no group of scholars to whom I would rather present the fruits of my work. In addition, I wish to thank Gert, not only for ceaselessly stimulating

me intellectually, but also for always making me feel welcome, and – of course – for her exquisite *cannelés*.

I am grateful to Michael Beaney, Florian Ganzinger, Jonas Held, Gino Margani, Adrian Moore, Stephen Mulhall, Gilad Nir, and Troy Vine for helping me think about Frege. A special mention goes out to the Layer Cake Boys Jonathan Gombin, Thomas Pendelbury, and Quentin Fisher for the wonderful and stimulating times we shared in Leipzig during that legendary Fall of 2018.

Thanks to the 120.048 gang – especially Annelies and Benjamin – for creating and maintaining the welcoming, light-hearted, and supportive work environment wherein I was able to flourish. One could not wish for better colleagues. Moving beyond the Blandijn, I wish to thank Benjamin De Mesel and Sybren Heyndels for hanging out with a *rijksfilosoof* such as myself without making me feel inferior.

From the bottom of my heart, I wish to thank Fons, Hannah, Lisa, Nicolas, Pascal, Wim, and Xof for their invaluable friendship, my parents for their unrelenting support, and Lisa for her love.

List of abbreviations

I use the following abbreviations to refer to the works of Frege, Russell, and Wittgenstein:

BS for the *Begriffsschrift* (Frege, 1970)
GL for the *Grundlagen* (Frege, 1968)
GGA for the *Grundgesetze* (Frege, 2013)
PMC for Frege's *Philosophical and Mathematical Correspondence* (Frege, 1980)
CP for Frege's *Collected Papers* (Frege, 1984)
PW for Frege's *Posthumous Writings* (Frege, 1979)
FLL for Frege's *Lectures on Logic* (Frege, 2004)
KS for Frege's *Kleine Schriften* (Frege, 1967)
NS for Frege's *Nachgelassene Schriften* (Frege, 1983)
BW for Frege's *Wissenschaftlicher Briefwechsel* (Frege, 1976)
PoM for Russell's *Principles of Mathematics* (Russell, 2010)
TLP for Wittgenstein's *Tractatus Logico-Philosophicus* (Wittgenstein, 1981)

For quoted passages from PMC, CP, and PW, I have included references to the German versions, and references to the original page numbers where applicable.

Table of contents

Introduction: Frege's predicament

> *Les comptes rendus d'histoire de la philosophie doivent représenter une sorte de ralenti, de figeage ou d'immobilisation du texte: non seulement du texte auquel ils se rapportent, mais aussi du texte dans lequel ils s'insèrent. Si bien qu'ils ont une existence double, et, pour double idéal, la pure répétition du texte ancien et du texte actuel l'un dans l'autre.*[1]
> *— Gilles Deleuze, Différence et répétition.*

There is a well-known predicament that arises in Frege's work. Frege's philosophical logic involves a sharp distinction between concepts and objects. There are objects, which are the meaning (*Bedeutung*) of proper names such as 'Seabiscuit', and there are concepts, which are the meaning (*Bedeutung*) of concept names such as 'ξ is a horse'. Whereas objects are *complete*, concepts are *unsaturated* and require completion by an object as argument. Any judgment involves the completion of what is unsaturated by a suitable argument. For instance, the judgment 'Seabiscuit is a horse' involves the completion of the concept *ξ is a horse* by the object Seabiscuit. Only objects can serve as arguments of concepts, and only concepts require completion. There is no such thing as a concept completing a concept, or an object completing an object, or a concept completing an object. As Frege puts it:

> Now it follows from the fundamental difference of objects from concepts that an object can never occur predicatively or unsaturatedly; and that logically, a concept can never stand in for an object. One could express it metaphorically like this: There

[1] Commentaries in the history of philosophy should represent a kind of slow motion, a congelation or immobilisation of the text: not only of the text to which they relate, but also of the text in which they are inserted – so much so that they have a double existence and a corresponding ideal: the pure repetition of the former text and the present text in one another.

> are different logical places; in some only objects can stand and not concepts, in others only concepts and not objects (CP, 281-282; KS, 270; 372).

Frege takes the distinction between objects and concepts to be absolute. Objects are *essentially* complete, and concepts are *essentially* unsaturated. It is a consequence of this view that there is no such thing as a variable encompassing both concepts and objects. There is no such thing as quantifying over both concepts and objects at once. Frege's metaphor of places brings this out nicely: whichever logical place is occupied by a variable, in that place can stand either objects, or concepts, but never both.

Frege's philosophical logic does not stop with the distinction between concepts and objects. It involves a whole hierarchy of logical categories. First, it should be noted that what have been called 'concepts' so far belong to the category of *first-level functions*. Whereas first-level functions map objects to objects in general, (first-level) concepts map objects to one of the truth-values – the True or the False. Thus, whereas the first-level function *the mother of* ξ maps X Æ A-XII to Grimes, and North West to Kim Kardashian, the first-level concept ξ *is the father of X Æ A-XII* maps Elon Musk to the True and Kanye West to the False. First-level functions that take objects as argument, however, are only the beginning. There are also second-level functions that take first-level functions as argument and map them to objects. If they map each of their arguments to either the True or the False, they are called second-level concepts. A crucial example is the second-level concept $(\forall x)(\varphi x)$. The judgment '$(\forall x)(x$ is a horse)', for instance, involves this second-level concept taking the first-level concept ξ is a horse as argument, which it maps to the False. An example of a first-level concept that it maps to the True, is $\xi = \xi$. This distinction between first-level functions and second-level functions is as absolute as the distinction between objects and first-level functions:

> Now just as functions are fundamentally different from objects, so also functions whose arguments are and must be functions are fundamentally different from functions whose arguments are objects and cannot be anything else (CP, 153; KS, 140; 26-27).

In addition to second-level functions that take first-level functions as argument, there are also third-level functions that take second-level functions as argument, and so on.

Another expansion is that there are not only first-level functions (or second-level functions) of one argument, but also first-level functions of two arguments – which may be called *relations* – first-level functions of three arguments, second-level functions of two

arguments, and so on. Again, the concomitant logical category distinctions are equally absolute:

> *Functions with two arguments* are just as fundamentally distinct from *functions with one argument* as the latter are from *objects* [Frege's emphases] (GGA, I, §21).

Finally, there are also what Frege calls *unequal-levelled functions*, which take multiple arguments from different categories (GGA, I, §22). An example would be $(\exists x)\big(\varphi(x,\xi)\big)$, which takes as arguments a first-level function of two arguments and an object. The judgment '$(\exists x)(x$ is the mother of X Æ A-XII)', for instance, involves this unequal-levelled concept taking the first-level relation ξ *is the mother of* ζ and the object X Æ A-XII as arguments, which it maps to the True.

In this way, the realm of *Bedeutungen* stratifies into a hierarchy of logical categories, each of which is *sui generis*: the logical categories are absolutely distinct in such a way that all forms of logical generality are limited to one logical category. There is no such thing as replacing an argument of a function with something that belongs to a different logical category. To use Frege's metaphor of places again: each logical place is intrinsically wedded to a single logical category, so that only what belongs to that category can stand in that place.

Let us now bring out the predicament arising from Frege's conception of the logical categories, focusing on the concept/object distinction again.[2] Against Frege, one could take that distinction to be relative to how concepts figure in judgments. For instance, one could think that the statements 'Seabiscuit is a horse' and 'The concept *horse* falls under the concept *animal*' involve the concept *horse* in two ways: first as unsaturated and taking an object as argument, second as complete and serving as argument of a first-level relation. On such an approach, it is the same concept that is at issue in both judgments. This is, as we shall see, how Russell thought about the matter. For Frege, however, such a view betrays confusion. It reveals that one has not properly grasped the absoluteness of the concept/object distinction. In fact, Frege does not take the second sentence to be about the concept ξ *is a horse* at all! As he puts it in what has come to be one of the most

[2] What I call *Frege's predicament* may be compared to what Sheffer has called the *logocentric predicament*, as it was taken up by Ricketts in his discussions of Frege and the *Tractatus*. Here is Sheffer's original formulation: "In order to give an account of logic, we must presuppose and employ logic" (Sheffer, 1926, 228). Here is one of the questions that Ricketts takes to arise from this predicament: "From what vantage point are we able to isolate and elucidate the basic notions that enter into the formulation of logic?" (Ricketts, 1985, 3).

notorious statements in the history of analytical philosophy: "the concept *horse* is not a concept" (CP, 186).[3]

It is at this point that Frege's predicament starts to present itself. If the concept *horse* is not a concept, it seems that Frege has to say the same about anything designated by an expression of the form 'the concept *X*'. But this seems to commit Frege to the conclusion that concepts are not concepts! Surely, it will be thought, this leaves him without much of a coherent view at all. Likewise, Frege seems forced to conclude that the concept *horse* is not unsaturated. But we saw that it is a cornerstone of his philosophical logic that concepts are unsaturated, even essentially so. Thus, it turns out that statements such as 'Concepts are unsaturated' cannot really succeed in expressing what they purport to express. Similarly, consider the statement: 'In some logical places only objects can stand and not concepts, in others only concepts and not objects'. It would seem that, for such a statement to express what it purports to express, it has to do exactly what Frege claims cannot be done: generalize over both objects and concepts. Frege wishes to simultaneously affirm *of* objects and deny *of* concepts that they can stand in certain logical places, or *vice versa*. But this means that Frege needs a form of logical generality – the generality of the notion of being able to stand in some logical place – that encompasses both objects and concepts.

Here is a way in which Max Black brings out Frege's predicament:

> [Frege's] contention that all functions are incomplete would lead one to suppose that he was making each of the following claims: that the sine function is incomplete, the property of solubility in water is incomplete, the relation of parenthood is incomplete. In fact, however, Frege intended not a single one of these assertions to follow from his claim that functions are incomplete. At first sight, this seems to be inconsistent with the customary conventions for the meaning of a sentence of the form 'All A are B'; and even the most sympathetic student of Frege's work may be hard put to it to explain Frege's contention (Black, 1968, 223-224).

My aim in this dissertation is to take up the role of such a sympathetic student, and to try to explain Frege's contention in a way that does not succumb to this predicament. Frege's

[3] We will have occasion to discuss this statement and its surrounding dialectic in detail in Chapter 2. Here, one may already sense a point that will become central in that discussion, namely that a proper understanding of what the statement 'The concept *horse* falls under the concept animal' seeks to express takes it to involve the first-level concept *ξ is a horse* serving as an argument – not of a first-level relation – but of a second-level relation that takes first-level concepts as arguments.

predicament can be formulated more generally as follows: it seems that Frege's conception of the logical categories is such that the sort of statements that we come out with in trying to give expression to that conception do not succeed in giving expression to it at all. The way in which those statements seem to undermine themselves suggests that it is only possible to provide a coherent statement of Frege's conception on the supposition that it is false, i.e. on the supposition that there are no absolute logical category distinctions of the kind Frege has in mind. And this seems to render inevitable the conclusion that Frege's conception of the logical categories is incoherent and must be relinquished.

To avoid this conclusion, we need an account of how statements such as 'Concepts are unsaturated' are supposed to fulfil their logico-philosophical function of expressing Frege's conception of the logical categories, without being self-undermining in the ways articulated above. Minimally, this means that such statements cannot be taken at face-value as the straightforward assertions that they seem to be. When Frege says that concepts are unsaturated, he cannot be straightforwardly asserting a subordination of a first-level notion of 'concepts' under a first-level notion of 'unsaturatedness' – as he *would* be in saying, for instance, that pineapple tastes good on pizza – since such a construal yields the sort of paradoxical results presented above.

If this negative observation is where Frege's *oeuvre* left us, there would be little more to do but to conclude that he had not seen the predicament which he was in. In that case, the task I have set myself in this dissertation – the task of developing a properly Fregean conception of the logical categories that overcomes the predicament – would be the task of trying to salvage the logico-philosophical wreck that Frege had bequeathed to us.[4] I wish to argue, however, that this is not where Frege's *oeuvre* left us. Frege was not blind to his predicament, but gave voice to it on multiple occasions. Here are two examples:

> It is not possible to give a definition of what a function is, because we have here to do with something simple and unanalysable. It is only possible to hint at what is meant and to make it clearer by relating it to what is known. Instead of a definition we must provide illustrations [*Erläuterung*]; here of course we must count on a meeting of minds [*entgegenkommende Verständnis*] (PW, 235; NS, 254).

[4] There are readers of Frege who think that this is basically our position. (Proops, 2013) is a good example, as we will see in Chapter 1. (Moore, 2012) also, I think, agrees with this sentiment.

> The peculiarity of functional signs, which we here call 'unsaturatedness', naturally has something answering to it in the functions themselves. They too may be called 'unsaturated', and in this way we mark them out as fundamentally different from numbers. Of course this is no definition; but likewise none is here possible. I must confine myself to hinting [*hinzuweisen*] at what I have in mind by means of a metaphorical [*bildlich*] expression, and here I rely on my reader's agreeing to meet me half-way [*entgegenkommende Verständnis*] (CP, 292; KS, 279-280; 665).

I am acutely aware that many Frege scholars will immediately respond that passages such as these cannot be read in the way that I am – at this point merely inchoately – suggesting that they can. For now, however, all I wish to emphasize is the minimal point that Frege did mark off his discourse about the logical categories – discourse in which terms such as 'function' and 'unsaturatedness' play a fundamental role – as having a distinct logico-philosophical status, a status that he characterizes by using such terms as *elucidation* (*Erläuterung*), *metaphorical* (*bildlich*), *hinting* (*hinzuweisen*), and *meeting of minds* (*entgegenkommende Verständnis*).[5] At the very least, this reveals that Frege was aware of the need for a distinction here. In this dissertation, I will bring together several strands from Frege's *oeuvre* in an attempt to bring out how he can fruitfully be read as having been very much alive to his predicament, to an extent that has not been fully appreciated in the literature. In line with a certain tradition of Frege scholarship in which I wish to position myself, I will henceforth use the term *elucidation* to characterize Frege's statements through which he seeks to bring to expression his conception of the logical categories. Although the term is, I think, apt, one could certainly use another one, such as *hints*, for instance. Whatever term one chooses will have advantages and disadvantages. One of the greatest disadvantages of using the term 'elucidation' will be discussed in the first chapter.

That Frege marks off his logico-philosophical discourse about the logical categories in this way, reveals that he seeks to distinguish it from straightforward assertions. As will become clear, elucidations do not express judgments, and the activity of elucidating is not an activity of judging. There is a reading of Frege's conception of the logical categories that is prevalent in the literature and that attempts to do justice to these observations. It is a reading that is nevertheless unsatisfactory. To improve upon that reading in a

[5] Diamond talks, in this regard, about Frege *placing* his remarks (Diamond, 1984, 354).

systematic way, is one core aim of this dissertation. On the reading in question, the recognition that the logical categories do not constitute a subject matter for judgment leads to the idea that what elucidations such as 'Functions are fundamentally different from objects' and 'Functions are unsaturated' seek to articulate, are what may be called *ineffable truths* about the logical categories. Because such truths are ineffable, however, those elucidations do not actually succeed in stating those truths – in the way in which judgments such as 'Grimes is the mother of X Æ A-XII' succeed in stating truths.[6] Understood in this way, elucidations constitute failed attempts to say what cannot be said, and as such they are logically defective. Nevertheless, they are construed as somehow succeeding in conveying these ineffable truths after all – not by stating them – but by ineffably gesturing at them *in and through* their failure to state them. I will call such a view *naïve ineffabilism*. My reasons for the qualification 'naïve' will become apparent. For now, let me just say this: insofar as the term 'ineffabilism' is used to signal nothing more than the observation that the logical categories do not constitute a subject matter for judgment, Frege's conception of the logical categories is indeed ineffabilist. The whole question, however, is how to proceed from there.

In its attempt to address that question, this dissertation will proceed as follows. The first chapter seeks to achieve more clarity about how to position the activity of elucidating the logical categories in Frege's *oeuvre*. There is a difficulty here that arises from the fact that there are, in fact, multiple notions of elucidation at work in Frege, which must be clearly distinguished. There is what I will call *the elucidation of primitive scientific terms*, which is concerned with scientific terms that stand at the beginning of the chain of definitions of a scientific system. Because these cannot be defined, Frege says that they must be elucidated. The aim of such elucidations is to secure agreement in *Bedeutung* among scientists who use the primitive terms in question. As Frege uses the logical category terms, however, they cannot be construed as having *Bedeutung* at all – as I will argue in more detail – so that the aim of elucidating the logical categories cannot be to secure such agreement in *Bedeutung*. This reveals that the logical categories, contrary to primitive scientific notions, are not a subject matter for judgment at all.

[6] In Frege's terminology: elucidations have no *Bedeutung*, they do not refer to a truth-value in the way that expressions of judgment do. Note that, in putting the point like that, I am abstracting from the fact that, in order to arrive at an expression that refers to a truth-value, we must leave out the judgment-stroke from an expression of judgment (CP, 149n; KS, 137n; 22n).

Still, there is a way of taking up this point that nevertheless makes the difference between primitive scientific notions and the logical categories too slight,[7] and this is what happens in naïve ineffabilist readings. Although they pay lip-service to the point that the logical categories do not constitute a subject matter for judgment, the logical categories are still conceived in such a way that they are very much like a subject matter for judgment. Only, rather than straightforwardly saying that elucidations of the logical categories express truths about the logical categories, we must say that they constitute failed attempts to state what, strictly speaking, cannot be said, and that it is *in and through* their failure that they nevertheless succeed in conveying what cannot be said. In this way, such failed attempts at saying what cannot be said turn out to behave exactly as they would behave if they were, after all, successful in saying it, securing agreement in a way that turns out to be remarkably similar to that of elucidations of primitive scientific terms. Against this, I will argue that to properly account for the difference between primitive scientific notions and the logical categories – and thereby between judgment and elucidation – is to give up on the idea that the elucidation of the latter must be conceived as proceeding through failed attempts to mimic the sort of thing that elucidations of primitive scientific terms do. The elucidation of the logical categories constitutes a *sui generis* activity that is not to be modelled on judgment.

The aim of the second chapter is to develop an initial understanding of the logical stratification of the realm of *Bedeutungen* into logical categories. Central to this chapter is the claim that a certain distinction in Frege has not been well understood, namely the distinction between the notion of making judgments about particular concepts and the notion of making judgments about the logical category of (first-level) concepts. From what has already been said – and will be discussed in more detail in the first chapter – it is clear that Frege regards the latter notion as inherently confused: this is the minimal sense in which Frege's conception of the logical categories is indeed ineffabilist. Many readers, however, believe that Frege regarded the former idea as equally confused. It will be argued that this constitutes a fateful mistake that renders impossible a proper understanding of Frege's conception of the logical categories. Rather than regarding the notion of making judgments about particular concepts as incoherent, Frege was at pains to *vindicate* that notion in his philosophical logic, through the *Begriffsschrift* device of

[7] I take this notion of making differences too slight from (Diamond, 2010, 555), who traces it to Frege.

higher-level subsumptions. This, I will show, is what was at issue in Frege's dialectic with Kerry and in the notorious grain of salt passage, which has been widely misunderstood as espousing a version of naïve ineffabilism.

I will show how Frege's vindication of the possibility of making judgments about concepts involves the observation that the logical stratification of the realm of *Bedeutungen* comes with the logical stratification of a cluster of fundamental logico-philosophical phenomena, such as logical subjecthood, aboutness, *Bedeutung*, and so forth. This gives rise to a corresponding cluster of logico-philosophical terms of art each of which has what I will call a *logically stratified elucidatory use*. Rather than regarding such elucidatory use as logically confused because it seems to transgress the logical category distinctions – as many readers of Frege have done – we must regard it as essential to Frege's philosophical logic. Properly recognizing how second-level subsumptions are *about* first-level functions just as much as first-level subsumptions are *about* objects, just is to recognize the logically stratified function/argument nexus that lies at the core of Frege's philosophical logic. It gives rise to what I will call the *functionality* of the realm of *Bedeutungen*, which characterizes the logical stratification of the realm of *Bedeutungen*, and gives us the logico-philosophical form of unity of the realm of *Bedeutungen*. Its form of unity *is* its functionality, which is nothing else than its logical stratification.

In Chapter 3, we will examine the dialectic between Frege and Russell on Frege's function/object distinction. Russell objected that elucidations such as 'A function can never take the place of an object' are self-undermining, since 'a can never take the place of an object' puts a in the place of an object. In some ways, Russell's worries are similar to Kerry's, since Russell also questions whether Frege's function/object distinction is compatible with the possibility of making judgments about functions. In another way, however, Russell's worries cut deeper than Kerry's. Unlike Kerry, Russell takes aim at Frege's elucidations themselves, thereby confronting Frege with the exact predicament articulated above. This means that, in replying to Russell's objection, Frege has to account for the logico-philosophical status of his elucidations themselves. This results in a fascinating stretch of text that constitutes one of the finest examples of Frege attempting to elucidate his conception of the logical categories. We will see how Frege – in addition to trying to show to Russell how his device of higher-level subsumptions vindicates the possibility of making judgments about functions – tries to elucidate to Russell the very distinction between judgment and elucidation itself. This results in Frege providing us with crucial insights into his conception of the logical categories and their elucidation.

Rather than addressing the statement 'A function can never take the place of an object' directly, Frege turns to the statement 'A function name can never take the place of a proper name'. In this way, Frege makes it clear that a proper understanding of the logical stratification of the realm of *Bedeutungen* essentially involves a turn to language, i.e. an understanding of the logical stratification of the *Begriffsschrift*. If, however, this turn to language is to have the elucidatory function that Frege ascribes to it, the logical stratification of the *Begrifsschrift* and the logical stratification of the realm of *Bedeutungen* must be related in the right way. This yields two main tasks for the remainder of the dissertation. The first task is to develop an account of the logical stratification of the *Begriffsschrift*. The second task is to develop an account of the requisite logico-philosophcial relation that obtains between the respective logical stratifications of the *Begriffsschrift* and the realm of *Bedeutungen*. Each of these, it will be shown, revolves around a form of *internality*.

To carry out the first task, we must come to an understanding of Frege's conception of the sign, a neglected topic in the literature. In the fourth chapter, I will argue that Frege's conception of the sign is incompatible with the contemporary syntax/semantics distinction. Conceiving of the sign in terms of that distinction always involves a form of externality between what is used in the expression of judgment, on the one hand, and its being used in the expression of judgment, on the other hand. On Frege's conception of the sign, however, its being used in the expression of judgment belongs to the very nature of what is so used. I will call this *the internality of its use to the sign*. To do justice to that internality, one must leave behind the contemporary syntax/semantics distinction as providing a proper framework for thinking about the sign.

Once that framework has been left behind, we are in a better position to understand Frege's conception of the logical stratification of the *Begriffsschrift*. It will be seen how Frege characterizes that stratification in terms of a relation of *fit* that obtains between signs. Proper names, for instance, are said to *fit* the argument places of first-level functions of one argument. In accordance with the internality of its use to the sign, it is crucial to understand this relation of fit as obtaining between signs as they are used in the expression of judgment. For the proper name to fit the argument place of a first-level function name, *is* for it to be used in the expression of judgment. The logical stratification of the *Begriffsschrift*, then, consists in the manifold ways in which *Begriffsschrift* judgments ramify into signs that fit together in the expression of judgment. This gives rise to what I will call the *compositionality* of the *Begriffsschrift*, which characterizes the logical

stratification of the *Begriffsschrift*. This compositionality, moreover, reveals how the categories of *Begriffsschrift* signs are not logically self-standing, but intrinsically intertwined. To understand the compositional nature of a proper name *is* to understand how it fits together with other signs in the expression of judgment. In this way, the compositionality of the *Begriffsschrift* gives us the form of logico-philosophical unity of the *Begriffsschrift*. Its compositionality *is* its unity, which is nothing else than its logical stratification.

The fifth and final chapter proceeds to develop an account of the logico-philosophical relation that obtains between the respective logical stratifications of the *Begriffsschrift* and the realm of *Bedeutungen* – between compositionality and functionality – a relation that must vindicate Frege's turn to language as essential to the elucidation of the logical categories. Frege himself brings out that relation by employing an elucidatory notion of *correspondence*. Through a rejection of what I will call a *realist* and an *idealist* construal of that correspondence, I will show how properly taking into account Frege's commitment to the internality of its use to the sign gives rise to a conception of this correspondence as itself constituting a form of internality of the logical stratification of the *Begriffsschrift* and the logical stratification of the realm of *Bedeutungen.* Properly understood, Frege's elucidatory turn to language does not consist in bringing in an order of the sign that is external to the realm of *Bedeutungen*, but rather consists in unfolding the internality of the respective logical stratifications of the realm of *Bedeutungen* and the *Begriffsschrift*.

One important way in which Frege gives voice to this internality, is through his use of a single notion of fit to characterize both the compositionality of the *Begriffsschrift* – as discussed above – and the functionality of the realm of *Bedeutungen*. Proper names *fit* first-level function names, and objects *fit* first-level functions. That we have here one notion of fit, earmarks the internality of the *Begriffsschrift* and the realm of *Bedeutungen*. At the same time, I will argue that it is crucial to realize that the relations of fit of *Begriffsschrift* signs and the relations of fit of *Bedeutungen* are not the same – as some readers of Frege have thought – but heterogeneous. Compositionality is not functionality. It is only through what I will call their *internal heterogeneity that* the phenomena of compositionality and functionality can give us the unity of both the *Begriffsschrift* and the realm of *Bedeutungen*. I will end by suggesting how this gives rise to a Fregean notion of showing, according to which it can be said that the logical stratification of the realm of *Bedeutungen* shows itself in the logical stratification of the *Begriffsschrift*, in such a way that what shows itself is indeed internally related to that in which it shows itself. In the end, what Frege's

elucidations of the logical categories elucidate, is nothing else than this logico-philosophical phenomenon of showing itself, which proceeds *in and through* the internal heterogeneity of the compositionality of the *Begriffsschrift* and the functionality of the realm of *Bedeutungen*.

This dissertation is an exercise in the history of philosophy. What does one seek to do in doing history of philosophy? There is, of course, no single answer to be given to this question. I can only say what I take myself to have been aiming to do in this dissertation. Conant describes Frege as "working within a conception of logic that it is exceedingly difficult for someone intellectually at home in the assumptions of contemporary philosophy to so much as get into view" (Conant, 2020, 340). My aim, simply put, has been to get that conception into view. Not, it will become clear, in all its details, but at least in what may be called its general outline, in a way that can serve as a fruitful basis for further research.

There is some sort of paradox internal to this endeavour, if one accepts as I do Conant's claim that Frege's conception of logic is intellectually alien to us in profound ways. The paradox is this: the very exercise of bringing that conception into view, seems to undermine the claim that it is alien. That is to say: one must bring into view that which is alien *as* alien, i.e. in such a way that the act of bringing it into view does not relieve it of its alienness. This is a difficult exercise. Whether I have succeeded may be gathered, I think, from the extent to which this dissertation arouses in its reader a sense of awe at the vastness and philosophical depth of what remains to be excavated from Frege. It is a sense of awe that I take to be characteristic of history of philosophy done well, and the relevant sort of vastness and philosophical depth that occasion it can only be found in what is indeed intellectually alien – that is to say, in that which is such that we genuinely stand to *learn* from it.

I end this introduction with some practical remarks. First, I will not use Frege's own *Begriffsschrift* notation, but will present all formulas in contemporary notation. My motivation for this is practical, not philosophical. Indeed, there is philosophical loss involved in this procedure, since my notation will not feature anything that corresponds to Frege's horizontal or his judgment-stroke. As I will flag at certain points, this dissertation does not take into account the additional philosophical complexities that arise from a proper consideration of Frege's way of using these signs. I consider this a task for further research.

Second, I will often leave Frege's terms '*Bedeutung*' and '*Sinn*' untranslated. A first reason for doing this is to ensure that there can be no confusion about what is meant. A second reason is that leaving them untranslated helps us to remain conscious of the fact that these are crucial terms of art in Frege's philosophical logic that come with far-reaching philosophical commitments, as will become clear throughout this dissertation.

Finally, I wish to note that – in developing what I take to be Frege's considered conception of the logical categories – I will not be concerned with the historical evolution of that conception, which would constitute a topic for further research. Although tracing the precise trajectory of the evolution of Frege's views is a delicate matter, there can be no doubt that – when he published the *Begriffsschrift* in 1879 – he had not yet arrived at his mature conception of the logical categories, however one construes that conception. For the purposes of this dissertation, I take his mature conception of the logical categories to be the one that underlies the *Grundgesetze*. In what follows, however, I will be using passages from all stages of Frege's development to support my account. This means that I will be using certain passages to articulate a view about which it could be debated whether Frege already held it – or in some cases, perhaps, whether he still held it[8] – when he wrote the passages in question. These are historical intricacies that I cannot investigate within the scope of this dissertation. I hope that I will be excused for granting myself the liberty of blankly stating that I do not think that such an investigation would undermine my account in substantive ways. At the same time, I would welcome, of course, an investigation of how a detailed construal of the evolution of his views stands to impact my reading of Frege, as I would welcome any kind of critical[9] discussion of the account that I am about to present.

[8] For an overview of some of the far-reaching ways in which Frege revised his views towards the end of his career, see (Narboux & Perrin, 2021).
[9] Or non-critical.

Chapter 1
The many senses of 'elucidation'

In the introduction, I presented *Frege's predicament*, consisting in the fact that Frege's conception of the logical categories seems to be such that the very statements that we come out with in seeking to articulate that conception turn out to be self-undermining. I pointed out that Frege's marking off his logico-philosophical discourse about the logical categories as consisting in *elucidations* shows that he was aware of this predicament. The aim of this first chapter is to gain some initial clarity about the difference between elucidation and judgment, as it figures in Frege's *oeuvre*.

Doing so requires properly taking into account the fact that the logical categories are not the only phenomenon that Frege characterizes as a subject matter for elucidation. In fact, Frege's most sustained discussions of elucidation do not concern the logical categories at all, but instead primitive scientific terms that stand at the beginning of the chain of definitions of a scientific system. Most of this chapter will be concerned with bringing out the depth of the difference between the elucidation of such primitive scientific terms, on the one hand, and the elucidation of the logical categories, on the other hand. It is a mistake to think that there is a single notion of elucidation to be found in Frege's *oeuvre* that is meant to capture all cases.

In the Introduction, I flagged that there is a disadvantage to choosing the term 'elucidation' as a characterization of Frege's logico-philosophical discourse about the logical categories. We now see what this disadvantage is: it engenders a confusion between the different notions of elucidation that figure in Frege's thought. Such confusion is further facilitated by the fact that there are purely negative characterizations of elucidation that seem to capture all cases. Here is one such negative characterization: elucidation is what is required to secure mutual agreement in those

cases where definition is unavailable because a certain phenomenon is logically primitive. Because of the availability of such a negative characterization – and because Frege is not as clear on these matters as one could hope[1] – interpreters have tended to underestimate or even completely miss the extent to which it obscures the many senses of 'elucidation' that are at play in Frege's *oeuvre*. There is not one sense in which the different topics for elucidation are such that definition is 'unavailable' for them, not one sense in which the relevant phenomena are 'primitive', and not one sense in which 'mutual agreement' is to be secured. Concomitantly, there is also not one overarching notion of elucidation that governs all cases.

A failure to see this leads one to think that Frege's remarks about elucidations of primitive scientific terms equally apply to elucidations of the logical categories. Because the aim of elucidations of primitive scientific terms is to secure mutual agreement in the *Bedeutung* of the terms in question, one will equally conceive of elucidations of the logical categories on the model of securing agreement in *Bedeutung*. This results in the sort of naïve ineffabilist reading discussed in the introduction, and that will figure as a main target in this chapter. Because the logical categories cannot be the *Bedeutung* of any term, their elucidation is now conceived as consisting in gesturing at ineffable truths that cannot properly be expressed in judgment, a gesturing that proceeds *through* failed attempts to say what cannot be said, thereby securing something that may be described as agreement in *ineffabilia*. As long as one conceives of elucidations of the logical categories on the model of elucidations of primitive scientific terms – and thereby on the model of judgment – such a naïve ineffabilist reading is all but inevitable. It can only be overcome by properly seeing the elucidation of the logical categories as a *sui generis* activity that is not to be modelled on judgment.

[1] Though I have no wish to incur the reproach of picking petty quarrels with a genius to whom we must all look up with grateful awe.

1.1 Elucidation of primitive scientific terms

Let us start by investigating Frege's conception of the elucidation of primitive scientific terms, a conception that is articulated most explicitly at three places in his *oeuvre*: a letter to Hilbert (PMC, 36-37; BW, 63), the posthumous text *Logic in Mathematics* (PW, 207; NS, 224), and Part II of the published paper *Foundations of Geometry* (CP, 300-302; KS, 288-289; 301-303).

On Frege's conception of science, it is essential to the scientific endeavour that there is, at all times, mutual agreement in the *Bedeutung* of scientific terms.[2] A situation in which, for instance, one group of biologists uses the term 'gene' to designate something else than another group of biologists, is intolerable and must be remedied. Either these groups must use different terms to designate their different *Bedeutungen*, or they must come to use the term 'gene' with the same *Bedeutung*.[3] It is because Frege takes mutual agreement in *Bedeutung* to be crucial, that he emphasizes as one of the advantages of his *Begriffsschrift* that it allows us to secure such agreement through the setting up of rigorously formulated definitions. Any novel term that is introduced into a scientific system must be explicitly defined using only terms for which the necessary agreement has already been secured, which in turn guarantees mutual agreement in the *Bedeutung* of the novel term.

This process bottoms out in terms that are used to define other scientific terms, but which are not themselves provided with a definition. It is such terms that are at issue in Frege's discussions of the elucidation of primitive scientific terms. When I talk about *primitive scientific terms* in what follows without further qualification, I always intend this specific kind of terms, i.e. scientific terms that stand at the beginning of the chain of definitions of a scientific system. One main burden of this chapter is to show that, although Frege also calls the logical category terms primitive, they are not primitive scientific terms in the sense just specified.

---

[2] I use the locution 'agreement in' rather than 'agreement about', since what is fundamental is that scientists use their terms *with* the same *Bedeutung*, in relation to which the fact that they agree *about* this *Bedeutung* is a secondary phenomenon.

[3] There is an additional question whether they must also come to mutual agreement in the *Sinn* of the term in question. Frege's remark that elucidations serve to "make sure that all who use them henceforth also associate the same sense [*Sinn*] with the elucidated word" (CP, 301) gives an affirmative answer. Still, it is not entirely clear how elucidations are to succeed in securing such agreement in *Sinn*. For my purposes, this issue can be left aside, and I will focus on mutual agreement in *Bedeutung*.

Here is Frege:

> My opinion is this: We must admit logically primitive elements that are indefinable. Even here there seems to be a need to make sure that we designate the same thing by the same sign (word). Once the investigators have come to an understanding about the primitive elements and their designations, agreement about what is logically composite is easily reached by means of definition. Since definitions are not possible for primitive elements, something else must enter in. I call it elucidation (CP, 300; KS, 288; 301).

Two examples of such logically primitive terms in Frege's *Grundgesetze* system are the identity sign and the second-level function-name '$\grave{\varepsilon}\varphi(\varepsilon)$' standing for the second-level function that maps first-level functions of one argument to their course-of-values. These are employed in defining other *Begriffsschrift* terms – such as, for instance, Frege's version of set membership '$\xi \cap \zeta$' (GGA, I, §34) – but are not themselves defined in Frege's system. Examples of logically primitive terms that could figure in other scientific systems may include 'point' in geometry, 'force' in physics, 'life' in biology, 'morpheme' in linguistics, etc.

So far, I have mainly talked about scientific *terms* being logically primitive because they stand at the beginning of the chain of definitions of a scientific system. This naturally raises the question whether this sense of a term's being logically primitive is parasitic on a more fundamental sense in which the *Bedeutung* of such terms is itself logically primitive. Such a logically primitive *Bedeutung*, it seems, would have to be such that it can only be the *Bedeutung* of a primitive scientific term, i.e. a scientific term that stands at the beginning of the chain of definitions of a scientific system. In this way, any term having a logically primitive *Bedeutung* will be a primitive scientific term. The other direction, however, is more complicated. We can, for instance, set up a system in which the term 'number' stands at the beginning of the chain of definitions. But this does not show that the concept *number* is itself logically primitive, since this does not show that the concept *number* can *only* be the *Bedeutung* of a term that stands at the beginning of the chain of definitions of the scientific system in which it figures. Indeed, it is a central claim of Frege's logicism that the concept *number* can be defined. There are thus two senses in which a scientific term may be called logically primitive. It can be logically primitive in the sense of having a logically primitive *Bedeutung*, or it can be logically primitive in the sense of standing at the beginning of the chain of definitions of a particular scientific system. The former implies the latter, but not conversely. This raises the question: is

Frege's conception of the elucidation of primitive scientific terms meant to apply to terms that are logically primitive in the weaker or the stronger sense? Frege himself is unclear on the matter. Usually, he talks in terms of not being defined, which suggests the weaker sense. In his letter to Hilbert, however, he does talk about a term's *Bedeutung* being logically primitive (PMC, 37), which suggests the stronger sense. Note that some of Frege's own primitive terms from the *Grundgesetze* are only primitive in the weaker sense, due to the interdefinability of the sentential logical operators and the quantifiers. Here, however, we need not worry about this issue, nor need we go into the question of how exactly to understand the notion of a *Bedeutung*'s being logically primitive. When I talk about 'primitive scientific *Bedeutungen*' in what follows, this can always be understood as '*Bedeutungen* of primitive scientific terms'.

The aim of the elucidation of primitive scientific terms is to secure the mutual agreement in *Bedeutung* that is so important to Frege, but which cannot here be secured through definitions. It proceeds in the first place by presenting statements in which the terms in question occur,[4] in the hope that they will help the reader latch onto the *Bedeutung* of the primitive terms in question. Here is how Frege continues after the passage above:

> It is this, therefore, that serves the purpose of mutual understanding among investigators, as well as of the communication of the science to others. We may relegate it to a propaedeutic. It has no place in the system of science; in the latter, no conclusions are based on it. Someone who pursued research only by himself would not need it. The purpose of elucidations is a pragmatic one; and once it is achieved, we must be satisfied with them (CP, 300-301; KS, 288; 301).

Once mutual agreement is secured, elucidation has achieved its purpose and the actual business of science can get underway. Like any other scientific term, primitive scientific terms possess *Bedeutung*. Indeed, if this were not the case, they would not belong in a scientific system at all. Once we have latched onto that *Bedeutung*, they can safely be used just like any other scientific term in further definitions and judgments. That is why their elucidation belongs to a *propaedeutic*: it is a *propadeutic* to their subsequent scientific use, which is logically on a par with the scientific use of any other scientific term, logically

[4] "Elucidations will generally be propositions that contain the expression in question, perhaps even several such expressions" (CP, 301; KS, 288; 302). I say 'in the first place', since it cannot be excluded that further statements – in which the term to be elucidated does not itself appear – may also fulfil an elucidatory function.

primitive or not. The only difference is that mutual agreement in *Bedeutung* is secured through elucidations instead of definitions.[5]

Whether elucidations succeed in helping readers latch onto the *Bedeutung* of a primitive scientific term, Frege regards as a contingent matter:

> And here we must be able to count on a little goodwill and cooperative understanding, even guessing; for frequently we cannot do without a figurative mode of expression. But for all that, we can demand from the originator of an elucidation that he himself know for certain what he means; that he remain in agreement with himself; and that he be ready to complete and emend his elucidation whenever, given even the best of intentions, the possibility of a misunderstanding arises. Since mutual cooperation in a science is impossible without mutual understanding of the investigators, we must have confidence that such an understanding can be reached through elucidation, although theoretically the contrary is not excluded (CP, 301; KS, 288; 301).

It is precisely in this way, moreover – i.e. in the contingency of its success – that elucidations of primitive scientific terms differs from definitions:[6]

> Let us turn to proper definitions! They, too, serve mutual understanding, but they achieve it in a much more perfect manner than the elucidations in that they leave nothing to guess-work; nor need they count on co-operative understanding and goodwill (CP, 301-302; KS, 289; 302).

[5] I do not wish to suggest that there are no important differences among primitive scientific terms. For instance, it may be suspected that a primitive *logical* term (e.g. the identity sign) requires a different treatment than, for instance, a primitive *physical* term (e.g. 'force'), and that there will be related differences between their respective elucidations. One salient example is Frege's horizontal: part of the reason why it cannot be defined, is that no *definiens* could replace it as what is to be attached to the judgment-stroke in the expression of a judgment. We have here what could be called a *notational constraint* that is absent from other cases of primitive scientific terms. This shows that the sense in which the horizontal is logically primitive cannot be understood solely in terms of its standing at the beginning of the chain of definitions. For this reason, it may be unwise to call it a 'primitive scientific term' at all. Like the judgment-stroke, the horizontal is *sui generis* in Frege's philosophical logic, and requires its own separate investigation. Similar remarks may obtain for the conditional stroke, whose verticality on the page – yielding the characteristic two-dimensionality of Frege's *Begriffsschrift* – may also constitute such a notational constraint. Note that, in the *Tractatus*, Wittgenstein will come to regard as deeply confused the Fregean aim of securing mutual agreement in *Bedeutung* for primitive *logical* terms, in a way that makes them importantly different from primitive terms from other sciences. See e.g. (TLP, 4.0312).

[6] A sidenote I cannot resist making: if one wanted, one could read much of Wittgenstein's later philosophy as a sustained critique of Frege's conception of the contingency of the success of elucidations of primitive scientific terms. I hope to develop this thought in future work.

Much remains to be said about Frege's conception of the elucidation of primitive scientific terms, which could be a topic for another dissertation. My cursory discussion suffices for our purposes, however, since the aim of this chapter is to bring out how the elucidation of the logical categories cannot be conceived on this model of securing mutual agreement in *Bedeutung*. Whatever the more fine-grained details of Frege's conception of the elucidation of primitive scientific terms turn out to be, this core difference will remain intact.

1.2 Logical category terms are not primitive scientific terms

Above, I presented the following general characterization of elucidation: Elucidation is what is required to secure mutual agreement in those cases where definition is unavailable because a certain phenomenon is logically primitive. We have now seen one way to understand this formula, namely with regards to primitive scientific terms. In this case, the relevant notion of 'being logically primitive' is that of standing at the beginning of the chain of definitions of a scientific system. Definition is unavailable because what stands at the beginning of the chain of definitions of a scientific system must *ipso facto* be undefined. Securing mutual agreement means securing agreement in *Bedeutung*, by helping readers latch onto the *Bedeutung* of the primitive term in question. This is done by providing suitably constructed statements containing the primitive scientific terms in question.

Nothing I have said so far excludes that logical category terms are primitive scientific terms, so that the elucidation of the logical categories is simply a case of the elucidation of primitive scientific terms. What we will come to see, however, is that *none* of the above characteristics of the elucidation of primitive scientific terms apply to the elucidation of the logical categories. Although we can say that, here as well, elucidation is required to secure mutual agreement because definition is unavailable due to a certain phenomenon being logically primitive, what has thereby been said in using this form of words shifts completely.

Let us start from the observation that Frege himself does indeed call the logical categories primitive:

> It is not possible to give a definition of what a function is, because we have here to do with something simple and unanalysable. It is only possible to hint at what is meant and to make it clearer by relating it to what is known. Instead of a definition we must provide elucidations; here of course we must count on a meeting of minds [Frege's emphasis] (PW, 235; NS, 254).

> When we have thus admitted objects without restriction as arguments and values of functions, the question arises what it is that we are here calling an 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 any empty place (CP, 147; KS, 134; 18).

If logical category terms were primitive in the same sense as primitive scientific terms, logical category terms should have logical categories as their *Bedeutung*. Let us investigate this idea for the logical category of objects. Such a primitive term would have to be a first-level concept subsuming all objects, e.g. '$O(\xi)$'. A first point to make is that '$O(\xi)$' would in fact not be logically primitive in Frege's system, but could be defined in many ways, for instance as '$O(\xi) = (\xi = \xi)$' or '$O(\xi) = (\forall\varphi)\big(\varphi(\xi) \vee \neg\varphi(\xi)\big)$'.[7] Thus, if '$O(\xi)$' had as its *Bedeutung* the logical category of objects, that category would not be logically primitive, as Frege claims it is.

More fundamentally, however, if the logical category of objects were logically primitive in this sense, then the aim of elucidating it would be to attain mutual agreement in the *Bedeutung* of '$O(\xi)$', on the model of the elucidation of primitive scientific terms. If we look at Frege's own elucidatory use of logical category terms, however, we see that it does not correspond to this model. This is because that use is *contrastive*: it concerns the sharp distinctions between the logical categories, which Frege often presents in terms of the distinction between what is unsaturated and what is complete. Here is one example:

> 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

[7] An anonymous referee of a paper in which I make the same point objected that such definitions would not capture the *Sinn* of the term 'object'. First, granting that there is such a determinate *Sinn* to be captured (which I do not think is as clear as this objection seems to suppose), on what basis could it be claimed that none of these definitions capture it? Second, why would the success of such a definition depend on its capturing the prior *Sinn* of the term 'object', and not just its *Bedeutung*?

> itself must be called incomplete, in need of supplementation, or 'unsaturated'. And in this respect functions differ fundamentally from [objects][8] (CP, 140; KS, 128; 6).

How would we use our primitive term '$O(\xi)$' to express, for instance, what Frege says in the final sentence? To bring out the fundamental difference between functions and objects, we want to say something like this: functions are unsaturated, whereas objects are not. Thus, we need to introduce a first-level predicate – say '$U(\xi)$' – to purportedly express that something is unsaturated. The claim that objects are not unsaturated then becomes: $(\forall x)\big(O(x) \supset \neg U(x)\big)$. At the same time, however, we want '$U(\xi)$' to express exactly what we affirm of functions when we say that they *are* unsaturated. But we cannot use '$U(\xi)$' to say anything about functions, since it only takes proper names as arguments. To say something about a (first-level) function, we need a second-level predicate, say '$\upsilon_\varepsilon\big(\varphi(\varepsilon)\big)$'. We can then try to express the claim that all functions are unsaturated as '$(\forall F)\left(\upsilon_\varepsilon\big(F(\varepsilon)\big)\right)$'. The result, however, is that we are using signs of different levels for 'unsaturated' in our statements. Thus, what we affirm of functions is not the same as what we deny of objects, since signs of different levels cannot have the same *Bedeutung*. We have thereby not succeeded in articulating a *difference* between functions and objects, something true of objects while false of functions.

It could be objected that we should use an unequal-levelled relation to express directly that functions differ fundamentally from objects, along the lines of '$(\forall x)(\forall F)\big(x \neq_\varepsilon F(\varepsilon)\big)$', in which the unequal-levelled relation sign '$\xi \neq_\varepsilon \varphi(\varepsilon)$' designates the requisite relation of being fundamentally different that obtains between objects and first-level functions of one argument. It is true, of course, that we can introduce an unequal-levelled relation $\xi \neq_\varepsilon \varphi(\varepsilon)$ which is such that '$(\forall x)(\forall F)\big(x \neq_\varepsilon F(\varepsilon)\big)$' turns out to be true. But this does not mean that '$(\forall x)(\forall F)\big(x \neq_\varepsilon F(\varepsilon)\big)$' gives us a statement of the fundamental difference between objects and first-level functions. Frege's commitment to the latter equally comes, for instance, with the commitment that objects are *not* fundamentally different from each other. But here we run into the same problem: we cannot express such a claim using the sign '$\xi \neq_\varepsilon \varphi(\varepsilon)$', since it is not of the right level.

[8] Frege has 'numbers' instead of 'objects' – due to the specific context – but it is clear that he would endorse the same statement about objects. Indeed, it is precisely the fact that numbers *are* objects that is operative here.

However one tries to devise *Begriffsschrift* terms to capture Frege's elucidatory statements about the logical categories, one will always run into the fact that such terms can never capture Frege's contrastive elucidatory use of the logical category terms in what we may call its full *elucidatory generality*, a form of generality that is essential to Frege's conception of the logical categories.[9] Contrary to elucidations of primitive scientific terms, then, Frege's elucidatory use of logical category terms cannot have the aim of securing agreement in *Bedeutung*. If we look at the Fregean realm of *Bedeutungen* and try to find a *Bedeutung* that fits Frege's contrastive elucidatory use of logical category terms, we find that there is no suitable *Bedeutung* for us to latch onto.[10] Frege himself never introduces a *Begriffsschrift* term that is supposed to stand for a logical category.[11] Given how centrally these figure in his accounts, this would be inexplicable if he actually believed that he could do so. Frege evidently believed that there is something essential about the logical categories that cannot be captured by *Begriffsschrift* terms such as '$O(\xi)$' and '$\xi \neq_{\varepsilon} \varphi(\varepsilon)$', and this lies in the sharp distinctions between the logical categories that he seeks to bring out through his contrastive elucidatory use of the logical category terms.

This shows that the sense in which the logical categories cannot be defined is entirely different from that of primitive scientific *Bedeutungen*. The latter *are* the *Bedeutung* of *Begriffsschrift* terms, only those terms stand at the beginning of the chain of definitions of a scientific system. Logical category terms, on the other hand, do not stand at the beginning of chains of definitions at all. The logical categories are not the *Bedeutung* of *any* term, and do not figure in scientific systems. In this way, the very idea of definition is completely out of place, since definitions serve to fix the *Bedeutung* of a term, and the logical categories do not even inhabit the realm of *Bedeutungen*.[12] We may define a

[9] In talking about an elucidatory form of generality, I am distinguishing it from logical forms of generality, which are limited to a single logical category. This notion of elucidatory generality anticipates what I will call Frege's *logically stratified elucidatory use* in Chapter 2.

[10] This has been observed many times (Wells, 1968, 400) (Hugly, 1973, 229) (Geach, 1976, 57-58) (Diamond, 1984, 357ff.) (Ricketts, 2010, 182) (Hale & Wright, 2012, 104) (Jolley, 2015, 8) (Jones, 2016, §5) (Conant, 2020, 450).

[11] This has been observed many times as well (Conant, 2002, 385) (Ricketts, 2010, 170, footnote 50) (Weiner, 2010, 59-60). There is one seeming exception in Frege's correspondence with Russell which we will discuss at length in Chapter 3.

[12] It could be objected that, insofar as logically primitive *Bedeutungen* cannot be defined, the very idea of definition is completely out of place with regards to them as well. One can certainly make sense of such a statement. But the sense in which the very idea of definition is out of place with regards to logically primitive *Bedeutungen* remains fundamentally different from the sense in which it is out of place with regards to the logical categories. Using the same words does not undo this difference.

predicate '$O(\xi)$' as '$O(\xi) = (\xi = \xi)$', but this does not give us a definition of a logical category. It simply gives us a definition of a first-level concept that subsumes all objects[13]. Another way to bring this out, is as follows. Suppose that we want to check whether something is an object. Can we use the purported definition '$O(\xi) = (\xi = \xi)$' to do so? No. It is rather the other way around: in order for us to be in a position to apply the definition at all, we must already know that we have on our hands an object, because only an object fits the argument place of the first-level function $\xi = \xi$.[14] The very idea that this formula could serve to identify objects shows that we have not properly grasped the notion of an object. Suppose I am in doubt whether the concept *horse* is an object or not? It makes no sense to then ask whether it is identical to itself or not to settle this question, since the meaningfulness of the question itself already presupposes that the concept *horse* is an object.[15]

Some readers of Frege seem to disagree and think that one can provide definitions of logical categories. Sullivan, for instance, seems to take '$(\forall x)(\varphi x = -\varphi x)$'[16] to provide a definition of the category of concepts (Sullivan, 2006, 102).[17] Similarly, (Dummett, 1981, 216-217) (Beaney, 1996, 200-201) (Noonan, 2006, 165) all seem to flirt with the idea that one can provide definitions of the logical categories. What can cause confusion here, is the fact that – in Frege's later *Grundgesetze* system – the earlier logical category of concepts has been superseded by the logical category of first-level functions of one argument, as described in *Funktion und Begriff*. In this way, Sullivan's formula indeed gives us a definition of first-level concepts, now conceived as a kind of first-level functions of

[13] See also (Conant, 2020, 839).
[14] Compare (Diamond, 1984, 354ff.).
[15] Can one then use the criterion whether 'The concept *horse* is self-identical' is meaningful? This question brings with it a host of philosophical subtleties that would take us too far afield here. For our current purposes, it is enough to observe that this criterion would not yield a *Begriffsschrift* term that has the logical category of concepts as its *Bedeutung*.
[16] The '—' is the content-stroke.
[17] I say 'seems' here, because I am not exactly sure what Sullivan intends this definition to do. He distinguishes between absolutely indefinable notions and notions that "happen to be chosen as primitives of, hence not capable of formal definition within, a particular formulation of logic" (Sullivan, 2006, 102). So far, it seems as if he is nicely capturing the distinction between logical categories and primitive scientific notions. But he then proceeds to say that the notion of a concept is "Frege's favourite example of something indefinable in the first, system-independent sense", on the one hand, while adding that it "*is* straightforwardly definable in the second sense" (Sullivan, 2006, 102), after which he proceeds to give the quoted definition. I find it difficult to make sense of what Sullivan says here. What is something that is at once 'absolutely indefinable' *and* 'straightforwardly definable in a particular formulation of logic'? Does Sullivan's definition give us a definition of what is absolutely indefinable? I assume it does not. But then I am unclear about what Sullivan takes it to do, at least with regards to the *logical category* of concepts.

one argument. But this is no longer a definition of a *logical category*, since first-level concepts do not constitute a logical category in the *Grundgesetze*, only first-level functions of one argument do.[18]

My discussion so far should give an initial impression of the depth of the difference between primitive scientific *Bedeutungen* and the logical categories. The issues raised so far will be investigated in more detail in the subsequent chapters. There is a feedback mechanism here: to embark on a proper investigation of Frege's conception of the logical categories, we must not confuse the logical categories with primitive scientific *Bedeutungen*. This being established, however, the resulting investigation will itself serve to further deepen our understanding of the fundamental difference between the two.

1.3 Naïve ineffabilism

I have argued that the logical categories, for Frege, are not the *Bedeutung* of some suitable (primitive) scientific term. They do not inhabit the realm of *Bedeutungen*.[19] Readers who believe that it is possible for a term to have a logical category as its *Bedeutung*, let alone to provide definitions of the logical categories, are mistaken. This does not mean, however, that it suffices for a correct understanding of Frege's conception of the logical categories to accept the merely negative point that they do not inhabit the realm of *Bedeutungen*. There is what I will call an *irresolute* way of taking up this point that is prevalent in the literature, but which also constitutes a misunderstanding of Frege's philosophical logic, resulting in the sort of naïve ineffabilist readings discussed in the Introduction.

[18] Unless this point is clearly kept in mind, confusion is bound to arise when one engages with Frege's talk of concepts as constituting a logical category. His earlier notion of concepts as constituting a logical category should be understood as on a par with his later notion of the logical category of first-level functions of one argument, *not* with any definable species of such functions.

[19] With an eye to the *Tractatus*, we could say that they form the *scaffolding* (*Gerüst*) of the realm of *Bedeutungen* (TLP, 6.124). Note that, in proposing this as a way of putting the matter, I in no way wish to commit myself to the claim that Wittgenstein's notion of scaffolding can be simply transposed onto Frege's conception of the logical categories.

Irresolute readers of Frege (in the sense currently at issue[20]) accept that the logical categories do not inhabit the realm of *Bedeutungen*. At the same time, however, such readers hold on to the idea that all elucidations are modelled on the aim of securing agreement in *Bedeutung*. Elucidations of the logical categories are then taken to constitute as attempt to state truths about the logical categories as if they *were Bedeutungen*, so as to secure agreement in those *Bedeutungen*. Only, because the logical categories are not *Bedeutungen*, such elucidations fail in their endeavour, since what they try to say cannot, in fact, be said. As a result, they are inherently logically defective. Still, it is *qua* logically defective attempts to say what cannot be said that they nevertheless manage to *convey* the ineffable truths that they seek to express, so that they nevertheless manage to secure mutual agreement with regards to the logical categories. In this way, such elucidations behave exactly as if the logical categories were *Bedeutungen*, only we must remember to add the requisite qualifiers to signal the purportedly ineffable nature of the logical categories.

In subscribing to such a naïve ineffabilist reading of Frege, irresolute readers shy away from resolutely thinking through the implications of the fact that the logical categories do not inhabit the realm of *Bedeutungen*. They profess to accept this, but still conceive of the elucidation of the logical categories as modelled on the elucidation of primitive scientific terms, as attempting to do exactly what elucidations or primitive scientific terms do. In conceiving of statements such as 'Concepts differ fundamentally from objects' as conveying what cannot be said, such readings take back with one hand what they profess to give away with the other. In the end, the supposed acknowledgment of the fact that the logical categories do not inhabit the realm of *Bedeutungen*, and the supposed acknowledgment that the logical category distinctions cannot be properly expressed results in a conception of elucidations of the logical categories on which they end up being remarkably successful in emulating exactly the sort of judgment which – on the record – they cannot really be. All of this, however, is an illusion: this conception of elucidations as conveying that which cannot properly be said , is a hopeless philosophical muddle. The whole idea of elucidations of the logical categories as failed attempts to do what elucidations of primitive scientific terms succeed in doing needs to go out the

[20] Although there is certainly a connection to be made to the distinction between resolute and irresolute readings of Wittgenstein's *Tractatus*, this is not my immediate concern here. In the final chapter, we will briefly touch upon some parallels between certain readings of the *Tractatus* and naïve ineffabilist readings of Frege.

window. What is needed, is a conception of elucidations of the logical categories on which they do not *fail* at all, but in which they succeed in doing exactly what they are supposed to be doing, namely to elucidate the logical categories. Of course, the naïve ineffabilist could respond that – on their conception – elucidations do succeed in elucidating the logical categories, since they succeed in conveying that which cannot be said. But this purported notion of success – which, to be clear, is entirely illusory – essentially depends on elucidations being such that they are first of all failed attempts – not to *elucidate* at all – but directly to state ineffable truths about the logical categories. It is in and through their failure to *say* that which cannot be said that they are, as it were, relegated to merely elucidating it, a failsafe that comes in handy whenever we happen to run into that which cannot be said.[21] As said, this is a hopeless philosophical mess. Resolutely thinking through the difference between a primitive scientific term and a logical category term – between what can be the *Bedeutung* of a scientific term and what does not even inhabit the realm of *Bedeutungen* – requires approaching the elucidation of the logical categories as a *sui generis* activity the success of which does not depend on a prior failure to mimic judgment.

Let me discuss some examples of irresolute approaches from the literature. In his discussion of Frege, Conant[22] acknowledges that the elucidation of the logical categories constitutes a distinctive issue (Conant, 2002, 386). He points out that, by Frege's own lights, the logical categories cannot be defined, and that the sort of articulations we will come up with will be such that they "cannot be translated into a proper *Begriffsschrift*" (Conant, 2002, 386). So far, so good. Conant becomes irresolute, however, when he adds: "Their meaning [i.e. of logical category terms] must be presupposed from the outset. The most one can do is to lead the reader to what is meant by such terms—what it is one's words are trying to gesture at—by means of a series of *hints*" (Conant, 2002, 386). Just

[21] I would like to relate this to the point, emphasized by Narboux in his discussion of the Tractarian notion of showing, that one should not regard the Tractarian claim that logical form cannot be represented as a *privation* (Narboux, 2014, 210). Underlying that point is Narboux's claim that the notion of *showing* is not meant to, as it were, stand in for the notion of saying where the latter cannot fulfil its function because something is such that it cannot be said. Such a reading of the Tractarian notion of showing makes a mistake that resembles – I do not say: 'is the same as' – the mistake of modelling Fregean elucidation on judgment in the way naïve ineffabilism does.

[22] To be fair to Conant, he frames his discussion of Frege as presented in the voice of Wittgenstein in the *Tractatus*. At the same time, however, he gives no indication that he takes Wittgenstein to be mistaken. I would claim that Wittgenstein already had a more sophisticated reading of Frege. I should note that Conant no longer endorses his former reading, as is clear from his (Conant, 2020) and personal conversation. My own reading of Frege draws from (Conant, 2020) at crucial junctures.

when we thought that the logical categories were out of the realm of *Bedeutungen*, Conant pulls them back in, reverting to the model of latching onto an intended *Bedeutung*:[23]

> According to Frege, in elucidating the meaning of terms such as 'object' and 'concept', we attempt to help our audience to latch on to the intended meaning of a term for something logically fundamental by coming out with forms of expression that misfire, and then helping our audience to see how and why they misfire (Conant, 2002, 387).

Such waffling is typical of an irresolute approach, which is inherently philosophically unstable. On the one hand, the logical categories cannot be referred to. On the other hand, we talk as if everything is as it would be if the logical categories could be referred to, only we insert words such as 'ineffable', 'inexpressible', 'expressions that misfire', and the like at the right junctures to pretend that we are still fully in control of what we are saying. That control, however, is illusory: there is no coherent philosophical standpoint on offer here. The irony is that Conant realizes this, and that he takes himself to be merely laying bare Frege's own confusions. What Conant fails to see, is that the confusions he believes he finds in Frege in fact arise from confusions in his own reading of Frege. He is confused about what is confused in Frege.

As one would expect, Conant bases his reading of Frege's conception of the elucidation of the logical categories on the passages – discussed above – in which Frege presents his conception of the elucidation of primitive scientific terms. For example, Conant refers to the passage from *Logic in Mathematics* to support the following statement:

> Every science must employ some primitive terms whose meanings must be presupposed from the outset. Even in a logically perfect language there will be some terms that are not (and cannot) be introduced by definition and that must remain undefinable. The purpose of elucidations is to convey the meanings of such terms (Conant, 2002, 386).

This is indeed a faithful rendition of Frege's conception of the elucidation of primitive scientific terms. Conant's failure to suitably distinguish that topic from the elucidation of the logical categories, however, can only result in a muddle.

[23] Compare also (Witherspoon, 2002, 104ff.).

There are plenty of examples of readers who take the passages concerned with elucidations of primitive scientific terms and proceed to apply them to elucidations of the logical categories. Proops, for instance, infers from the fact that elucidations of primitive scientific terms can consist in judgments in which those terms are used with their intended *Bedeutung* that Frege "cannot have regarded elucidations as nonsensical" (Proops, 2013, 93). He then proceeds to apply this point to elucidations of the logical categories, concluding that Frege regarded those as judgments as well, and that he thereby failed to face up to the predicament that arises from his conception of the logical categories. I would counter that Frege did face up to that predicament, that he developed a specific approach to the elucidation of the logical categories in an attempt to overcome it, and that Proops' claim rests on an unjustified generalization of what Frege says about elucidations of primitive scientific terms. Another example is Joan Weiner, who quotes a passage from the *Foundations of Geometry* article as if it was straightforwardly meant to apply to elucidations of the logical categories (Weiner, 2006, 200), and then encounters various difficulties.[24] Such examples can be easily multiplied: again and again, readers of Frege fail to distinguish clearly between the elucidation of primitive scientific terms and the elucidation of the logical categories, taking Frege's remarks about the former – and the concomitant aim of securing mutual agreement in *Bedeutung* – to set the framework for understanding the latter.

Although most readers agree (as Conant does) that naïve ineffabilism is philosophically hopeless,[25] it is fair to say, I think, that there is widespread agreement among Frege scholars that Frege did not have much more to say about the elucidation of the logical categories than what naïve ineffabilism provides. The real work to investigate whether we might not find a philosophically more subtle account in Frege remains to be done. When interpreters find Frege talking about 'elucidation', they never stop to ask what *kind* of elucidation might be at issue, as if the mere fact that he uses one term is sufficient to

[24] To be sure: it is to Weiner's credit that she sees such difficulties to begin with. Nevertheless, as will be indicated below, Weiner does not fully escape the deleterious effects of taking such passages to apply to elucidations of the logical categories.

[25] See (Diamond, 1981) (Diamond, 1984, 367) (Long, 2001, 101) (Conant, 2002) (Witherspoon, 2002, 111ff.) (Johnston, 2007, §2.2) (Hale & Wright, 2012, 100ff.) (Moore, 2012, Chapter 8, §III) for some arguments against this kind of position. I am inclined to agree with Sullivan that "no one would knowingly defend it" (Sullivan, 2003, 206). Note that one exception is (Geach, 1976), who seems to want to endorse what may be regarded as a more sophisticated version of naïve ineffabilism, both as a reading of Frege and as a philosophical position. I will discuss his account in detail in the final chapter.

conclude that he is always talking about roughly the same activity.[26] I am not aware of any reader of Frege who shows themselves to be explicitly aware that an exegetical assumption is made here, let alone who questions it. Once one does question it, however, one starts to approach Frege's remarks about the logical categories more carefully, and one begins to see that Frege's approach to the elucidation of the logical categories is quite different from his approach to the elucidation of primitive scientific terms, as it *must* be if his conception of the logical categories is not to fall into the hopeless sort of naïve ineffabilism that will serve as a target throughout this dissertation.

One could object that – for all I have said – it remains the case that Frege nowhere registers the fact that there are different notions of 'elucidation' at issue in his work, and that I am simply wrong in claiming otherwise. It is true that, insofar as one wants from Frege a direct and clear statement of the difference between the elucidation of primitive scientific terms and the elucidation of the logical categories, one will not find it. Because of this, Frege's talk *about* elucidation can easily generate the impression that there is only one notion of elucidation at issue, which is why it has led interpreters astray. When we look at Frege's *practice* of elucidation, however, clear differences start to appear between the elucidation of primitive scientific terms and the elucidation of the logical categories. As we will see, Frege's approach to elucidating the logical categories in no way corresponds to what he says about elucidations of primitive scientific terms.

Moreover, once one starts distinguishing the two, some more subtle differences in Frege's discussions of elucidation also start to appear. For instance, Frege's talk about elucidation as constituting a propadeutic to actual science only appears with regards to the elucidation of primitive scientific terms. This notion of elucidation as a propadeutic has been given a lot of weight in the work of Joan Weiner.[27] Although Weiner is certainly an astute reader of Frege – whose work has, to be clear, been invaluable to my own research – and although she does not blankly ascribe naïve ineffabilism to Frege,[28] I would

---

[26] It is, of course, no coincidence that Frege uses the same term. It possesses the sort of negative unity captured by the characterization that was presented in the introduction to this chapter. But such a negative characterization obscures the deep differences between the multiple activities of elucidation at issue in Frege's *oeuvre*. Once the elucidation of primitive scientific terms has been set aside, moreover, it turns out that – although, to be sure, he does use it in this way – Frege does not use the term 'elucidation' as often with regards to the logical categories as one might expect. What he does consistently do, is mark off his logico-philosophical discourse about the logical categories as having a different status than expressions of judgment, in the myriad sorts of ways that were touched upon in the Introduction.

[27] See e.g. (Weiner, 2005, 339) (Weiner, 2006, 202f.), (Weiner, 2010, 61)

[28] Although she does at one point suggest that Frege's elucidations of the logical categories constitute "failed attempts to express the inexpressible" (Weiner, 2010, 59).

still claim that her approach to the elucidation of the logical categories is insufficiently resolute, failing to fully register the difference between it and the elucidation of primitive scientific terms.

Weiner writes: "What marks a discussion as elucidatory is neither its form nor its content but, rather, its role in the project. The mark of elucidation is its contribution to the propadeutic" (Weiner, 2010, 61). Insofar as one understands 'propadeutic' merely negatively in the sense of 'not belonging to science proper', this characterization of elucidation is fine as it goes. But such a purely negative characterization once again threatens to obscure the fundamental difference between the sense in which elucidations of primitive scientific terms do not belong to science proper and the sense in which elucidations of the logical categories do not belong to science proper, a difference that Weiner fails to address. When she writes, for instance, that "Frege's introduction of the term 'elucidation' is meant to highlight the difference between these attempts to communicate the meanings of terms and actual definitions" (Weiner, 2010, 60), she does not seem to realize that this notion of 'communicating meanings' only applies to the elucidation of primitive scientific terms. The sense in which elucidations of primitive scientific terms do not belong to science proper – but instead to the propadeutic – is indeed that they serve to communicate *Bedeutungen* that cannot be presented through definitions but about which mutual agreement must be secured before science can get underway. The sense in which elucidations of the logical categories do not belong to science proper, however, is entirely different: it is because the logical categories do not belong to the subject matter of science at all. There is *no* subsequent scientific use of logical category terms for elucidations of the logical categories to be a propadeutic to. I take it to be no coincidence, then, that Frege does not employ the notion of a propadeutic with regards to elucidations of the logical categories. Of course, one is free to say that elucidations of the logical categories belong to the propadeutic. Only, if sense is to be made of what one has thereby said, this will have to result in an understanding of the notion of 'propadeutic' that differs sharply from the understanding of that notion that is relevant to elucidations of primitive scientific terms. It is precisely this difference to which Weiner fails to do justice in her writings.

1.4 Logico-philosophical elucidation

In this chapter, I have presented some initial considerations to show that there is a sharp difference, in Frege, between the elucidation of primitive scientific terms, on the one hand, and the elucidation of the logical categories, on the other hand. I have also presented some indications to show that this distinction has been neglected or underestimated in the literature, yielding distorted readings of Frege. These points will receive further confirmation as Frege's conception of the logical categories is further unfolded throughout this dissertation.

My presentation so far may generate the impression that there are, at bottom, two notions of elucidation at issue in Frege: the elucidation of primitive scientific terms and the elucidation of the logical categories. I should make it clear, therefore, that this is not a claim to which I am committing myself. If we revert back to the negative characterization of elucidation as what is required to secure mutual agreement in those cases where definition is unavailable because a certain phenomenon is logically primitive, then we quickly see that there is a whole array of phenomena that constitute matters for elucidation. Besides primitive scientific terms and the logical categories, these include truth, judgment, *Bedeutung*, *Sinn*, and many others.

One salient dividing line is the following: only in the case of the elucidation of primitive scientific terms does elucidation concern a phenomenon that constitutes a subject matter for judgment, i.e. does elucidation concern that which inhabits the realm of *Bedeutungen*. The same sort of arguments that I have presented in this chapter can be used to establish that none of the other phenomena just mentioned – truth, judgment, *Bedeutung*, *Sinn*, etc – constitute a subject matter for judgment, so that their elucidation also cannot be conceived as a *propadeutic* to a subsequent scientific use of the corresponding primitive terms. Based on this observation, I propose to refer to the elucidation of primitive scientific terms as *scientific elucidation*, and to characterize these other activities of elucidation – which concern that which is not a subject matter for judgment, and which is thereby *ineffable* in this minimal sense – as *logico-philosophical elucidation*. The latter concerns certain fundamental logico-philosophical terms of art of Frege's philosophical logic, terms of art that have no *Bedeutung* and are not employed in the development of a scientific system.

In line what was said above, this notion of 'logico-philosophical elucidation' should not be taken as being univocal. We should not suppose that the logico-philosophical elucidation of the logical categories is the same sort of activity as, for instance, the logico-philosophical elucidation of judgment. In this dissertation, my concern lies primarily with Frege's conception of the logical categories, and therefore with the logico-philosophical elucidation of the logical categories.[29] This is not to say that one can cleanly separate, for instance, the elucidation of the logical categories and the elucidation of judgment. The former unavoidably involves a concern with judgment as well, and *vice versa*, as will be clear from my own discussion.[30] At the same time, the fact that these activities are inevitably intertwined, need not mean that there are no crucial differences between them. A sustained investigation of Frege's notions of truth, judgment, etc, and his concomitant conception of the elucidation of those notions, is a task for further research. Such research, to be sure, may very well impact my own reading of Frege's conception of the logical categories in substantive ways. My only reply is to say that one has to start somewhere, and that I hope that this dissertation can serve as a fruitful basis for further investigations of Frege's philosophical logic, investigations that would undoubtedly necessitate important revisions in my own account.[31]

[29] For ease of presentation, I will often leave the qualifier 'logico-philosophical' implicit, since the subject matter for elucidation already serves to distinguish between scientific and logico-philosophical elucidation.

[30] Compare: "No elucidation of logically primitive notions (such as 'content,' 'object,' 'concept,' 'thought,' 'judgment,' 'inference,' and 'truth') is possible except in terms of others—a proper understanding of any of them ultimately depends on a proper understanding of all of them. Hence, no logically fundamental aspect of the logical order can be reduced to something non-logical and none can be elucidated except in the light of its place within the unity of that order. This is also why Frege will insist that no fundamental logical notion can be defined" (Conant, 2020, 795-796). Compare also (Ricketts, 1986, 71).

[31] Indeed, this gesture of marking a certain topic as falling outside the scope of this dissertation while at the same time admitting that a proper investigation of that topic is likely to have important repercussions for my account, is a gesture that I will be forced to make several times throughout this dissertation. That Frege's work forces me to do this, reveals that it possesses the sort of unity that, I think, characterizes all great works in philosophy, a unity that is the effect of a sustained attempt to think through – without compromise – certain philosophical issues in their utmost detail. One could say that it is the mark of philosophical *vision*. Against what some may be inclined to think, I would say that such vision may be found in the best work in contemporary analytical philosophy as well, such as the work of the great Timothy Williamson.

Chapter 2
The logical stratification of subjecthood

In the first chapter, we have seen how naïve ineffabilist readings of Frege arise from making too slight the difference between elucidation and judgment, between the logical categories and *Bedeutungen*. There is, however, a passage from Frege's *oeuvre* that seems to directly undermine my claim that Frege was not a naïve ineffabilist. This is the notorious passage from *Über Begriff und Gegenstand* where Frege asks his reader for a grain of salt:

> I do not at all dispute Kerry's right to use the words 'concept' and 'object' in his own way, if only he would respect my equal right, and admit that with my use of terms I have got hold of a distinction of the highest importance. I admit that there is a quite peculiar obstacle in the way of an understanding with the reader. By a kind of necessity of language, my expressions, taken literally, sometimes miss my thought; I mention an object, when what I intend is a concept. I fully realize that in such cases I was relying upon a reader who would be ready to meet me half-way – who does not begrudge a pinch of salt (CP, 193; KS, 177; 204).

Frege here evinces a predicament: he is trying to express a thought, yet his expressions miss his thought. As far as I can tell, all scholars agree that this passage espouses a form of naive ineffabilism as a response to that problem. Since most scholars also agree that naive ineffabilism is philosophically hopeless, the passage is generally regarded as exhibiting a deplorable laxity on Frege's part, who is seen as failing to fully face up to the philosophical difficulties posed by his position.[1] The naive ineffabilist reading depends,

[1] E.g. (Dummett, 1981, 212) (Ricketts, 2010, 186) (Moore, 2012, 218-219) (Hale & Wright, 2012, 91ff.).

however, on a certain answer to the question *what* thought it is, exactly, that Frege takes his expressions to miss. On the naïve ineffabilist reading, the predicament at issue in this passage is the predicament presented in the Introduction, so that what Frege's expressions miss, is the absolute distinction between concepts and objects as such.[2] Read in this way, the passage will indeed inevitably appear as construing that distinction as some kind of ineffable truth, thereby yielding a version of naïve ineffabilism. I will argue, however, that this is a mistake.

There is, in fact, another predicament that also arises from Frege's conception of the logical categories, but which must be kept clearly distinct from the issue of the seemingly self-undermining nature of Frege's elucidations of the logical categories. This other predicament does not concern statements about, for instance, the logical category of concepts as such, but rather statements about *particular* concepts, such as the statement 'The concept *horse* is realized'. What we have here, is a statement that purports to express a judgment about a particular concept. Only it cannot do so, since the expression 'the concept *horse*' is a proper name that refers to an object. Thus, it seems that Frege's absolute distinction between concepts and objects precludes the possibility of making concepts into logical subjects, and thereby precludes the possibility of making judgments about concepts.

Frege's way of dealing with this second predicament is, moreover, not naïve ineffabilist at all. Rather, Frege believes that the *Begriffsschrift* device of second-level subsumptions provides us with the means for making judgments about concepts in a way that the natural language device of using the singular definite article does not allow for. Properly construed, a sentence such as 'The concept *horse* is realized' must be seen as a defective attempt to express what can be perfectly rigorously expressed in a second-level subsumption such as '$(\exists x)(x$ is a horse)'.

Sentences such as 'The concept *horse* is realized', then, have an entirely different status than sentences such as 'Concepts differ fundamentally from objects'. We should not confuse the notion of making a judgment about the logical category of concepts – which,

[2] As already pointed out in the previous chapter, it is important to keep in mind that the concept/object distinction, *qua* logical category distinction, was superseded by the distinction between objects and first-level functions of one argument. In the *Grundgesetze*, concepts are a kind of first-level functions. Because Frege's reply to Kerry proceeds in terms of the concept/object distinction, however, and because this is also the terminology that is employed in most of the secondary literature that will concern us in this chapter, I will mostly talk in terms of the concept/object distinction. It must be kept in mind that the term 'concept' is accordingly used as a logical category term, not as giving us a kind of first-level functions.

we have seen in the first chapter, Frege does regard as incoherent – with the notion of making a judgment about particular concepts, concluding that Frege regarded the latter notion as equally confused, as many readers of Frege have done. To properly understand what is at issue in the grain of salt passage – and thereby in Frege's dialectic with Kerry – we must distinguish between these two distinct activities in Frege: elucidating the logical categories, on the one hand, and making judgments about concepts, on the other hand. Rather than regarding the latter as confused, Frege was in fact at pains to *vindicate* it.

My reading of Frege is committed to the idea that both objects and concepts can be said to be logical subjects, that both can be what judgments are *about*. This reveals that a term such as 'logical subject' is not used as a scientific terms with *Bedeutung*, but instead has what I will call a *logically stratified elucidatory use*. That such a use must be clearly distinguished from the scientific use of terms with *Bedeutung*, is a point familiar from the first chapter. Properly recognizing how second-level subsumptions are *about* first-level functions just as much as first-level subsumptions are *about* objects, just is to recognize the logically stratified function/argument nexus that lies at the core of Frege's philosophical logic. It gives rise to what I will call the *functionality* of the realm of *Bedeutungen*. This functional stratification also gives us the logico-philosophical form of unity of the realm of *Bedeutungen*. Logical stratification *is* a form of unity, and to properly understand it, we must understand how the logical stratification of the realm of *Bedeutungen* comes with the logical stratification of a cluster of fundamental logico-philosophical phenomena, such as logical subjecthood, aboutness, *Bedeutung*, and so on. Only in this way can we properly understand how Frege seeks to vindicate the possibility of making judgments about concepts, a vindication that involves a reconciliation of the following commitments:

(1) Concepts are essentially predicative.
(2) Concepts can be made into the logical subjects of judgments.

What is at stake in Frege's dialectic with Kerry – and thereby also in the grain of salt passage – is precisely this reconciliation, a reconciliation that is made possible by Frege's conception of the logical stratification of logical subjecthood.

2.1 The road to naive ineffabilism

Let us start by investigating a bit more closely how the grain of salt passage and similar passages seem – at first blush – to support naive ineffabilist readings of Frege. Here is the grain of salt passage again:[3]

> I do not at all dispute Kerry's right to use the words 'concept' and 'object' in his own way, if only he would respect my equal right, and admit that with my use of terms I have got hold of a distinction of the highest importance. I admit that there is a quite peculiar obstacle in the way of an understanding with the reader. By a kind of necessity of language, my expressions, taken literally, sometimes miss my thought; I mention an object, when what I intend is a concept. I fully realize that in such cases I was relying upon a reader who would be ready to meet me half-way – who does not begrudge a pinch of salt (CP, 193; KS, 177; 204).

Two aspects of the passage – which reappear throughout similar passages in Frege's *oeuvre* – are pertinent. The first is Frege's talk about his expressions missing his thought. Elsewhere, Frege talks about his thought being 'obscured' or 'falsified', about 'making a mistake', 'errors of thought', or about his words 'missing their intended target'.[4] This seems to confirm the naive ineffabilist picture of there being ineffable truths that Frege's statements are trying to but ultimately fail to express. In attempting to state the concept/object distinction, for instance, we try to say what really cannot be said, so that the statements we come out with miss their intended target.

The second aspect is Frege's insistence that language is to blame for the difficulty. Apart from Frege saying that the fact that his expressions miss his thought is a 'necessity of language' in the grain of salt passage, we elsewhere find him talking about an 'awkwardness of language', the 'inappropriateness of language', an 'idiosyncracy of language', and the like[5]. This too seems to accord well with naive ineffabilism, since it suggests that there are certain expressive limitations of language that debar us from talking about the logical categories. If only we *could* talk about them, we would, but due

[3] Two scholars who explicitly invoke it in support of naive ineffabilist readings are (Conant, 2002, 392) (Witherspoon, 2002, 108-109).
[4] See (PW, 119-120, 177-178) (PMC, 142, 193, 255, 273).
[5] See (FLL, 127) (PW, 119, 177-178, 239, 250, 270, 273) (CP, 185, 194, 282).

to certain limitations of language they remain ineffable – they lie beyond the limits of language – so that we can only gesture at them through logically defective elucidations.

It is not hard to see how such passages can inspire naive ineffabilist readings of Frege according to which the elucidation of the logical categories proceeds through statements that constitute failed attempts to say what cannot be said. Such a reading, however, presupposes that these passages are in fact concerned with purported statements about the logical categories, such as 'Concepts are fundamentally different from objects'. I will now argue that this is mistaken. In fact, there are two predicaments facing Frege's conception of the logical categories, and these must be clearly distinguished. Once this is seen, Frege's talk about 'missed thoughts' and 'necessity of language' loses the ineffabilist implications that it has seemed to virtually all readers of Frege to have.

2.2 Judgments about concepts

To see what is actually at issue in the grain of salt passage, it is helpful to take a brief excursion to Russell's *Principles of Mathematics*. In the *Principles of Mathematics*, there is one all-encompassing logical category of *terms* (PoM, §47). That category includes both what Russell calls *things* and *concepts*. Whereas things can only occur non-predicatively in propositions, concepts can occur as a *verb*, i.e. predicatively (PoM, §48). Concepts *can* occur predicatively, but they do not have to. In the proposition 'Socrates is human', for instance, the concept *humanity* occurs predicatively and therefore cannot be replaced by a thing. Because it occurs predicatively, *humanity* is not the logical subject of this proposition; the proposition is not about *humanity*. The proposition 'Humanity belongs to Socrates', on the other hand, *is* about *humanity*, which is the logical subject of this proposition. This is an illustration of Russell's principle that "every constituent of every proposition must, on pain of self-contradiction, be capable of being made a logical subject" (PoM, §52). 'On pain of self-contradiction', because this is the only way to make coherent the fact that we can and do – as I just did in my discussion of *humanity* – make judgments about concepts, to make coherent the fact that they are objects of thought. Even just saying '*humanity* cannot be made into a logical subject' would already be self-defeating, since *humanity* occurs as a logical subject in this very statement.

We will return to such Russellian lines of thought in the next chapter. For now, it should be noted that, according to Frege, concepts are *essentially* predicative, so that there is no overarching logical category that encompasses both concepts and objects, as there is for Russell. For Frege, there is no such thing as having a concept appear non-predicatively in a judgment. It follows that there is no form of logical generality that encompasses everything that can be a constituent of a judgment, as there was for Russell. Rather, such constituents are stratified into distinct logical categories, and logical generality remains wedded to a single logical category. In that sense, not even logic can talk about 'everything' in the way it can for Russell.[6]

This raises the following problem for Frege: if a concept cannot occur non-predicatively, it seems that it cannot be made into a logical subject, so that it is impossible to make judgments about concepts. But if it is impossible to make judgments about concepts, it becomes unclear with what right we recognize concepts as constituents of judgments to begin with. To repeat Russell's principle: "every constituent of every proposition must, on pain of self-contradiction, be capable of being made a logical subject" (PoM, §52). Many take it for granted that Frege disagrees with Russell here. I will argue that there is an important sense in which Frege *agrees* with Russell.

If one believes that Frege took it to be impossible to make concepts into logical subjects, it will start to seem that it is not only truths about the *logical category* of concepts that are ineffable, but also truths about *particular* concepts, in such a way that any attempt to make a judgment about a particular concept ends up in the same sort of naïve ineffabilist gesturing. We will see that this is indeed the sort of conclusion that some interpreters arrive at, whereby the oil spill of naïve ineffabilism is made to spread even further through Frege's *oeuvre*. If this is to be avoided, a solution must be found to the problem of making judgments about concepts. That is to say: we must find in Frege the possibility of combining the following two ideas:

(1) Concepts are essentially predicative.
(2) Concepts can be made into the logical subjects of judgments.

I will argue that Frege did indeed seek to combine (1) and (2), that this is what was at stake in his dialectic with Kerry, and that it is only by seeing this that we can come to an adequate understanding of the nature of Frege's logico-philosophical innovations

[6] One may compare my discussion to (Kremer, 2006, 183ff.).

underlying the *Begriffsschrift*. To see how Frege sought to reconcile (1) and (2), we have to see how his philosophical logic involves a radical rethinking of the notion of a logical subject itself, as it figures in (2).

Let us start by looking at a characteristic passage where Frege articulates the problem of making judgments about concepts:

> If I want to speak of a concept, language, with an almost irresistible force, compels me to use an inappropriate expression which obscures—I might almost say falsifies—the thought. One would assume, on the basis of its analogy with other expressions, that if I say 'the concept *equilateral triangle*' I am designating a concept, just as I am of course naming a planet if I say 'the planet Neptune'. But this is not the case; for we do not have anything with a predicative nature. Hence the meaning of the expression 'the concept *equilateral triangle*' (if there is one in this case) is an object (PW, 119-120; NS, 130).

First of all, we should note that the issue is indeed how to make judgments about a particular concept, not how to articulate the concept/object distinction itself. The problem presents itself as follows: when we try to make a judgment about the concept *equilateral triangle* – when we try to make that concept into a logical subject – we are brought to use the proper name 'the concept *equilateral triangle*', which is used non-predicatively and thereby refers to an object instead. The essentially predicative nature of the concept is lost. According to Frege, this is marked by the definite article: "the singular definite article always indicates an object" (CP, 184).[7]

The above passage is not an exception: the problem of making judgments about concepts was a recurring concern for Frege. In the relevant passages, he consistently talks about the problem of 'speaking of a concept', 'designating a concept', 'mentioning a concept', and the like, and consistently mentions constructions involving the singular definite article.[8] Frege did not believe, however, that this problem requires taking recourse to naive ineffabilism. The problem arises because "language brands a concept as an object, since the only way it can fit the designation for a concept into its grammatical

[7] See also (CP, 387). I will have more to say about this below.
[8] See (CP, 193, 282) (FLL, 127) (PW, 119, 193, 239, 250, 270, 273).

structure is as a proper name" (PW, 177).[9] What is true for ordinary language, however, is not true of the *Begriffsschrift*. The *Begriffsschrift* does have a systematic way to fit the designation for a concept into its grammatical structure without turning it into a proper name. The solution to this problem, then, does not lie in naïve ineffabilism, but instead in the development of a logical notation that offers the expressive tools needed to make judgments about concepts, as the *Begriffsschrift* does.

These tools are provided by the distinction between first-level concepts and second-level concepts, which is one of the cornerstones of the *Begriffsschrift*.[10] First-level concepts are used to make statements about objects of the straightforward kind, such as 'Seabiscuit is a horse'. In such a statement, an object is subsumed under a first-level concept, so that the logical subject is the object. Call these *first-level subsumptions*. Second-level concepts, on the other hand, are used to make statements about first-level concepts. In such a statement, a first-level concept is subsumed under a second-level concept, so that the logical subject is the first-level concept. Call these *second-level subsumptions*. An example is '$(\exists x)(x$ is a horse)', which subsumes the first-level concept ξ *is a horse* under the second-level concept $(\exists x)(\Phi x)$, saying about it that at least one object falls under it, i.e. that it is realized.[11]

Thus, we see how Frege's way of dealing with the problem of making judgments about concepts leads us further into the logical stratification of the realm of *Bedeutungen*,

[9] This may seem a strange statement: can't ordinary language straightforwardly fit the designation for a concept into its grammatical structure, for instance in sentences such as 'Seabiscuit is a horse', in which the concept ξ *is a horse* is designated? Yes, but Frege is thinking about the question of how to make judgments *about* a concept. It is in our ordinary language attempts to do this that we are irrevocably driven to use proper names such as 'the concept *horse*', so that we do not succeed in doing what we want to do. There seems to be one salient exception to this: statements such as 'There exists at least one horse'. We will see below how to understand the status of such ordinary language statements.

[10] Frege repeatedly invokes it in his correspondence with Russell and his criticism of Hilbert (PMC, 136, 141) (CP, 280ff., 307-308). For some other places where it is discussed, see (CP, 153) (GGA, I, §21ff.) (PW, 182). One reader of Frege who explicitly seeks to do justice to the importance of this distinction for Frege, is (Macbeth, 2005). Juxtaposing her reading with mine would lead me too far astray here.

[11] There is a complication here: Frege's conception of the horizontal as a first-level function makes it the case that all judgments are first-level subsumptions. In this way, there is a sense in which even Frege remains wedded to the old subject/predicate logic. Recall that Frege says, in his original presentation of the *Begriffsschrift* in 1879, that it can be conceived as having one predicate 'is a fact' for *all* judgments (BS, §3). Investigating this complication – which would certainly complicate my own account – is a task for another occasion. It brings into play Frege's conception of the force/content distinction, which requires him to find a *single* sign for the expression of assertoric force, i.e. the judgment-stroke. It is because Frege needs such a single sign, that he must regard all judgments as in the end sharing a first-level subsumption structure. That is to say: Frege is committed to the claim that the sense in which it may be *judged that* Socrates is bald, and the sense in which it may be *judged that* there exists a square root of 4 – i.e. the sense in which both a first-order subsumption and a second-order subsumption may be judged to be true – are *the same*, a sameness that is secured by the horizontal.

bringing in the category of second-level concepts.[12] Second-level subsumptions are of a logically different kind than first-level subsumptions, and Frege did not forget to lament the fact that these are often conflated.[13] The *Begriffsschrift*, however, clearly distinguishes between first-level subsumptions and second-level subsumptions through its device of argument places. A first-level subsumption has the form '$F(a)$', so that the argument-place 'a' can only be occupied by a proper name. A second-level subsumption has the form '$\mathrm{M}_{\xi}\big(\varphi(\xi)\big)$'[14] (GGA, I, §25), so that the argument-place can only be occupied by a first-level function of one argument. And this gives us the above-mentioned systematic way to fit the designation for a concept into the grammatical structure of the *Begriffsschrift* without turning it into a proper name.

It may be thought that this is all rather straightforward. To a limited extent it is, since we have ourselves benefited from Frege's insights. Frege's contemporaries, however, were far from clear about these matters, as Frege never tired of pointing out. At the same time, we should not be too quick in declaring ourselves enlightened by Frege. As we will see, contemporary readers of Frege also struggle to get clear about the exact nature and import of these logico-philosophical innovations underlying Frege's *Begriffsschrift*. We would do well, then, to not underestimate the difficulty of bringing full clarity to these matters, which is a task that will occupy us for the remainder of this dissertation. Let us start by taking a renewed look at the grain of salt passage.

---

[12] There will then be the additional issue of making second-level concepts into logical subjects, which will involve third-level concepts, and so on.

[13] One paradigmatic case of such conflation is that between first-level subsumption and subordination between first-level concepts. See (CP, 183n, 190, 216) (PMC, 68, 100, 109) (PW, 18, 213).

[14] The argument-place is here marked by the expression '$\varphi(\,)$' (GGA, I, §25). The inserted Greek letter 'ξ' does *not* itself figure as an argument-place in this second-level expression. That is why it is bound by the subscript of the 'M_{ξ}', in a way similar to that in which a quantifier binds a variable. It goes together with the argument-place, since its presence marks the fact that only expressions standing for first-level functions of one argument can occupy the argument-place of '$\mathrm{M}_{\xi}\big(\varphi(\xi)\big)$'. Frege calls such an argument-place an *argument place of the second kind* (GGA, I, §23). The result of inserting the first-level concept expression '$\xi > 3$' into the argument-place of '$\mathrm{M}_{\xi}\big(\varphi(\xi)\big)$', for instance, is '$\mathrm{M}_{\xi}(\xi > 3)$'. If we take the universal quantifier as our second-level concept, then we get '$(\forall x)(x > 3)$', which is a false second-level subsumption.

2.3 Frege's grain of salt

Here is the grain of salt passage again:

> I do not at all dispute Kerry's right to use the words 'concept' and 'object' in his own way, if only he would respect my equal right, and admit that with my use of terms I have got hold of a distinction of the highest importance. I admit that there is a quite peculiar obstacle in the way of an understanding with the reader. By a kind of necessity of language, my expressions, taken literally, sometimes miss my thought; I mention an object, when what I intend is a concept. I fully realize that in such cases I was relying upon a reader who would be ready to meet me half-way – who does not begrudge a pinch of salt (CP, 193; KS, 177; 204).

The crucial phrase is: 'I mention *an* object, when what I intend is *a* concept'. The obstacle Frege is concerned with here is not an obstacle to stating the concept/object distinction as such[15], but an obstacle to making judgments about concepts.[16] We are now in a position to take the ineffabilist sting out of the grain of salt passage. Consider a statement such as 'The concept *horse* is realized'. In the *Begriffsschrift*, as we have seen, this is expressed by '$(\exists x)(x$ is a horse)'. In fact, the whole solution is already contained in my use of the phrase 'this is expressed by'. Frege's talk about 'missed thoughts', correctly understood, is not ineffabilist at all: there is a thought that is in perfect logical order, and that Frege's statements are meant to express – a second-level subsumption – but which Frege's statements fail to express due to the logical deficiency of the grammar of ordinary language. Taken in accordance with the grammar of ordinary language, the sentence 'The concept *horse* is realized' expresses a first-level subsumption. If we grant Frege his grain of salt, however, we recognize that Frege aims to express a second-level subsumption, and we will understand him accordingly.

---

[15] To be clear: I do not wish to say that there is no such obstacle, only that it is not the focus of the grain of salt passage. The obstacles facing purported statements of the concept/object distinction give rise precisely to Frege's predicament that was discussed in the Introduction.

[16] These obstacles correspond to the third and fourth problem discussed in (Proops, 2013). As far as I know, Proops is the only interpreter to have clearly distinguished these. He also connects this to the grain of salt passage (Proops, 2013, 86f.). Proops does not discuss, however, how the *Begriffsschrift* device of second-level subsumptions allows us to overcome the second obstacle. He concludes that "although [...] the problem came to seem inevitable to Frege, it is equally clearly something he thought we could learn to live with" (Proops, 2013, 88). Frege *did* think we could learn to live with it, but this is because he thought the *Begriffsschrift* solves the problem. In this way, it is not inevitable at all, but really is due to an awkwardness of ordinary language.

We should understand Frege's talk about the inappropriateness of language in a similar way. That Frege's thought is missed, *is* due to the expressive limitations of ordinary language, which "by its use of the definite article to stamp as an object what is a function and hence a non-object, proves itself to be the source of inaccurate and misleading expressions and so also of errors of thought" (PW, 273).[17] The *Begriffsschrift*, on the other hand, gives us a logically perspicuous notation wherein first-level subsumptions and second-level subsumptions are clearly distinguished, so that we can unproblematically make judgments about particular first-level concepts rather than objects. In this way, the limitations of ordinary language are overcome. When Frege asks his readers for a grain of salt, he is asking them to look beyond the inappropriateness of his ordinary language expressions, an inappropriateness that is due to the limitations of the grammar of ordinary language. What lies beyond, however, is not some mysterious kind of ineffable truth, but a logically perspicuous notation that allows us to express the very thoughts that Frege's inaccurate statements are meant to convey.

To be sure, readers often recognize that statements such as 'the concept *horse* is realized' – which purport to be about particular concepts – are at issue in the grain of salt passage. This recognition, however, tends to remain philosophically inert, because they nevertheless fail to register the underlying difference between such statements, on the one hand, and elucidatory statements such as 'Concepts are fundamentally different from objects', on the other hand. A good example is again Conant[18], who writes about the grain of salt passage:

> [Frege's] words miss his thought (and end up being nonsense); so there is a thought they are aiming at: an understanding of what his words intend to say depends upon his reader latching onto the thought his words fail properly to express. This failure is due, according to Frege, to 'a kind of necessity of language.' If he is to convey the thought he here seeks to convey, *he has no alternative but to have recourse to (elucidatory) nonsense* [my emphasis] (Conant, 2002, 392).

Conant mentions both tropes of the 'missed thought' and the 'necessity of language', but reads them as endorsements of naïve ineffabilism: there is a thought Frege is aiming at,

[17] In Carnap's lecture notes, we find: "It is a deficiency of *everyday language* that we have to talk as if the function were an object" [my emphasis] (FLL, 127).

[18] Other readers who directly tie the grain of salt passage or similar passages to the problem of stating the concept/object distinction itself include (Witherspoon, 2002, 106ff.) (Weiner, 2006, 202) (Ricketts, 2010, 186) (Moore, 2012, 218-219) (Hale & Wright, 2012, 91-92) (Bengtsson, 2018, 107) (Travis, 2020, 222-223).

but *any* attempt to express that thought *must* fail and thereby issue in nonsense. Thus, Conant extends his naïve ineffabilist reading of elucidations such as 'Concepts are fundamentally different from objects' to statements such as 'the concept *horse* is realized', failing to distinguish between the two. This is further confirmed in the following passage:

> Frege's discussion turns [...] on the idea that we know *what* it is we are trying to say (*when we employ an expression such as 'the concept prime'*), but when we try to say 'it,' we realize that what we are trying to say *requires* that what we actually say be something nonsensical [last two emphases mine] (Conant, 2002, 389).

Again, the idea is that what Frege is trying to say in using expressions such as 'the concept *horse*' is of such a nature that it *requires* saying something nonsensical. But it requires no such thing: what Frege seeks to do, can be done by using the *Begriffsschrift* device of second-level subsumptions. In lumping together statements such as 'The concept *horse* is realized' with statements such as 'Concepts are fundamentally different from objects', Conant is unable to see this.

Seen in the right light, the grain of salt passage and similar passages offer no grounds for a naive ineffabilist reading of Frege. Instead, they mark a moment in Frege's logico-philosophical investigations where he realizes that, on the one hand, he must employ ordinary language in ways that come natural to us in order to communicate his views while, on the other hand, such employment threatens to be misleading when what is at issue are judgments about concepts. The grammar of ordinary language forces Frege to render his judgments about concepts in such a way that they appear as first-level subsumptions, whereas they are meant to be second-level subsumptions. For his readers to understand this, they must understand that statements about concepts are in fact second-level subsumptions. But it is precisely this that ordinary language renders so difficult to properly understand. Frege is forced to use the logically deficient expressive means of ordinary language in an attempt to instil in his readers a sense of the logical deficiency of those means. This is a difficult exercise, and it gives rise to the second predicament facing Frege's conception of the logical categories. It is in connection with this exercise that the notions of missed thoughts and awkwardness of language, as well as Frege's request for a grain of salt, have their own elucidatory role to play – not as espousing naive ineffabilism, but as indications that Frege's readers are to look beyond the logically deficient grammar of ordinary language to the logically proper *Begriffsschrift* notation.

2.4 The logical stratification of subjecthood

Conant's reading is not the only reading of Frege that is tainted by a failure to distinguish between the elucidation of the logical categories, on the one hand, and making judgments about particular concepts, on the other hand. I would like to discuss in some detail another such reading, advanced by Kelly Dean Jolley. This will allow us to achieve further clarity about the way in which the device of higher-level subsumptions is central to Frege's philosophical logic, and how it is intertwined with the logical stratification of the realm of *Bedeutungen*.

Jolley aims to show how certain approaches to Frege may *seem* to do justice to the concept/object distinction, but end up undermining it. Here are some of the ways in which Jolley characterizes the sort of views that he wishes to criticize:

> Valberg is trying to secure a way to talk about concepts. He wants to find a way to talk about something other than objects, individuals (Jolley, 2007, 52).

> Kerry thinks that there is no reason, if a concept is one of the items that make up a thought, why that concept cannot become the object of a thought (Jolley, 2007, 59)

> Sellars thinks that he has found a way of forming sentences such that something other than a singular term is the grammatical subject of the sentence (Jolley, 2007, 67).

For Jolley, then, the very idea of making concepts into logical subjects constitutes a neglect of the absoluteness of the concept/object distinction. A proper understanding of that distinction is such that it renders incoherent the very idea of making judgments about concepts.

I wish to show that this is a mistake, which renders impossible a proper understanding of Frege's conception of the logical stratification of the realm of *Bedeutungen*. Notice, to begin, how Jolley has a tendency to slide between talk about the *notion* of an object or a concept, and talk about *objects* or *concepts*, or to advance statements that are ambiguous between the two. Consider, for instance, these two passages:

> Frege is unrestingly awake to the difficulties of understanding what he says about concepts. He notes early in his essay that his explanation of concepts 'is not meant

> as a proper definition'. The reason for this is that concepts are logical simples (Jolley, 2007, 1).

> As is made clear in Frege's debate with Kerry, as well as in other places in Frege's writing, concepts cannot be defined. If I attempt to make a concept a definiendum, I fail sharply to separate object and concept—even more, my attempt is defeated: the definiendum in a definition sentence is an object (Jolley, 2007, 95).

It is true that, in his debate with Kerry, Frege makes it clear that the *notion* of a concept cannot be defined because it is logically simple. Jolley glosses this by saying that concepts are logical simples. Now, it may seem quite uncharitable to read Jolley as saying that *particular* concepts are logical simples, so that particular concepts cannot be defined either. Yet it is precisely this reading that is confirmed in the second passage. In fact, however, nothing could be further from the truth: Frege's Grundgesetze definitions *are* definitions of concepts, and Frege explicitly discusses the requirements on a definition of a concept in passages such as (GGA, II, §56).

This is not an isolated occurrence. Here is Jolley again:

> For Wittgenstein, both Frege's *objects* and *concepts* are formal concepts, i.e., each is a way of talking about the values of particular propositional variables. This has the result that neither *objects* nor *concepts* can be talked about (Jolley, 2007, 19).

What are formal concepts, are the *notions* of an object and a concept. In a way similar to the above passages, however, Jolley glosses this by saying that neither *objects* nor *concepts* can be talked about, which is ambiguous between a use of the terms 'objects' and 'concepts' to talk about the logical categories of objects and concepts, or about particular objects and concepts. Of course, Jolley cannot have believed that there is no such thing as talking about particular objects; he must mean that we cannot talk about the *notion* of an object. But the very fact that he has a tendency of phrasing his claims in these unhappy ways, is already revelatory of the way in which he does not clearly distinguish between the notion of a concept and an object, on the one hand, and particular concepts and objects, on the other hand, which leads him to make confused claims such as the above claim that concepts cannot be defined.

Surely, if one asked Jolley whether there is a distinction between the notion of a concept and particular concepts, he would say that it is obvious that there is. His careless way of speaking, however, reveals that the distinction does not appear to him to be logically salient, does not appear to him to matter all that much. What generates this

appearance, is precisely the fact that Jolley takes Frege to be committed to the claim that we cannot make judgments about concepts, that concepts cannot be logical subjects. This results in the idea that *neither* the logical category of concepts *nor* particular concepts can be logical subjects, so that the distinction between them comes to seem as being of lesser importance. Both statements that purport to be about the notion of a concept and statements that purport to be about particular concepts are, according to Jolley, inherently logically defective, failing to express what they are confused attempts to express. This is exactly the sort of blurring between statements such as 'Concepts are fundamentally different from objects' and statements such as 'The concept *horse* is realized' that we also found in Conant. It is a blurring that Frege himself would have regarded as logically disastrous, and which renders it impossible to see what Frege took to be at stake in the logico-philosophical innovations underlying the *Begriffsschrift*.

It could be replied that, if we can make judgments about particular concepts, then we can also make general judgments about concepts, and that this gives us judgments such as 'Concepts are fundamentally different from objects'. This leads to the same sort of idea already discussed in the previous chapter: that a statement such as '$(\forall x)(\forall F)\big(x \neq_{\varepsilon} F(\varepsilon)\big)$' articulates the fundamental distinction between concepts and objects. For the reasons presented in that chapter, such generalizations cannot, however, give us statements about the *logical categories*. We do not get at the logical categories by generalizing, and no *Begriffsschrift* judgment serves to capture Frege's contrastive elucidatory use of the logical category terms. To appreciate the depth of the difference between statements such as 'Concepts are fundamentally different from objects' and statements such as 'The concept *horse* is realized', is precisely to appreciate how the possibility of taking the latter as a second-level subsumption in no way threatens to undermine the elucidatory status of the former, in no way threatens to undermine the distinction between concept and object itself.

Against Jolley, I claim that Frege thought that we can make judgments about concepts, and that concepts can be logical subjects. I have said that, in first-level subsumptions, an object is the logical subject, whereas in second-level subsumptions, a first-level concept is the logical subject. Jolley could object, however, that I am thereby using the notion of a logical subject in a way that transgresses the logical category distinctions, applying it both to objects as they figure in first-level subsumptions and to first-level concepts as they figure in second-level subsumptions. In Frege's logico-philosophical framework, the

objection continues, there is no room for such notions. In other words: the notion of a logical subject must be limited to objects after all, in accordance with Jolley's reading.

To say that Frege's philosophical logic entails that the notion of a logical subject must be limited to objects, is to regard that notion as a scientific notion. It is to claim that any legitimate use of the term 'logical subject' must be such that its range of application is limited to a single logical category, as is the case for any legitimate scientific use of a term. But the notion of a logical subject is not a scientific notion, it is a logico-philosophical notion. Logical subjecthood is exactly the sort of fundamental logico-philosophical phenomenon that constitutes a subject matter for *elucidation* rather than judgment, in such a way that the proper elucidatory use of the term 'logical subject' is indeed not limited to a single logical category. Rather than this being a logical defect in its elucidatory use, it is *essential* to it.

This is not an isolated phenomenon. Frege's philosophical logic involves a whole cluster of fundamental logico-philosophical terms of art that are used in such ways. Below, we will consider '*Bedeutung*' as another example. For now, I wish to point out that the same is true of the logical category terms. This may seem surprising, given that each logical category subsumes, of course, only the elements of a single logical category. This is true, but this does not mean that their *elucidatory use* is also limited to a single logical category. We already saw this in Chapter 1: Frege's elucidatory use of the logical category terms is essentially *contrastive*. We need to be able to use the logical category term 'concept', for instance, to say *both* that *ξ is a horse* is a concept *and* that the concept *horse* is not a concept.

Thus, we find that it is essential to Frege's elucidatory use of certain fundamental logico-philosophical terms of art that they are used with an elucidatory range of application that is not limited to a single logical category. I will call such an elucidatory use a *logically stratified elucidatory use*. All the logical category terms have such a logically stratified elucidatory use. The same is true of terms such as 'function', 'argument', '*Bedeutung*', and – as I have argued – 'logical subject'.

Against my claim that Frege is committed to the logical stratification of subjecthood, it could be objected that he nowhere states that he is rethinking the notion of a logical subject in the way I have proposed, and that we nowhere find him saying that it is a cornerstone of the *Begriffsschrift* that it offers us the tools to make concepts – as well as objects – into logical subjects. It is true that 'logical subject' is a term that Frege is loath to use, for reasons he explains in the following passage:

> Hence, the words 'relation of a subject to a predicate' designate two quite different relations, according as the subject is an object or is itself a concept. Therefore, it would be best to banish the words 'subject' and 'predicate' from logic entirely, since they lead us again and again to confound two quite different relations: that of an object's falling under a concept and that of one concept being subordinated to another (PW, 120; NS, 130).

Frege says that the words 'subject' and 'predicate' lead us to confuse subsumption with subordination. This is the because these notions are wedded to the old subject/predicate logic, in which there is no room for Frege's higher-level subsumptions. Frege's worry is that, if he would say that a concept serves as the logical subject in a subordination between concepts, his readers will misunderstand him as thereby implying that, in such a subordination, the concept becomes an object. Since Frege's aim is precisely to show that subordination of concepts does not involve *Bedeutungen* that are objects, he wants to avoid such a misunderstanding at all costs. At the same time, he himself indicates in the first sentence of this very passage the possibility of using the notion of a logical subject in such a way that we can say both of an object and of a concept that it is a logical subject. He nevertheless decides to abjure such a use because of the confused associations with which the notion of a logical subject has been saddled in the old subject/predicate logic.

I have decided not to follow Frege in this, and to use the notion of a logical subject in articulating his views. One reason is that – although Frege does not use the notion in a systematic way – many of his interpreters latch onto it to make claims about Frege's philosophical logic that I wish to contest. We already saw how this is the case for Jolley. Similarly, we find Bengtsson saying that "[Frege] seems to preclude the possibility of making a first order concept the subject of a thought" (Bengtsson, 2018, 107).[19] She bases this claim on a passage where Frege says the following about the statement 'There is at least one square root of 4': "Although our sentence does not present the concept as a subject, it says something about it; it can be regarded as expressing that a concept falls under a higher one. But this does not in any way efface the distinction between object and concept" (CP, 189). The sentence in question is a second-order subsumption. In saying that, in that sentence, the concept *square root of 4* is not the subject, Frege is pre-empting precisely the danger I mentioned above: if he said that it *was* the subject, he would likely

[19] Compare also Black, who says about Frege's statement that functions are unsaturated that it "might be taken simply to mean that it is logically impossible to make a function the subject of an assertion" (Black, 1968, 240).

be misunderstood as implying that concepts can be objects. Because he is at pains to avoid such a confusion and to emphasize precisely his sharp distinction between concepts and objects, he makes his point by saying that the concept is here not the subject, in the sense that it is not an object. But he immediately adds that the statement in question does say something *about* the concept, in accordance with my reading.

It could be objected that this is all a merely verbal matter. If we use 'logical subject' in the sense of 'object', then of course it is true that concepts cannot be logical subjects. This, it could be said, is what Bengtsson is doing, and this does not block her from acknowledging any of the claims about Frege's device of second-order subsumptions that I have made, apart from the fact that she would not express them by saying that concepts are made into logical subjects. That the matter is not merely verbal, however, becomes apparent once we observe how Bengtsson's discussion continues after she has made the above statement.[20] She proceeds to discuss Frege's statement that the concept *horse* is not a concept. After quoting the grain of salt passage, she writes:

> We notice that Frege says that the expressions he uses when he wants to *say something about a concept* 'miss his thought'. The latter phrase suggests that expressions or sentences that he uses when 'he intends a concept but mentions an object' do not express thoughts, i.e., that they are nonsensical [my emphasis] (Bengtsson, 2018, 107).

Bengtsson combines her previous observation about the purported impossibility of making concepts into logical subjects, on the one hand, with Frege's claim that sentences that purport to say something about a concept miss his thought, on the other hand, to conclude that any attempt to say something about a concept must issue in nonsense, thereby again arriving at the sort of view that lumps together statements such as 'Concept are fundamentally different from objects' and statements such as 'The concept *horse* is realized'. In this way, she fails to do justice to the fact that Frege – in the same breath as saying that the concept *square root of* 4 is not a subject in the statement 'There is at least one square root of 4' – also says, without any qualification, that the statement in question is a statement *about* the concept *square root of* 4, and that this does not efface the distinction between concept and object. Frege is saying precisely that his sharp

[20] That the matter is not merely verbal in Jolley, was already clear from his claim that concepts cannot be defined.

distinction between concept and object is *compatible* with the claim that we can make statements about concepts, a claim that I – more emphatically than Frege – also wish to express by saying that we can make concepts into logical subjects, precisely to clearly position myself against readings such as Bengtsson's and Jolley's.

We find another version of the same misreading in Witherspoon, who writes:

> [Frege] forms sentences that appear to have a concept as their logical subject; but this appearance is deceptive, because concepts are to be distinguished from logical subjects. His sentences are in a certain sense ill-formed; they necessarily fail to express the distinction he wants to draw between concepts and objects. Nevertheless, he thinks that his sentences can serve as hints, which, if his readers meet him halfway, can lead them to grasp his crucial but unstatable insight (Witherspoon, 2002, 109).

Like Bengtsson, Witherspoon claims that concepts cannot be logical subjects. In the transition from the first to the second sentence in this passage, we see Witherspoon jump from a statement concerning sentences that appear to have a concept as their logical subject to the claim that *they* – those sentences – fail to express the distinction between concepts and objects. This is an even more blatant confusion between statements about the logical category of concepts, on the one hand, and statements about particular concepts, on the other hand. The former are elucidations and must be clearly distinguished from expressions of judgment. The latter, on the other hand, *are* expressions of judgment.

Finally, it must be noted that – although Frege does not systematically use the notion of a logical subject in the way that I do to articulate his views – he does systematically use the notion of *aboutness* in this way, having no qualms about saying that we can make judgments about concepts. One need but recall his claim from the *Grundlagen* that "a statement of number is an assertion about a concept" (GL, §46) to realize this. That Frege does not use the notion of a logical subject in this way is, to repeat, because he fears that – if he were to say that concepts can be logical subjects – he would be misunderstood as saying that concepts can be objects. Ironically, this is how Jolley, Bengtsson, and Witherspoon misunderstand him, albeit in a converse way. That is to say: they take Frege's reluctance to say that concepts can be logical subjects to reveal that Frege believed that such a view would *in fact* obviate the concept/object distinction, whereby they take him to be committed to the sort of amalgamation of logical subjects and objects characteristic of the old subject/predicate logic. They take Frege's not saying that

concepts can be logical subjects to amount to him saying that concepts cannot be logical subjects, a fateful shift which, I believe, makes it impossible to achieve clarity about Frege's philosophical logic.

2.5 The functionality of the realm of *Bedeutungen*

One could reply to my discussion that Frege's logically stratified use of logico-philosophical terms of art such as 'function' and 'about' should be understood as merely pragmatic, in something like the following way: Frege realizes that such a use is logically improper, but it would be too cumbersome to disambiguate such terms in all contexts, so he continues to use them in this logically improper way, counting on his readers to grant him another grain of salt. In a logically correct presentation, Frege's statements would have to be categorically disambiguated. Instead of talking, for instance, about the unsaturatedness of functions as such, we can only talk about the first-level-single-argument-unsaturatedness of first-level-functions-of-one-argument, the second-level-single-first-level-argument-of-one-argument-unsaturatedness of second-level-functions-of-one-first-level-function-of-one-argument, and so on, where each of these are categorically different notions whose range of application is limited to a single logical category.

This is, in effect, to once again conceive of such terms on the model of primitive scientific terms, assimilating them as closely as one can to terms that straightforwardly have *Bedeutung*, and thereby to assimilate their elucidatory use as closely as one can to the expression of judgment. I take it to be crucial to a correct understanding of Frege's philosophical logic to see that such an approach is out of the question for him. Frege nowhere suggests that his logically stratified elucidatory use should be disambiguated in this way. This is no accident: such a disambiguation would rob that use of its elucidatory function. On such a conception, terms such as 'first-level function' and 'second-level function' would be construed as merely equivocal. There would be no overarching logico-philosophical unity underlying Frege's logically stratified use of the term 'function' to characterize both first-level functions and second-level functions, since any purported such unity would have to be dismantled through disambiguation. But this means that the logical categories of first-level functions and second-level functions come to be logically

independent of each other. That first-level functions are unsaturated, for instance, is now to be seen as completely logically independent from the question whether second-level functions are also unsaturated. Indeed, this way of putting the matter already inserts too much unity into the matter, since this use of the term 'also' neglects the fact that the notion of unsaturatedness would also have to be categorically disambiguated. What first-level functions are in being unsaturated, is logically entirely independent of what second-level function are in being unsaturated.

Frege would have been appalled by such a view, let alone by the suggestion that it was his. Rather than saying that he should have disambiguated his logically stratified use of his terms, we should regard him as using those terms in exactly the way in which he believed they *should* be used. That is: we should regard their logically stratified use as essential to the elucidation of the logical stratification of the realm of *Bedeutungen*, and with it the logical stratification of a whole cluster of interconnected logico-philosophical phenomena, such as logical subjecthood, *Bedeutung*, unsaturatedness, and so on. The logical stratification of the realm of *Bedeutungen* is not such as to result in a handful of logically independent logical categories[21] – and a concomitant handful of logically independent notions of unsaturatedness, aboutness, and so on – but rather characterizes the logico-philosophical *unity* of the logical categories and thereby of the realm of *Bedeutungen* itself. To say that a first-level subsumption is *about* an object, whereas a second-level subsumption is *about* a first-level function, to say that an object can be an *argument* of a first-level function whereas a first-level function can be an *argument* of a second-level function, to say that both first-level functions and second-level functions are *functions*, such logically stratified elucidatory use of the terms 'about', 'argument', and 'function' brings out the unity of the logical categories, a unity that is characterized by what I will call the *functionality* of the realm of *Bedeutungen*. Functionality *is* the mode of logical stratification of the *Begriffsschrift*, and it consists in nothing else than the logically stratified function/argument nexus that lies at the core of Frege's philosophical logic and that replaces the old subject/predicate nexus in which anything that is a logical subject must *ipso facto* be an object. To understand Frege's conception of the logical categories, one must see how the logical stratification of the function/argument nexus is essentially tied to the logical stratification of logical subjecthood that I have sought to articulate, so

[21] The realm of *Bedeutungen* is not a layer cake!

that the function/argument nexus is not limited to first-level functions and objects, but rather constitutes – in its logical stratification that yields the functionality of the realm of *Bedeutungen* – the way in which all logical categories are essentially interconnected. Rather than standing in need of disambiguation, Frege's logically stratified elucidatory use of his terms of art serves to bring out this form of unity of the realm of *Bedeutungen*, serves to bring out its functionality.

Note, finally, that Frege himself has ways of marking this unity of the realm of *Bedeutungen*. Consider, for instance, the following passage:

> The relation of an object to a first-level concept that it falls under is different from the (admittedly similar) relation of a first-level to a second-level concept. (To do justice at once to the distinction and to the similarity, we might perhaps say: An object falls *under* a first-level concept; a concept falls *within* a second-level concept.) (CP, 190; KS, 174; 201).

We have here the distinction between a first-level subsumption and a second-level subsumption. What Frege wishes to make clear, is that this distinction harbours an underlying unity, a unity that is marked by Frege's two-fold use of the verb 'to fall' and his notion of similarity. Elsewhere, Frege talks about "how a concept can be related to a *second level concept* in a way analogous to that in which an object is related to a concept under which it falls" (PW, 254; NS, 275). Again, Frege's use of the term 'analogous' is meant to bring out the functionality of the realm of *Bedeutungen*. We will encounter further examples in later chapters.

2.6 The logical stratification of *Bedeutung*

Let us consider another fundamental logico-philosophical phenomenon: *Bedeutung*. Here is a characteristic sort of remark from Wright:

> Nowhere does it seem cleanly to have been noticed that Frege was simply never at liberty to introduce *Bedeutung* into the semantics of predication, at least not if that is to involve having *the very same relation* link predicates and concepts as ties a singular term to its bearer (Wright, 1998, 254-255).

Since Frege's logic precludes the idea that it is the same relation of *Bedeutung* that relates proper names to objects and predicates to concepts, Wright concludes that there is something logically defective about Frege's use of the term '*Bedeutung*', a use that is logically stratified in the way just discussed.

Wright seems to think that we have here a fateful oversight by Frege, who did not realize that he is not at liberty to use his terms in this way, given his own views.[22] I submit that it is ludicrous to think that Frege overlooked the fact that he was using such terms as '*Bedeutung*' in ways that are categorically ambiguous. Rather, we must conclude that Frege self-consciously employed his terms in this way, and that he took this to be essential to the elucidatory use to which he wished to put them. Indeed, passages such as the following show that Frege was very much aware of Wright's objection:[23]

> We should really outlaw the expression 'the meaning of the concept-word A', because the definite article before 'meaning' points to an object and belies the predicative nature of a concept. It would be better to confine ourselves to saying 'what the concept word A means', for this at any rate is to be used predicatively: 'Jesus is, what the concept word 'man' means' in the sense of 'Jesus is a man' (PW, 122; NS, 130).

Frege's acknowledgment that an expression such as 'the meaning of the concept-word A' faces the same problem as expressions such as 'the concept *horse*' is an acknowledgment of Wright's objection. At the same time, his response is not to withdraw that use, or to disambiguate it, since he proceeds to use the very same term in his proposed expression 'what the concept word A means'. The exact same objection would apply here: 'means' cannot mean the same in 'what the concept word A means' as in 'what the proper name B means'. This shows that Frege regards his logically stratified elucidatory use of the term '*Bedeutung*' to be justified *in spite* of Wright's objection, which reveals that we do not have here the sort of use that is to be modelled on the use of scientific terms with *Bedeutung*, where categorical ambiguity must indeed be ruled out at all costs. To describe Frege's

[22] Wright adds, moreover, that "*almost* all of more than 100 years of sophisticated secondary literature on this topic" [original emphasis] has missed this point (Wright, 1998, 254-255n). Wright is mistaken. To give just one prominent example: Dummett – who figures centrally in Wright's account – was aware of the point: "The relation of reference between, for instance, a relational expression and the relation for which it stands is not supposed by Frege to be the same relation as that between a name and its bearer, but only an analogous one" (Dummett, 1981, 171). For more recent examples, see (Kimhi, 2018, 80n), as well the passage from Jolley that I will discuss shortly.

[23] See also (PMC, 136-137).

logically stratified elucidatory use as 'categorically ambiguous' is already misleading, since the latter notion only applies where it is *judgment* that is at issue, not elucidation. Because elucidation is not judgment, logically stratified use is not categorically ambiguous use.

Wright's position can be characterized as a position that has no room for my notion of logically stratified elucidatory use, a mode of use that must be conceived as fundamentally different from the scientific use of terms with *Bedeutung* in the expression of judgment. That Frege extensively engages in such elucidatory use, however, reveals that such use plays an essential role in his philosophical logic. As discussed in the first chapter, elucidation must be taken as a *sui generis* activity that is not to be modelled on judgment. Frege would see the danger as lying precisely *in* any attempt to blur the boundary between logico-philosophical elucidation and judgment, regarding as deeply confused the idea that his logically stratified elucidatory use of terms such as 'function' should be assimilated as closely as possible to the scientific use of terms with *Bedeutung* in the expression of judgment.[24]

As could be expected from my prior discussion, Jolley formulates a similar complaint:[25]

> Frege's talk of properties as the *Bedeutungen* of predicates is itself waffly. The absolute distinction between concepts and objects should suffice to worry anyone about this talk. If both have *Bedeutungen*, then they are congeneric: but then they are not absolutely distinct. No concept-phrase can stand for something in a sense univocal with that in which an object-phrase can (Jolley, 2015, 124n).

Frege's talk is only waffly, however, if he means to use the term '*Bedeutung*' in a way that is suitable to scientific terms with *Bedeutung*. Seen in this way, Frege's talk about *Bedeutung* will appear as if it has to be categorically disambiguated in the way discussed above.

---

[24] That philosophy should not be modelled on science is, of course, a claim with a venerable tradition, a tradition in which, we now see, Frege has a role to play.

[25] Compare also: "The [concept *horse* paradox] engages with an ineradicable tendency of ours: a tendency to understand all genuine distinctions as, somehow and (in some acceptation of 'ultimate') ultimately, distinguishing between objects and (in some acceptation of 'ultimate') ultimately, distinguishing between objects of which we can (must) take cognitive—perceptual or conceptual—grasp" (Jolley, 2007, 50). There is a strange way in which Jolley's statement turns back on himself. Jolley refuses to recognize a distinction between objects appearing as logical subjects and concepts appearing as logical subjects. Why? Because he believes that such a distinction would make concepts into objects. But this means that it is *Jolley* who is refusing to understand as a genuine distinction anything that is not a distinction between objects. In a similar way, he refuses to understand a distinction between objects as *Bedeutung* and concepts as *Bedeutung*, since he believes that this would equally turn concepts into objects.

Others have raised similar worries about the fact that Frege uses the term 'name' to characterize both proper names and function names. Here is Max Black:[26]

> Frege, in spite of [the] clear distinction in the uses of [proper names and function names], still wants to think of the functional expressions as *names* for functions (cf. his use of the term '*Functionname*' and his tendency to say that function names and function expressions 'stand for' something). Of course, if one intends to use the term 'name' in such a way that 'objects' can, but functions cannot, have names, and yet thinks of the functional expression as after all still, somehow, naming a function, confusion will be bound to result; one will try to speak of functions as if they were and at the same time were not 'objects.' The use of the figurative expressions '*unvollständig*' and '*ungesättigt*' is an attempt, but an unsuccessful one, to have the matter both ways—to call functions objects, albeit peculiar, 'unsaturated,' ones (Black, 1968, 246).

The point is the same: Black is neglecting the fact that Frege's use of terms such as 'name' and 'stand for' is a logically stratified elucidatory use that does not commit Frege to the claim that all names, or that what they stand for, belong to a single logical category.

A proper understanding of Frege's philosophical logic requires us to see that there is a whole cluster of fundamental logico-philosophical phenomena that are logically stratified. These include *Bedeutung*, subsumption, unsaturatedness, logical subjecthood, and many others. The logical stratification of the realm of *Bedeutungen* reverberates through each of these phenomena. What is at stake in the elucidation of the logical categories, then, is not just an understanding of the logical stratification of the realm of *Bedeutungen* as such, but equally an understanding of how each of these phenomena is in turn logically stratified. We have what may be called an *elucidatory cluster of logically stratified phenomena* at the basis of which lies the fundamental logico-philosophical phenomenon of the functionality of the realm of *Bedeutungen*.

Although there are certainly relevant differences between these logically stratified phenomena, the activities of elucidating them can nevertheless, I think, be said to form what is properly called a *genus* of logico-philosophical elucidation. That is to say: we have here a sense of 'elucidation' – the logico-philosophical elucidation of logically stratified phenomena – that can properly be said to capture a more or less univocal notion of

[26] Compare also (Klement, 2004, 8).

elucidation, a notion of elucidation that pertains to each of these logically stratified phenomena without thereby ramifying into different senses of 'elucidation' in the way in which the notions of 'elucidation of primitive scientific terms' and 'elucidation of the logical categories' pertain to different senses of 'elucidation', as discussed in the previous chapter.

What is at issue in the logical stratification of each of these phenomena, is what I above called the functionality of the realm of *Bedeutungen*. Wright, Jolly, or Black have no way to properly acknowledge that functionality as it figures in the work of Frege, since they have no way to properly acknowledge the fact that Frege takes the function/argument nexus of judgment to be logically stratified, giving rise to logically stratified notions of 'function', 'unsaturatedness', 'argument', 'subsumption', and so on. That a first-level subsumption is *about* an object and a second-level subsumption *about* a first-level function just is one way of bringing out that *both* subsumptions involve a function/argument nexus, involve the completion of what is unsaturated by a suitable argument, and so on, and thereby of bringing out the form of logico-philosophical unity that characterizes the realm of *Bedeutungen*, a unity without which we fall into the disastrous idea that each of the logical categories are to be regarded as constituted logically independent of each other, in such a way that all attempts to bring out their unity are to be regarded as logically defective and calling for disambiguation. A proper understanding of this unity requires taking Frege's logically stratified elucidatory use of his terms as the *sui generis* use that it is, which in turn requires taking the elucidation of the logical categories as the *sui generis* activity that it is, so that the former is not modelled on scientific use and the latter is not modelled on judgment.

2.7 Russell's principle

Let us return to Russell's principle that "every constituent of every proposition must, on pain of self-contradiction, be capable of being made a logical subject" (PoM, §52).[27] It will be clear that readers such as Jolley, Bengtsson, Witherspoon, and Wright take Frege to disagree with Russell. Above, I claimed that there is an important sense in which Frege *agrees* with Russell's principle. I can now make good on that claim.

Russell believes that, for something to appear as a logical subject, it must *ipso facto* appear non-predicatively. Once this assumption is in place, the principle that every constituent of every proposition must be capable of being made a logical subject entails that every constituent of every proposition must be capable of appearing non-predicatively in a proposition. From this, it follows that Russell's logical category of *terms* – which encompasses all entities that can appear non-predicatively in a proposition – must be taken to encompass all possible constituents of all propositions. On Russell's understanding, then, the principle entails that every constituent of every proposition must – whatever else it might be – *also* be a term. All constituents are thereby drawn into the same basic logical nexus of term/assertion, which is another version of the old subject/predicate nexus. For all of Russell's insights – and for all his emphasis on the importance of relations – there is this crucial sense in which he remains wedded to the old subject/predicate logic.[28]

Frege, on the other hand, rejects Russell's assumption that logical subjects must always appear non-predicatively. Instead, he takes logical subjecthood to be logically stratified, so that logical subjects need not be objects. On this renewed understanding, however, it is *also* correct to say that every constituent of every proposition must be capable of being made a logical subject. For each constituent, there is a logical category of functions that

---

[27] Compare: "Contrary to an ancient opinion, we can say what is not. We can also speak of what is not an individual. Everything, whether or not an individual, is a possible subject of discourse, though not every subject of discourse is an individual" (Valberg, 1971, 135). Recall that Valberg is one of Jolley's targets. Notwithstanding genuine problems with Valberg's exposition – which I cannot discuss in detail here – Valberg's way of unpacking his statement is closer to Frege than to Russell. Jolley's qualms with Valberg's account are, in the main, Russellian qualms that are foreign to Frege.

[28] To what extent this remains true of Russell's later views is an issue that I cannot investigate here. What is true, is that Russell continued – for some time – to stick by an understanding of his principle that is importantly different from Frege's. See (Klement, 2004) for an excellent discussion of the role played by that principle in Russell's philosophy before his encounter with Wittgenstein.

takes that constituent as argument, giving rise to judgments with that constituent as logical subject.

I believe it is important to emphasize this sense in which Frege accepts the principle, because it comes with a proper understanding of the logical stratification of subjecthood, giving us a clear account of how exactly Frege's approach differs from Russell. In this way, we can say that both Frege and Russell accept the principle, while making clear that the logico-philosophical views underlying that acceptance – and thereby their understanding of the principle itself – are radically different.[29]

It is a commonplace that Frege sought to replace the subject/predicate nexus of judgment with the function/argument nexus[30]. One has not grasped the nature of this replacement, however, unless one has grasped the functionality of the realm of *Bedeutungen* as I have articulated it above, a functionality that involves the logical stratification of a whole cluster of logico-philosophical phenomena and that characterizes the logico-philosophical form of unity of the realm of *Bedeutungen*. We have seen how readers such as Jolley, Bengtsson, Witherspoon, and Wright fail to see this. It seems correct to say that Frege invented this notion of functionality.[31] In any case, it constitutes a cornerstone of his philosophical logic.[32]

---

[29] It could be said that they both accept the principle *qua* form of words, but understand the words differently. Still, we must be able to make sense of the statement that Frege and Russell ascribe to a different understanding of the *same* principle in a way that goes beyond the mere observation that we have the same sensible form of words. There is a logico-philosophical unity to their different ways of understanding this form of words that cannot be captured in terms of mere equivocity. This form of unity is, in turn, reflected in the unity of our capacity, as historians of philosophy, to attain a comparative understanding *of* Frege's and Russell's way of understanding this principle, and by extension a comparative understanding of their philosophical logic in which their respective views are understood together in a way that does not render them merely incommensurable. To my knowledge, the monumental (Conant, 2020) is the work in the history of philosophy that most saliently exhibits this capacity of the historian of philosophy, by uniting within one overarching historical mode of understanding the logico-philosophical views of such diverse thinkers as Descartes, Leibniz, Kant, Frege, and Wittgenstein, to name just those that figure most prominently in Conant's account. Understanding the form of unity that characterizes such historical understanding – and thereby the nature of the activity of the historian of philosophy – is, I believe, the hardest problem in philosophy.

[30] Note that I let 'function' replace 'subject' as first member. For an account of why this better captures some aspects of Frege's replacement of subject/predicate with function/argument – by placing Frege's views against the background of Kant's Copernican revolution – see (Van de Vijver, 2013).

[31] Which is not to say that there are no parallels with other philosophers. In a sense, it could be said that the phenomenon of logical stratification is already thematized as early as in the works of Aristotle, whose notions of 'being' and 'form' – to give just two prominent examples – may be said to be logically stratified. Of course, there are deep differences here, and some may already object that my using the same term 'logical stratification' to characterize aspects of Aristotle's views is bound to be misleading. See (Kimhi, 2018) for some relevant reflections.

[32] Frege's logico-philosophical innovations also come with a new understanding of logical generality, a complex issue that I cannot discuss in detail here. Let me just make some statements. The notion of 'generality' is folded

2.8 Frege's dialectic with Kerry

I started this chapter with a discussion of the notorious grain of salt passage. The time has come to return to it a final time, and to investigate in more detail the nature of the surrounding dialectic between Frege and Kerry. It is well-known that Kerry objects to Frege's sharp distinction between concept and object by advancing a purported counterexample: the statement 'The concept *horse* is a concept easily attained'. As Kerry sees it, the expression 'the concept *horse*' is here used to refer to a concept. But it is a proper name, so that it also refers to an object, showing Frege's distinction to be unfounded. Notoriously, Frege replies: "Quite so; the three words 'the concept 'horse'' do designate an object, but on that very account they do not designate a concept, as I am using the word" (CP, 184). In other words: "the concept *horse* is not a concept" (CP, 186).

It is true, of course, that Kerry rejects Frege's sharp distinction between concept and object. The important question, however, is *why*? What is it, exactly, that Kerry finds so problematic about Frege's absolute distinction between concept and object? Insofar as that question is taken up at all in the literature, Kerry is often portrayed as falling prey to some sort of naive psychologistic conception of meaning, according to which the mere fact that we use the term 'concept' in the expression 'the concept *horse*' suffices to have it refer to a concept.[33] Frege is then presented as the more enlightened philosopher who

---

into Frege's logically stratified notion of subsumption. In its simplest form, a general judgment about objects comes to be understood as a second-level subsumption, i.e. a judgment about a concept. The relation between such a second-level subsumption and first-level subsumptions is encapsulated in Basic Law IIa (GGA, I, §20), which licenses the inference from '$(\forall x)(\varphi(x))$' to '$\varphi(a)$'. Such inferences can only be properly understood if it is understood how the former statement is *about* the first-level function that features in the latter. Frege's approach to logical generality is in turn connected to the way in which he seeks to overcome the empiricist reliance on abstraction. As empiricists see matters, our capacity for general judgments must be grounded in our capacity for particular judgments. In Fregean terms, this would mean that our capacity for second-order subsumptions must be grounded in our capacity for first-order subsumptions. But such a claim makes no sense from Frege's point of view, since these capacities constitute inseparable moments of our overarching capacity for judgment as such, a capacity the unity of which allows for no such empiricist division. Compare: "Frege argued vigorously against any notion of abstraction as needed to get from particulars to general notions. Indeed, elimination of any role for abstraction is central not just to Frege's antipsychologism, but also to his anti-Kantianism. To eliminate abstraction is to eliminate the question, how do we attain the general? Frege replaces it with the question of the relation between the (already given) general and the particular, a question to be answered by logic" (Goldfarb, 2001, 33). What is 'already given', is our capacity to make second-level judgments about concepts. One need only compare Frege's conception of generality with Russell's early views about denoting from *Principles of Mathematics* to develop a sense of the depth of the difficulty of these issues. Russell's view that terms such as 'all men' must either denote an ambiguous object or ambiguously denote several objects (PoM, §62) again betrays an important affinity between his views and the old subject/predicate logic: because 'all men' can occur as subject, it *must* be understood on the model of denoting an object.

[33] A good example is (Jolley, 2007, 13).

realizes that the logic of our language is not such as to yield to our psychological whims. We cannot, by sheer will, make a proper name refer to a concept. The very idea rests on a confused admixture of the psychological and the logical.[34]

Against such accounts, I would like to propose what I consider to be a more interesting reading of the dialectic, which makes Kerry's objection come out as more principled and more germane to Frege's own concerns than has hitherto been allowed.[35] It is a reading that becomes possible once one has clearly delineated, as I have tried to do, the issue of making judgments about particular concepts as a distinct moment in Frege's philosophical logic. Rather than seeing Kerry's objection as arising from a naive psychologistic conception of meaning, we should see it as arising from a concern with this issue. Kerry's own purported counterexample – the statement that the concept *horse* is a concept easily attained – constitutes an attempt to say something about a particular concept, i.e. the concept *horse*. It does not yield the objection that Frege cannot coherently state the concept/object distinction as such. About this latter issue, Kerry has little to say – unlike Russell, as we will see in the next chapter. As I understand Kerry's worry, it is the following: if Frege's concept/object distinction is accepted, we can no longer make judgments about concepts. The importance of Kerry's statement that the concept *horse* is a concept easily attained, does not lie in the bare technical point that the expression 'the concept *horse*' seems to do double duty as a proper name and a concept name. It rather lies in the fact that – as Kerry sees it – this is what our expressions *must* be capable of doing if we are to be able to make judgments about concepts at all.

The only manner of talking about concepts that Kerry can think of consists in having them appear non-predicatively as the *Bedeutung* of a proper name. Because Frege's absolute concept/object distinction precludes this, that distinction appears to Kerry to constitute a rejection of the very possibility of making judgments about concepts, and Kerry believes that no coherent philosophical logic is possible once that possibility is

[34] It is interesting to note that what I have called Kerry's purported naive psychologistic idea of meaning is in fact encountered in many discussions of the dialectic. It underlies the oft-made claim that Frege's statement that the concept *horse* is not a concept is *paradoxical*. Hale and Wright, for instance, talk about "phrases like 'the concept *horse*' which intuitively ought to stand for concepts" (Hale & Wright, 2012, 85). Beaney talks about "the natural way of referring to concepts" (Beaney, 1996, 189). Burge characterizes Frege's claim that 'the concept *horse*' does not stand for a concept as "deeply counterintuitive" (Burge, 2005, 21). Geach says that 'the concept *horse*' "would have to stand for a concept if it stood for anything" (Geach, 1961, 156). Compare also (Parsons, 1986, 450) (Wright, 1998, 245, 257) (Textor, 2010, 128) (Proops, 2013, 76, 78) (Travis, 2020, 245).

[35] I should flag that my concern is with Kerry's objection as Frege understood it. Thus, my claim is that Frege himself took Kerry's objection to be more principled than has been allowed. For a discussion of Kerry's own views, see (Picardi, 1993).

rejected. Kerry's challenge to Frege, then, is to address the seeming fact that his concept/object distinction precludes the possibility of making judgments about concepts.

It could be objected against my proposal that Frege's reply does not address Kerry's challenge as I have just construed it, so that this cannot have been how Frege understood Kerry's objection. I disagree. There is, in fact, a long stretch of *Über Begriff und Gegenstand* that is concerned precisely to show that Frege's concept/object distinction does not preclude the possibility of making judgments about concepts, that is – in other words – concerned precisely to show how to reconcile the following two statements:

(1) Concepts are essentially predicative.
(2) Concepts can be made into the logical subjects of judgments.

Frege begins his reply to Kerry by elucidating the concept/object distinction itself and by elucidating how an adequate grasp of that distinction precludes an understanding of Kerry's statement 'The concept *horse* is a concept easily attained' as a counterexample to that distinction – an endeavour in which Frege's claim that the concept *horse* is not a concept plays a central role. All of this is standard fare. Frege does not stop there, however. Having shown how a correct understanding of the concept/object distinction reveals that Kerry's sentence does not constitute a counterexample to that distinction, Frege continues:

> Thus Kerry does not succeed in filling the gap between concept and object. Someone might attempt, however, to make use of my own statements in this sense. I have said that to assign a number involves saying something about the concept; I speak of properties ascribed to a concept, and I allow that a concept may fall under a higher one. I have called existence a property of a concept (CP, 188; KS, 172-173; 199).

In the first sentence, Frege summarizes what he has just done, i.e. to show that Kerry's purported counterexample fails to undermine the concept/object distinction. Frege acknowledges, however, that this cannot be the full story, because he himself seems to employ many statements just like Kerry's, in which something appears to be said about a concept. This generates the impression that Frege is caught in a dilemma: either he sticks to the claim that Kerry's statement 'The concept *horse* is a concept easily attained' does not say anything about a concept, but then neither do his own statements, which undermines the coherence of his logico-philosophical discussions. Or, he claims that his own statements *do* succeed in saying something about concepts, but then it seems that

the same must be granted of Kerry's statement. The burden of the ensuing discussion is to disarm this dilemma. If my reading was correct, one would expect Frege to do this by discussing the difference between first-level subsumptions and second-level subsumptions, thereby clarifying how 'saying something about a concept' should be understood in terms of second-level subsumptions in which the concept retains its predicative nature, rather than first-level subsumptions in which we only have objects. This is exactly what Frege does. He starts by saying:

> How I mean this to be taken is best made clear by an example. In the sentence 'there is at least one square root of 4', we are saying something, 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 being said. But notice carefully that what is being said here is not the same thing as was being said about the concept (CP, 188; KS, 173; 199).

Frege announces that he will explain how we must take his own discourse in which statements just like Kerry's seem to be made. He does this through the discussion of an example – the sentence 'there is at least one square root of 4' – which he says must be construed as saying something about a concept. He contrasts this sentence with the sentence 'The concept *square root of* 4 is realized', which is not about a concept, but about an object. Frege then adds that what is said about the object in the latter is not the same as what is said about the concept in the former. In saying this, Frege is starting to point us towards the distinction between a first-level subsumption and a second-level subsumption. In bringing us to recognize that these statements say something different about their respective subjects, we are starting to see how the mere fact that the former is about a concept does not entitle us to think that the concept must thereby appear non-predicatively, as in the latter. *Even if*, Frege adds for rhetorical purposes, we would wish to say that these statements express *the very same thought*, we would still be forced to say that they differ with respect to what is being said of what.[36] In other words: the mere fact

[36] I say 'for rhetorical purposes', because I do not think this claim of sameness of thought can be made coherent within Frege's broader philosophical logic. Indeed, it makes it impossible to make sense of Frege's claim that statements such as 'The concept *square root of* 4 is realized' miss his thought, a claim that is consistently repeated

that the statement that there is at least one square root of 4 is a statement about a concept, does not by itself undermine the concept/object distinction, a point that Frege himself explicitly makes:

> Although our sentence does not present the concept as subject, it says something about it; it can be regarded as expressing that a concept falls under a higher one. But this does not in any way efface the distinction object and concept (CP, 189; KS, 173; 200).

In his discussion, Frege explores in detail the logical difference between the two statements 'There is at least one square root of 4' and 'The concept *square root of* 4 is realized', emphasizing the different ways in which – in each of them – something is said about something:

> In the sentence 'there is at least one square root of 4' it is impossible to replace the words 'square root of 4' by 'the concept *square root of 4*'; i.e. what is suitably said of the concept does not suit the object [...] Hence what is here said concerning a concept can never be said concerning an object; for a proper name can never be a predicative expression, though it can be part of one. I do not want to say it is false to say concerning an object what is said here concerning a concept; I want to say it is impossible, senseless, to do so (CP, 189; KS, 173-174; 200).

Frege is elucidating precisely the logical difference between a first-level subsumption and a second-level subsumption, the logical difference between how something is being said about something in each of these kinds of statements.

That Frege's main concern does indeed lie with the distinction between first-level subsumptions and second-level subsumptions, becomes explicit when he concludes his discussion:

> What has been shown here in one example holds good generally; the behaviour of the concept is essentially predicative, *even* where something is being said about it; consequently it can be replaced there only by another concept, never by an object. Thus what is being said concerning a concept does not suit an object. Second-level concepts, which concepts fall under, are essentially different from first-level

throughout his *oeuvre*. I cannot present a full discussion of this issue here. For my purposes, what is important is the difference Frege wishes to emphasize between these statements, irrespective of whether they must be taken to express the same thought.

> concepts, which objects fall under. The relation of an object to a first-level concept that it falls under is different from the (admittedly similar) relation of a first-level to a second-level concept. (To do justice at once to the distinction and to the similarity, we might perhaps say: An object falls *under* a first-level concept; a concept falls *within* a second-level concept.) The distinction of concept and object thus still holds, with all its sharpness [first emphasis mine] (CP, 189-190; KS, 174; 201).

What Frege is at pains to show in these passages – as revealed nicely by the 'even' in the first sentence – is precisely how we can reconcile his concept/object distinction with the possibility of making judgments about concepts. He is at pains to show that he has a way of reconciling (1) and (2) above, and is thereby responding to exactly the challenge that, on my reading, Kerry's objection raises. As far as I can tell, this is all but completely missed in the literature, whereby it becomes impossible to see Frege's remarks in the right light. It is no coincidence that – notwithstanding the vast body of literature on Frege's dialectic with Kerry – these particular passages have received comparatively little attention.[37]

Frege started his discussion with the promise that he would explain how to understand such claims of his as the claim that ascriptions of number are statements about concepts. Frege's explanation consists precisely in the point that statements about concepts are second-level statements, and that properties of concepts are second-level concepts, so that the fact that ascriptions of number are (second-level) statements about concepts does not efface the concept/object distinction[38].

Let us take a final look at the grain of salt passage:

---

[37] When readers refer to them, this is usually with regards to Frege's claim that 'There is a square root of 4' and 'The concept *square root of* 4 is realized' express the same thought, a remark that is then made to do some heavy lifting to support the claim that Frege believed that thoughts in general have no set logical form that determines the ways in which they are decomposable, as in (Travis, 2020, 271). There is, however, little in Frege's *oeuvre* that supports such a claim, and much that speaks against it. One further passage that may be taken to support this claim occurs where Frege says that a statement of the form '$(\forall x)\big(\varphi(x) = \psi(x)\big)$' expresses the same thought as '$\grave{\varepsilon}\varphi(\varepsilon) = \grave{\varepsilon}\psi(\varepsilon)$' (CP, 143). Note, however, that this claim of sameness of sense is absent from the later *Grundgesetze* discussions of Basic Law V. Moreover, even if these claims of sameness of sense express Frege's considered view, it must be remembered that they pertain to a very special case, where what is at issue is the conversion of judgments about functions into judgments about their courses-of-values (and *vice versa*). This does not warrant the more general claim that thoughts as such may be multiply decomposable in all sorts of ways.

[38] Frege also invokes the distinction between what he calls a *mark* and a *property* of a first-level concept, where the former yields another first-level concept that is subordinated to it, while the latter yields a second-level concept under which it falls (CP, 190). See also (GL, §53). There is a historical tradition – prior to Frege – of the use of such notions in logic, a tradition in relation to which Frege seeks to position his own logico-philosophical innovations. Frege's use of the notion of a 'property' as it applies to concepts, for instance, seems to be genuinely novel. In any case, this is a topic that warrants further investigation. I am grateful to professor Jean-Philippe Narboux for bringing this point to my attention.

> I do not at all dispute Kerry's right to use the words 'concept' and 'object' in his own way, if only he would respect my equal right, and admit that with my use of terms I have got hold of a distinction of the highest importance. I admit that there is a quite peculiar obstacle in the way of an understanding with the reader. By a kind of necessity of language, my expressions, taken literally, sometimes miss my thought; I mention an object, when what I intend is a concept. I fully realize that in such cases I was relying upon a reader who would be ready to meet me half-way – who does not begrudge a pinch of salt (CP, 193; KS, 177; 204).

We can now pinpoint more precisely the 'peculiar obstacle' that Frege has in mind. Frege says that ascriptions of number are statements about concepts. Still, when such ascriptions of number are advanced in his discussions, this is usually done in ordinary language, by employing expressions such as 'the concept *square root of* 4'. Here is an example from *Grundlagen*: "If I say 'The King's carriage is drawn by four horses', then I assign the number four to the concept 'horse that draws the King's carriage'" (GL, §46). Thus, although Frege means to make it clear that ascriptions of number are second-level statements, his own presentation makes them appear as first-level statements. As Frege puts it: "By a kind of necessity of language, my expressions, taken literally, sometimes miss my thought; I mention an object, when what I intend is a concept" (CP, 193). Frege's own statements make it appear as if he is not adhering to his concept/object distinction. Frege adds: "I fully realize that in such cases I was relying upon a reader who would be ready to meet me half-way – who does not begrudge a pinch of salt" (CP, 193). What he means, is this: a discerning reader, who is ready to meet Frege half-way, will acknowledge that Frege – in saying things such as that ascriptions of number are statements about concepts – has in mind second-level judgments, which are judgments about concepts. Such a reader will realize that seemingly first-level statements such as 'The concept *horse that draws the King's carriage* has the number four' must in fact be understood as second-level statements, in which a concept serves as the logical subject. Once this is properly understood, the whole suggestion that such statements – including Kerry's statement that the concept *horse* is a concept easily attained – undermine the concept/object distinction, evaporates.[39]

[39] A properly Fregean treatment of Kerry's statement would not only have to take into account the fact that it must be taken as a second-order statement, but also the fact that the notion of being easily attained is

In this way, I am committing myself to the claim that, in the end, Frege wished to construe statements such as 'The concept *horse that draws the King's carriage* has the number four' as misleading expressions of second-level subsumptions.[40] He wished to help his readers move beyond the surface grammar appearance that these are first-level subsumptions. In this, I diverge from readers of Frege who take his insistence that the use of the singular definite article results in a proper name to be categorical. Those readers then proceed to be struck by the peculiarity of that insistence, given how Frege did not seem to have an issue with abandoning surface grammar appearance in other cases.[41] This tension can be resolved by reading Frege's insistence as being – not categorical – but hypothetical in the following way: it is only once one has acquired a proper grasp of the distinction between first-level and second-level subsumptions that one is in the position to understand a seeming first-level subsumption as a second-level subsumption. Given such a grasp, there is nothing that stops one from using such sentences as 'The concept *horse* is realized' as second-order subsumptions, and to insist that they should be rendered as such in the *Begriffsschrift*.[42] Frege would be the last to say that it is somehow impossible to use 'The concept *horse* is realized' in this way.[43]

As I read Frege, there is no rule written in stone that the singular definite article always results in an expression that *must* be understood as a proper name wherever it appears in a statement. Frege's true concern is this: the way in which the singular definite article is used in ordinary language is tied to the old subject/predicate logic, and thereby to the

intensional. Without being able to discuss this in full detail here, I expect that the solution would lie in taking that notion as a second-level concept that subsumes *senses* of first-level concept names, senses that are themselves incomplete. Note that Frege needs a treatment of such higher-level intensional contexts in order to account for some of his own statements about concepts, such as the following passage from the *Grundlagen*: "Often it is only after immense intellectual effort, which may have continued over centuries, that humanity at last succeeds in achieving knowledge of a concept in its pure form, stripping off the irrelevant accretions which veil it from the eyes of the mind" (GL, VII). I am grateful to professor Michael Kremer for bringing this issue to my attention.

[40] Once seen in this way, it does become true, of course, that 'There is at least one square root of 4' and 'The concept *square root of* 4 is realized' express the same thought. But this cannot be what Frege had in mind in the passage discussed above, since he also wished to hold there that the former says something about a concept, whereas the latter says something about an object.

[41] This point was put to me by an anonymous referee, who thereby constitutes one such reader. Another good example is (Beaney, 1996, 190-191).

[42] Travis seems to make a similar suggestion (Travis, 2020, 245-246).

[43] Which is not to say that he would not still wish to say that there is something logically defective about such sentences. This issue will reappear in Chapter 4. Note that Frege calls the principle that the singular definite article always indicates an object a *Kennzeichen*, i.e. something like a *characteristic feature or mark*, which seems to allow for exceptions. Frege himself suggests the possibility of exceptions in his discussion of 'The Turk besieged Vienna' and 'The horse is a four-legged animal' (CP, 185).

confused amalgamation of logical subjects and objects that Frege wants to overcome. As long as one is wedded to this logically confused perspective, one's use of expressions such as 'the concept *horse* is realized' will remain equally confused, so that it cannot properly be construed as a second-level subsumption. For that, one first has to attain the logico-philosophical understanding that the logically deficient grammar of ordinary language makes it so difficult to attain. For Frege, the capacity to drive a wedge between the surface grammar of a statement and the judgment expressed by it requires such an understanding. That is why Frege wavers between saying things such as that the intended thought of statements such as 'the concept *horse* is realized' is 'missed' or 'falsified' on the one hand, and things such as that it is merely 'obscured' on the other.[44] The former emphasizes the confused logical perspective with which ordinary language saddles us, while the latter is said from the position of someone who has already attained the requisite logico-philosophical understanding. Kerry, of course, occupies the confused perspective: his understanding of his own use of the expression 'the concept *horse*' betrays exactly the sort of confusions Frege wishes us to overcome – and this precludes a construal of his statements as second-level subsumptions.

In a similar way, one could object that statements such as 'There exists at least one horse' *is* a second-level subsumption, so that the distinction between first-level and second-level subsumptions is already present in ordinary language, which is thereby not wedded to the old subject/predicate logic at all. Again, however, the capacity to recognize 'There exists at least one horse' as a second-level subsumption requires one to have already grasped the *Begriffsschrift* distinction between first-level and second-level subsumptions. Only then can one come to see 'There exists at least one horse' as a statement *about* a concept in such a way that the confused amalgamation of logical subjects and objects is broken.[45]

As we saw, Kerry is not alone in his amalgamation of logical subjects and objects. Jolley, Bengtsson, and Witherspoon – each in their own way – side with Kerry on this score. A crucial difference between Kerry and them, however, is that Kerry accepts, while the others reject, the following conditional claim: If Frege's concept/object distinction is such

[44] See, for instance, the "obscures—I might almost say falsifies" (PW, 119) from the passage about the concept *equilateral triangle* discussed above.

[45] To illustrate the importance of a background grasp of the *Begriffsschrift* to achieve clarity about such ordinary language statements, consider what goes into one's understanding of such sentences as Dummett's 'A philosopher is what 'ξ is a philosopher' stands for', which is to be construed as equivalent to 'There exists at least one philosopher' (Dummett, 1986, 217).

that it precludes the possibility of making judgments about concepts, it must be rejected. As we have seen, this is a point on which Kerry and Frege *agree*. In rejecting this conditional claim, then, such readers completely miss what is at stake in the dialectic between Frege and Kerry. In insisting that the very idea of making judgments about concepts is incoherent, readers such as Jolley take themselves to be unfolding and further defending Frege's insights against Kerry's confusions. Without being aware of it, however, it is in fact Kerry with whom they share a crucial premise: that a logical subject must always appear non-predicatively in a judgment. If one combines that premise with Frege's claim that concepts are essentially predicative, then one does end up with the claim that the very idea of making judgments about concepts is logically incoherent. Frege himself, however, is at pains to avoid such a view. Rather than unfolding Frege's own logico-philosophical ideas, then, such readers remain subject to the very logical confusions that Frege wished to overcome, confusions that issue in the amalgamation of the notions of logical subject, object, and aboutness that is characteristic of the old subject/predicate logic. Another symptom of this is that none of them pay much attention to Frege's device of higher-level subsumptions.[46] If my account is on the right track, however, the logical categories are logically intertwined in such a way that a proper understanding of any of Frege's logical category distinctions is inseparable from an understanding of the logical stratification of the realm of *Bedeutungen* as such, which in turn involves an understanding of the logical stratification of a whole cluster of logico-philosophical phenomena, including logical subjecthood.

[46] This is one way in which the customary focus on *the* concept/object distinction threatens to be misleading. It suggests that Frege's logic revolves around this one core distinction. It does not: it revolves around the logical stratification of the realm of *Bedeutungen*, of which the concept/object distinction is best described as constituting one inseparable moment. A second, related way in which it threatens to be misleading is that it can suggest that Frege replaces the subject/predicate nexus with the concept/object nexus. He does not. Again: he replaces it with the logically stratified function/argument nexus. Against this background, one can describe the concept/object nexus as what, in Frege's philosophical logic, most closely resembles the old subject/predicate nexus, but only if one first goes through the replacement of the subject/predicate nexus with the logically stratified function/argument nexus.

Chapter 3
Frege *vis-à-vis* Russell on function and object

With the insights into the logical stratification of the realm of *Bedeutungen* from Chapter 2 under our belt, we now turn to one of the finest examples of the elucidation of the logical categories from Frege's *oeuvre*, which occurs in his correspondence with Russell. Russell is Frege's most potent interlocutor by far, and their correspondence is a document of rare philosophical richness, of which the full measure remains to be taken in the literature. I will present a close reading of a selection of passages – which constitute a mere fraction of that correspondence[1] – that offer us unique insights into Frege's conception of the logical categories and their logico-philosophical elucidation. We will see how they give rise to an initial framework for a further investigation of that conception, an investigation that will occupy us in the final two chapters of this dissertation.

The passages in question begin with Russell voicing an objection to Frege's sharp distinction between functions and objects.[2] They continue with a reply by Frege to that objection, a subsequent reply by Russell to Frege's reply, and a final rebuttal by Frege. These passages are interwoven with other discussions in the correspondence. Still, they can be fruitfully read as constituting a more or less self-standing discussion, which is what I will do here. Although more could be said about how they fit into the broader dialectic between Frege and Russell, I do not think that this would require me to significantly alter

[1] For a discussion of some other passages from the correspondence concerning an incipient version of Russell's theory of types, see my (Vanrie, 2020).
[2] In this chapter, we are back to the function/object distinction rather than the concept/object distinction. It must be kept in mind that the notion of a function itself stratifies into distinct logical categories, a point that will become important in Frege's reply to Russell.

my account. As a preliminary, let me present the passages in full, so as to give an initial overview of the stretch of text that we will be discussing in detail throughout this chapter.

Russell's objection (first formulation)

> From what you say on p. 37, that a function name can never take the place of a proper name (I am speaking of the Basic Laws), there arises a philosophical difficulty. I know very well what good reasons there are to be found for this view; yet it is self-contradictory. For 'ξ can never take the place of a proper name' is a false proposition if ξ is a proper name, but otherwise it is not a proposition at all. If there can be something which is not an object, then this fact cannot be stated without contradiction; for in the statement, the something in question becomes an object. It therefore seems to me doubtful whether the φ in φx can be regarded as anything at all. But at this point we are plunging into philosophical logic (PMC, 134; BW, 216).

Frege's first reply

> Concerning your doubts regarding my proposition that a function name can never take the place of a proper name, we must distinguish sharply between a name or sign and its meaning. When I use a proper name in a proposition, I am not talking about this proper name but about the object it designates. But it can happen that I want to talk about the name itself; I then enclose it within quotation marks. In order to bring out the unsaturatedness of the function name, let me leave the argument place empty for once. I can then say:
>
> '( ) · 3 + 4' is a function name
>
> and
>
> '( ) · 3 + 4' can never take the place of a proper name.
>
> You are correct in saying:
>
> "ξ can never take the place of a proper name' is a false proposition if ξ is a proper name';
>
> but you are not correct in continuing:
>
> 'but otherwise it is not a proposition at all'.
>
> On the other hand, it is correct to say:
>
> If 'ξ' is not a proper name, then 'ξ can never take the place of a proper name' is not a proposition.

Here "( ) · 3 + 4" – with two sets of quotation marks – takes the place of 'ξ'. While '( ) · 3 + 4' is a function name, "( ) · 3 + 4" is a proper name, and its meaning is the function name '( ) · 3 + 4'. In the proposition 'Something is an object', the word 'something' takes an argument place of the first kind and stands for a proper name. Thus whatever we put in place of 'something', we always get a true proposition; for a function name cannot take the place of 'something'. Here we find ourselves in a situation where the nature of language forces us to make use of imprecise expressions. The proposition 'A is a function' is such an expression: it is always imprecise; for 'A' stands for a proper name. The concept of a function must be a second-level concept, whereas in language it always appears as a first-level concept. While I am writing this, I am well aware of having again expressed myself imprecisely. Sometimes this is just unavoidable. All that matters is that we know we are doing it, and how it happens. In a conceptual notation we can introduce a precise expression for what we mean when we call something a function (of the first level of one argument), e.g.: '$\acute{\varepsilon}\varphi(\varepsilon)$'. Accordingly, '$\acute{\varepsilon}(\varepsilon \cdot 3 + 4)$' would express precisely what is expressed imprecisely in the proposition '$\xi \cdot 3 + 4$ is a function'. Whatever we now put in place of '$\varphi( )$', we always get a true proposition because we can only put in names of functions of the first level of one argument, for the argument place here is of the second kind. Just as in language we cannot properly speaking say of a function that it is not an object, so we cannot use language to say of an object, e.g. Jupiter, that it is not a function. You are correct in thinking that a function cannot properly be treated as something; for, as I said before, the word 'something' stands for a proper name. Instead of using the imprecise expression 'ξ is a function', we can say: "( ) · 3 + 4' is a function name'. We cannot properly say of a concept name that it means something; but we can say that it is not meaningless. It is clear that function signs or concept names are indispensable. But if we admit this, we must also admit that there are some that are not meaningless, even though, strictly speaking, the expression 'the meaning of a function name' must not be used (PMC, 135-137; BW, 217-219).

Russell's objection (second formulation)

Concerning function names, there still seems to me to be a difficulty. If we leave aside names altogether and speak merely of what they mean, then we must admit that there is no proposition in which a function takes the place of a subject. But the proposition 'A function never takes the place of a subject', is self-contradictory; and

> it seems to me that this contradiction does not rest on a confusion of a name with what it means (PMC, 138; BW, 220).

Frege's second reply

> The difficulty in the proposition 'A function never takes the place of a subject' is only an apparent one, occasioned by the inexactness of the linguistic expression; for the words 'function' and 'concept' should properly speaking be rejected. Logically, they should be names of second-level functions; but they present themselves linguistically as names of first-level functions. I have, I believe, dealt with this in my essay 'On Concept and Object'. If we want to express ourselves precisely, our only option is to talk about words or signs. We can analyse the proposition '3 is a prime number' into '3' and 'is a prime number'. These parts are essentially different: the former complete in itself, the latter in need of completion. Likewise, we can analyse the proposition '4 is a square number' into '4' and 'is a square number'. Now it makes sense to fit together the complete part of the first proposition with that part of the second proposition which is in need of completion (that the proposition is false is a different matter); but it makes no sense to fit together the two complete parts; they will not hold together; and it makes just as little sense to put 'is a square number' in place of '3' in the first proposition. This difference between the signs corresponds to a difference in the realm of meanings; although it is not possible to speak of it without turning what is in need of completion into something complete and thus falsifying the real situation. We already do this when we speak of 'the meaning of 'is a square number''. Yet the words 'is a square number' are not meaningless. The analysis of the proposition corresponds to an analysis of the thought, and this in turn to something in the realm of meanings, and I should like to call this a primitive logical fact. This is precisely why no proper definition is possible here (PMC, 141-142; BW, 224).

There is a crucial difference between this dialectic between Frege and Russell, and the dialectic between Frege and Kerry discussed in the previous chapter. Kerry's worry about Frege's concept/object distinction was that it precludes us from making judgments about concepts, a result that Kerry believes should be avoided at all costs. Frege agrees that the result must be avoided, but aims to show that it does not issue from his concept/object distinction. As Frege sees it, Kerry's worry rests on a logically confused adherence to a version of the old subject/predicate logic, in which the notions of an object and a logical

subject are amalgamated. It is this old conception that must be displaced by the logical stratification of subjecthood that lies at the heart of the *Begriffsschrift*.

In replying to Kerry, Frege is elucidating such fundamental logico-philosophical phenomena as the logical stratification of subjecthood, the concept/object distinction, the device of second-level subsumptions, etc. Frege's reply to Kerry thereby offers us a clear instance of the activity of logico-philosophical elucidation, as Frege conceives it. Frege's dialectic with Russell, however, moves us deeper still into Frege's conception of the logical categories. This is because Russell does not merely raise the issue of making judgments about concepts (functions) – although, to be sure, that issue constitutes an important moment in his objection and will continue to figure prominently in my account – but also the issue of elucidation itself. Whereas the dialectic with Kerry is focused on statements such as 'The concept *horse* is realized', Russell questions the status of elucidations such as 'A function never takes the place of an object' themselves. Those statements, Russell claims, are self-contradictory, so that they cannot serve as suitable articulations of Frege's conception of the logical categories. We could say that, whereas Kerry is worried about the *consequences* of Frege's conception, Russell raises the issue of the *intelligibility* of that conception. In this way, Russell squarely confronts Frege with the predicament that was presented in the Introduction in a way that Kerry never did.

3.1 Russell's objection (first formulation)

I start with a discussion of Russell's first formulation of his objection. I divide it into three parts, which Frege will address in order:

> [Part I] From what you say on p. 37, that a function name can never take the place of a proper name (I am speaking of the *Basic Laws*), there arises a philosophical difficulty. I know very well what good reasons there are to be found for this view; yet it is self-contradictory. For 'ξ can never take the place of a proper name' is a false proposition if ξ is a proper name, but otherwise it is not a proposition at all. [Part II] If there can be something which is not an object, then this fact cannot be stated without contradiction; for in the statement, the something in question becomes an object. [Part III] It therefore seems to me doubtful whether the φ in φx

> can be regarded as anything at all. But at this point we are plunging into philosophical logic (PMC, 134; BW, 216).

We immediately see the above-mentioned difference with Kerry appear: rather than focusing on statements about specific concepts, Russell focuses on statements that purport to articulate the function/object distinction itself. He considers two versions of such a statement, both of which Frege will discuss, and which it is important to keep apart:

> (D1) A function name can never take the place of a proper name.
> (D2) There is something that is not an object.

The salient difference between (D1) and (D2) is that the former takes signs as its subject matter, whereas the latter concerns the realm of *Bedeutungen* itself. Russell's objection to both is that they are self-contradictory. This is clearest for (D2): In saying that 'something is not an object', whatever comes to stand for the 'something' is put in the position of the logical subject of predication, so that it must *ipso facto* be an object. The very act of saying that there is something that is not an object, then, presupposes that the something in question *is* an object, resulting in contradiction. Russell reasons similarly for (D1): in saying of a function-name that it can never take the place of a proper name, that function name *does* come to occupy the place of a proper name, resulting in contradiction.

We can read Russell's objection as raising the issue of the possibility of making judgments about functions with regards to Frege's elucidations themselves. Those elucidations seem to purport to make judgments about functions, the upshot of which is that functions are fundamentally different from objects. Russell, however, thinks that the only way to make judgments about functions is to have them appear as objects, so that Frege's elucidations are self-undermining. In one way, then, Russell's objection is similar to Kerry's: both rest on the amalgamation of logical subjects and objects characteristic of the old subject/predicate logic, and both involve a version of the worry that, given Frege's function/object distinction, it becomes impossible to make judgments about functions.

There remains the crucial difference, however, that Russell turns this worry against Frege's elucidations themselves. To disarm Russell's objection, it does not suffice to point out the possibility of making judgments about functions, since elucidations cannot be cast in the role of such judgments. In replying to Russell's objection, then, Frege will have to address the logico-philosophical status of his elucidations of the logical categories themselves, and this is why Russell's challenge cuts deeper than Kerry's.

3.2 Frege's first reply

3.2.1 Reply to Part I

Frege starts by replying to Part I of the first formulation of Russell's objection, the part that concerns (D1). With regards to (D1), Frege charges Russell with a confusion between sign and *Bedeutung:*[3]

> Concerning your doubts regarding my proposition that a function name can never take the place of a proper name, we must distinguish sharply between a name or sign and its meaning. When I use a proper name in a proposition, I am not talking about this proper name but about the object it designates. But it can happen that I want to talk about the name itself; I then enclose it within quotation marks. In order to bring out the unsaturatedness of the function name, let me leave the argument place empty for once. I can then say:
>
> '( ) · 3 + 4' is a function name
>
> and
>
> '( ) · 3 + 4' can never take the place of a proper name.
>
> You are correct in saying:
>
> [(A)] '"ξ can never take the place of a proper name" is a false proposition if ξ is a proper name';
>
> but you are not correct in continuing:
>
> 'but otherwise it is not a proposition at all'.
>
> On the other hand, it is correct to say:
>
> [(B)] If 'ξ' is not a proper name, then 'ξ can never take the place of a proper name' is not a proposition.
>
> Here '"( ) · 3 + 4"' – with two sets of quotation marks – takes the place of 'ξ'. While '( ) · 3 + 4' is a function name, '"( ) · 3 + 4"' is a proper name, and its meaning is the function name '( ) · 3 + 4' (PMC, 135-136; BW, 217-218).

Russell had claimed: "'ξ can never take the place of a proper name' is a false proposition if ξ is a proper name, but otherwise it is not a proposition at all" (PMC, 134). It is unclear, however, whether Russell means to use the second occurrence of 'ξ' to refer to the proper

[3] I have inserted some labels into the quotation for my discussion below.

name that takes up the position of the first occurrence of ξ, or the *Bedeutung* of that proper name.[4] Russell is right about the following: the expression that comes to occupy the place of 'ξ' in 'ξ can never take the place of a proper name' must be a proper name, if the result is to be a meaningful proposition. If we try to put a function name there instead – by writing, for instance, '$(\,) \cdot 3 + 4$ can never take the place of a proper name' – the result is nonsense.[5] That is what Frege expresses in the statement labelled (B) in the quotation above.

If we put a proper name in the place of 'ξ', the resulting proposition will not, however, be about that proper name, but about its *Bedeutung*. Let us call this proper name *A*. There are three salient possibilities: (1) *A* refers to an object that is not a name, as in 'Socrates can never take the place of a proper name'; (2) *A* refers to a proper name, as in ''Socrates' can never take the place of a proper name'; (3) *A* refers to a function name[6], as in ''$(\,) \cdot 3 + 4$' can never take the place of a proper name'. In case (1), the resulting proposition is true, but it is not relevant to Frege's and Russell's concerns. In case (2), the resulting proposition is false, and this is what Frege expresses in the statement labelled (A) in the quotation above. In case (3), however, the resulting proposition is true, according to Frege, and this is why Russell's clause 'but otherwise it is not a proposition at all' is incorrect.

Thus, Russell's claim that (D1) is self-contradictory rests on a failure to distinguish between a function name, on the one hand, and a proper name that is used to designate that function name, on the other hand. When Frege advances (D1) in the *Grundgesetze* (GGA, I, §21), he cannot be charged with making a self-contradictory statement, at least not on the grounds Russell advances. This is not to say that there are no further questions to be asked about the logico-philosophical status of statements such as (D1). Indeed, as will soon become clear, statements such as (D1) play a crucial role in the elucidation of the logical categories, as Frege conceives it, and will occupy us in great detail in the

[4] This unclarity is no doubt related to Russell's early view that the constituents of propositions *are* what the propositions are about.

[5] Unless, of course, we use the expression '$(\,) \cdot 3 + 4$' as a proper name. There are complex philosophical issues in the offing here, which will become more prominent in the next chapter.

[6] That Frege took (3) to be a possibility, is clear from the passage, where he explicitly says that ''$(\,) \cdot 3 + 4$'' (with double quotation marks) is a proper name that has a function-name as its *Bedeutung*. This shows that Frege did not take function names to be themselves functions, as some readers – most notably Geach – have claimed. Geach's reading will be discussed in detail in the final chapter.

following chapters. Russell's reasons for seeing problems with (D1), however, hold no water.

3.2.2 Reply to Part II

Next, Frege addresses Part II of Russell's objection, which concerns (D2) and does not rest on a confusion between sign and *Bedeutung*. Frege begins his discussion of (D2) as follows:

> In the proposition 'Something is an object', the word 'something' takes an argument place of the first kind[7] and stands for a proper name. Thus whatever we put in place of 'something', we always get a true proposition; for a function name cannot take the place of 'something' (PMC, 136; BW, 218).

Russell believes that Frege's function/object distinction commits him to (D1) and (D2), and he believes that (D1) and (D2) are self-contradictory. In the case of (D1), Frege agrees that the function/object distinction commits him to (D1), but he takes Russell's reasons for thinking that it is self-contradictory to be mistaken. With (D2), the situation is different. As the above quotation shows, Frege agrees that (D2) is self-contradictory. Thus, he will have to show that Russell's reasons for thinking that the function/argument distinction commits him to (D2) are mistaken.

Why does Russell think that Frege is committed to (D2)? Russell regards (D2) as what one must come out with as a statement of Frege's function/object distinction. Thus, we can advance two principles that are at work in Russell's argument:

> (FUNCTIONS) Frege must be able to say about functions that they are not objects.
> (ABOUTNESS) The only way to make judgments about functions, is to have them appear as objects.

Combined with the function/object distinction, (FUNCTIONS) and (ABOUTNESS) are incompatible. We already know, of course, Frege's solution to this problem, since (ABOUTNESS) is the very supposition that Frege takes to be one of the confusions inherent in the old subject/predicate logic, and that is overcome through the *Begriffsschrift* device of higher-level subsumptions. It should be no surprise, then, that

[7] See (GGA, I, §23) for Frege's taxonomy of different argument places. An argument place of the first kind is one that is to be occupied by a proper name.

Frege proceeds to point out to Russell the *Begriffsschrift* distinction between the first and the second level:

> Here we find ourselves in a situation where the nature of language forces us to make use of imprecise expressions. The proposition '*A* is a function' is such an expression: it is always imprecise; for '*A*' stands for a proper name. The concept of a function must be a second-level concept, whereas in language it always appears as a first-level concept. While I am writing this, I am well aware of having again expressed myself imprecisely. Sometimes this is just unavoidable. All that matters is that we know we are doing it, and how it happens (PMC, 136; BW, 218).

Russell takes Frege to be committed to (D2) as a judgment about functions. Since he takes such a judgment to require having functions appear as objects, he regards it as self-undermining. The *Begriffsschrift* device of second-level subsumptions, however, reveals how to make judgments about functions without having them appear as objects, so that statements such as (D2) need not be *ipso facto* self-undermining for the reasons Russell thinks. Indeed, Frege proceeds to make this more explicit by involving the *Begriffsschrift* notation.

> In a conceptual notation we can introduce a precise expression for what we mean [*was man meint*] when we call something a function (of the first level of one argument), e.g.: '$\acute{\varepsilon}\varphi(\varepsilon)$'[8]. Accordingly, '$\acute{\varepsilon}(\varepsilon \cdot 3 + 4)$' would express precisely what is expressed imprecisely in the proposition '$\xi \cdot 3 + 4$ is a function'. Whatever we now put in place of '$\varphi(\,)$', we always get a true proposition because we can only put in names of functions of the first level of one argument, for the argument place here is of the second kind (PMC, 136; BW, 218).

Frege's introduction of the second-level concept '$\acute{\varepsilon}\varphi(\varepsilon)$' makes explicit how the *Begriffsschrift* notation allows one to make judgments about functions without having them appear as objects. In juxtaposing 'ξ is a function' with '$\acute{\varepsilon}\varphi(\varepsilon)$', Frege seeks to make clear to Russell that the fact that a statement such as 'ξ is a function' only takes objects as argument – and is thereby self-contradictory – is an accidental fact of ordinary

[8] '$\acute{\varepsilon}\varphi(\varepsilon)$' is *not* the expression for courses-of-values, which uses the reverse *spiritus*. It is a new *Begriffsschrift* sign that, as far as I know, appears nowhere else in Frege's *oeuvre*. The notation resurfaces at (PMC, 161), but is there used differently. As used here, '$\acute{\varepsilon}\varphi(\varepsilon)$' designates a second-level concept that yields the True whichever first-level function of one argument is put into its argument-place, so that we have $(\forall F)\big(\acute{\varepsilon}F(\varepsilon)\big)$.

language that is not essential to logic, and that is overcome in the *Begriffsschrift* notation, so that we can leave behind the confused principle (ABOUTNESS) of the old subject/predicate logic.[9]

Strictly speaking, this suffices to deal with Russell's objection, since Russell has presented his argument in such a way that it relies on (ABOUTNESS). Without (ABOUTNESS), the fact that (D2) is self-contradictory does not directly undermine the function/object distinction. If this were the end of the matter, Russell's objection would essentially be the same as Kerry's, and Frege's dialectic with Russell would be comparable to his dialectic with Kerry. There is, however, a further layer of philosophical complexity to Russell's argument which distinguishes it from Kerry's, and which remains to be addressed.

So far, I have said that Frege is happy to accept (D2) as self-contradictory, since this only undermines the function/object distinction if (ABOUTNESS) is also accepted. Russell, however, could rightly be dissatisfied with such a response. After all, Frege himself advances statements exactly like (D2) throughout his *oeuvre*, and seems to acknowledge that they give voice to commitments engendered by his function/object distinction. That is what seems to give us the other principle (FUNCTIONS) above. If Frege really is so happy to grant that (D2) is self-contradictory, then he should be equally happy to retract such statements, and this would amount to retracting the function/object distinction itself.

If we stick to a general formulation, it is correct that Frege is committed to statements such as (D2). The crucial question, however, is what sort of commitment to what sort of statement is at issue here. Russell takes Frege's commitment to (D2) to be a commitment to the possibility of making judgments about the logical categories.[10] For Russell, this is obvious, since it seems obvious to him that the only way to articulate a conception of the logical categories is through expressions of judgment. In other words, Russell has no room for a distinction between judgment and elucidation. Frege, however, rejects such a construal of his commitment to statements such as (D2). According to Frege, the sort of

[9] Compare: "In §49 of his book, *The Principles of Mathematics*, Mr. B. Russell does not want to concede that a concept is essentially different from an object; concepts, too, are always supposed to be *terms*. He supports his argument here with the contention that we find it necessary to use a *concept* substantively as a *term* if we want to say anything about it, e.g. that it is not a *term*. In my opinion, this necessity is grounded solely in the nature of our language and therefore is not a properly logical one" (CP, 282n).

[10] Recall: "If there can be something which is not an object, then this *fact* cannot be *stated* without contradiction" [my emphases] (PMC, 134). Russell's use of the terms 'fact' and 'stated' is the only use he has in view: a use that is wedded to judgment.

grasp of the logical categories that we are after in philosophical logic is not such as to be articulated through the activity of judgment, but rather through the activity of logico-philosophical elucidation. Indeed, a proper understanding of the logical categories reveals precisely *that* the idea of making judgments about the logical categories is inherently confused, for the reasons already discussed in Chapter 1, and which will be further fleshed out below.

From Frege's point of view, then, (FUNCTIONS) is crucially ambiguous between two construals:

> (FUNCTIONS-JUDGMENT) Frege must be able to express in judgment the fact that functions are not objects.
> (FUNCTIONS-ELUCIDATIONS) Frege must be able to elucidate the fundamental difference between functions and object.

While Frege accepts the latter, he rejects the former. In this way, (D2) comes to be seen as ambiguous between a construal according to which it is supposed to vindicate (FUNCTIONS-JUDGMENT) or (FUNCTIONS-ELUCIDATIONS), between a construal of it as a judgment or as an elucidation. If (D2) is construed as a judgment, the notion of being an object that figures in it must be construed as a first-level concept under which all objects fall, a first-level concept such as '$\xi = \xi$'. On such an understanding, (D2) can be expressed in the *Begriffsschrift* notation as $(\exists x)(x \neq x)$, which is indeed self-contradictory. If (D2) is construed as an elucidation, on the other hand, the notion of being an object that figures in it pertains to the logical category of objects, and this is not a first-level concept at all, nor any other kind of concept (or function).

When I say that Frege is happy to accept that (D2) is self-contradictory, what I mean is that he is happy to accept that the first construal of (D2) is self-contradictory. What about the second construal? Here, the notion of self-contradiction does not apply. Since elucidations are not judgments, they cannot contradict each other, or themselves, in the usual sense.[11] What is at issue, is whether (D2), construed as an elucidation, properly contributes to the logico-philosophical elucidation of the logical categories. To criticize

[11] I do not wish to claim that terms such as 'contradiction', 'truth', and 'falsity' can play no role in the assessment of elucidatory statements. It is helpful, of course, to say that 'Concepts are objects' and 'No concept is an object' contradict each other, and that the former is false while the latter is true. Indeed, I use such locutions myself. It is important to keep in mind, however, that the operative notions of contradiction, truth, and falsity are not the same as those that apply to *Begriffsschrift* judgments.

Frege's elucidatory use of (D2), one would have to show that it does not, by showing, for instance, that it issues in confusion in the way that Kerry's claim that 'The concept *horse* is a concept' does. For Russell, however, there are no two ways of construing (D2). The only way in which Russell can understand Frege's notion of elucidation is as marking off a specific kind of judgment, i.e. those judgments that have a purely logical subject matter. He cannot understand Frege's notion of elucidation as marking off a logico-philosophical activity that is not that of judging at all, because Russell's conception of philosophical logic does not allow for the idea that any other activity than judging could be essential to it.

To bring further clarity to these matters, let us take another look at Frege's introduction of the second-level concept $\acute{\varepsilon}\varphi(\varepsilon)$, and his juxtaposition of that concept with 'ξ is an object'. On a first approach, it may seem that Frege himself is trying to vindicate, through his introduction of $\acute{\varepsilon}\varphi(\varepsilon)$, the possibility of making judgments about the logical categories after all. In other words: it may seem that Frege is trying to answer Russell's objection by showing how his own elucidatory statements – such as (D2) – may be construed as judgments in which the second-level concept $\acute{\varepsilon}\varphi(\varepsilon)$ figures.

Such a reading, however tempting it may be, would be mistaken, and is in fact foreclosed by Frege himself in the sentence immediately following his introduction of $\acute{\varepsilon}\varphi(\varepsilon)$:

> Just as in language we cannot properly speaking say of a function that it is not an object, so we cannot use language [*dieser Bezeichnung*] to say of an object, e.g. Jupiter, that it is not a function (PMC, 136).

There is an important error in the translation. With '*dieser Bezeichnung*', Frege is not referring to language as such, but to the specific *Begriffsschrift* expression '$\acute{\varepsilon}\varphi(\varepsilon)$' which he has just introduced. Thus, Frege is quite explicit that '$\acute{\varepsilon}\varphi(\varepsilon)$' cannot be used to capture his own contrastive elucidatory statements, so that we should not take his introduction of that *Begriffsschrift* expression as an attempt to render those elucidations as *Begriffsschrift* judgments.

How, then, are we to understand Frege's juxtaposition of 'ξ is a function' with '$\acute{\varepsilon}\varphi(\varepsilon)$'? That juxtaposition, as I understand it, is to be situated on the level of the first construal of (D2) discussed above, i.e. the construal of 'There is something that is not an object' as a first-level subsumption in which the notion of an object takes the guise of a first-level concept. On such a construal, we saw, the statement 'There is something that is not an

object' is self-contradictory. But, insofar as we take this construal of (D2) as an attempt to express a true judgment about functions, the conclusion is not that the endeavour to make such judgments about functions is hopeless, but rather that this requires the *Begriffsschrift* device of second-level subsumptions. In this way, we arrive at the second-level concept '$\acute{\varepsilon}\varphi(\varepsilon)$', which is stipulated to be such that all first-level functions of one argument fall under it. In other words, the following holds: $(\forall F)\big(\acute{\varepsilon}F(\varepsilon)\big)$. In the same way that '$\xi$ is an object' can be construed as a first-level concept that subsumes all objects, '$\acute{\varepsilon}\varphi(\varepsilon)$' is a second-level concept that subsumes all first-level functions of one argument.

It is true that, whatever comes to stand for the 'something' in 'Something is not an object' must *ipso facto* be an object. But, as Frege sees it, this is not a general fact about any statement whatsoever. If we juxtapose 'Something is not an object' with '$(\exists F)\big(\neg\, \acute{\varepsilon}F(\varepsilon)\big)$', we see that the latter has an analogous property: whatever comes to stand for the '$F(\varepsilon)$', must *ipso facto* be a first-level function of one argument, and the resulting statement is also self-contradictory. But no one in their right mind would suggest that this shows that there are no objects, or that the very idea of making judgments about object is incoherent, or some such. By parity of reasoning, neither does the fact that 'Something is not an object' is self-contradictory show any such thing. Russell is wrong to think that the fact that (D2) is self-contradictory reveals some inherent logical obstacle to expressing truths about Fregean functions. That is what Frege seeks to make clear by introducing the second-level concept $\acute{\varepsilon}\varphi(\varepsilon)$.

Since Russell's objection is based on the construal of (D2) as a judgment, it is natural for Frege to take over that construal in his reply. When Frege says that '$\acute{\varepsilon}(\varepsilon \cdot 3 + 4)$' serves to express precisely what is expressed imprecisely in the proposition '$\xi \cdot 3 + 4$ is a function', he is thinking about such a construal. Still, because of the specific statement at issue, there is a real danger that the issue of how to make judgments about functions – such as $(\exists F)\big(\acute{\varepsilon}F(\varepsilon)\big)$ – will be conflated with the issue of how to understand Frege's logico-philosophical elucidations, of which (D2) may *also* be taken as an example. Since Russell sees no difference between elucidation and judgment, he sees no such danger, and the move of construing (D2) as a judgment appears entirely innocuous to him – insofar as it appears *as* a move at all. For Frege, however, this move makes all the difference in the world. That is why, after revealing the confusions underlying Russell's argument, he proceeds to note that nothing he has said about '$\acute{\varepsilon}\varphi(\varepsilon)$' should be taken to violate the distinction between judgment and elucidation, by observing that '$\acute{\varepsilon}\varphi(\varepsilon)$' cannot serve to

capture his contrastive elucidatory statements about the logical categories. Frege is pointing out that it would be a serious confusion to take '$\acute{\varepsilon}\varphi(\varepsilon)$' as a primitive *Begriffsschrift* term standing for the logical category of first-level functions of one argument. If we construe (D2) as an elucidation, the introduction of '$\acute{\varepsilon}\varphi(\varepsilon)$' does not bring us any closer to rendering that elucidation as a judgment.

Still, it may be tempting to retreat to the following conclusion: although we cannot use '$\acute{\varepsilon}\varphi(\varepsilon)$' to make *contrastive* statements about the logical categories, at least we can use it to express what may be called *true categorical subsumptions* of first-level functions of one argument under the logical category of first-level functions of one argument. This sort of view is advanced by Noonan, who says, on the one hand, that "we can say of an object that it *is* an object using a proposition of the form 'x is an object'", while also holding that "we cannot [...] construct any legitimate way of saying that a particular object is *not* a first-level function" (Noonan, 2006, 165).[12]

I believe such a view to be inherently confused. It rests on seeing the subsumption of an entity under a logical category as possessing a peculiar ambiguity: insofar as we are subsuming an entity under its own logical category, the relevant notion of subsumption is the regular notion of subsumption that figures in *Begriffsschrift* judgments, so that such a subsumption can be expressed in such judgments. *This* notion of subsumption is such that the very idea of subsuming (positively or negatively) an object under a second-level concept makes no sense. It follows that when we say, for instance, that an object is not a first-level concept, the notion of subsumption that is at issue in such a statement is entirely different. It follows that the 'is' that figures in the statement 'Grimes *is* an object', is not the 'is' that figures in the statement 'ξ *is the mother of X Æ A-XII is not* an object'. On this view, then, whatever an object's *being* an object consists in, this is logically unrelated to the notion of being that is at issue in a function's *not being* an object. The result is that we have a dismantling of the logico-philosophical unity of Frege's logically stratified elucidatory use of the logical category terms, leading once again to the sort of disambiguation that was considered and rejected in the previous chapter, this time under the guise of there being something logically special about a specific kind of application of

[12] Diamond seems to flirt with a similar view: "If the concept script allows the predication of category terms at all, we shall then be able to say only of the things that actually are in the category in question that they are. Whenever we say that something is *not* in a category we shall be wrong, because only of what is in a particular category shall we be able to say that it is not in that category, and category predicates will not have a genuine classifying use" (Diamond, 1984, 359).

logical category terms, namely those applications to what belongs to the logical category in question. When what is at issue is the elucidation of the logical categories, the elucidatory notion of being at issue in the statements that objects *are* objects and that functions *are not* objects, must be one and the same.

Whatever is at issue in *Begriffsschrift* judgments such as '$(\forall x)(x$ is an object)' and '$(\forall F)\big(\acute{\varepsilon}F(\varepsilon)\big)$', it is entirely foreign to what is at issue in Frege's elucidations of the logical categories. Insofar as we wish to hold on to the claim that '$\acute{\varepsilon}(\varepsilon \cdot 3 + 4)$' ascribes a logical category, we are using the notion of a logical category in a way that severs it from Frege's elucidatory notion of a logical category that is at issue in the logical stratification of the realm of *Bedeutungen*. We *can*, if we want, use the term 'logical category' to characterize the *Bedeutung* of '$\acute{\varepsilon}\varphi(\varepsilon)$', which is a second-level concept that subsumes all first-level functions of one argument. And we can, if we want, call the resulting judgments 'true categorical subsumptions'. But we have thereby not rendered any less fundamental the difference between what is such as to be the *Bedeutung* of a *Begriffsschrift* term, on the one hand, and the logical categories as they are at issue in Frege's logico-philosophical elucidations, on the other hand.

We can bring this out in a more technical manner as well. Taking our cue from the *Begriffsschrift* formula '$\acute{\varepsilon}\varphi(\varepsilon)$', we could introduce another second-level concept '$\acute{\pi}\varphi(\pi)$' that is *false* whatever first-level function is put into its argument place, and claim that '$\acute{\pi}(\pi \cdot 3 + 4)$' expresses precisely what is expressed imprecisely when we say (falsely) '$\xi \cdot 3 + 4$ is an object', so that '$\neg\acute{\pi}(\pi \cdot 3 + 4)$' expresses the true statement that this function is not an object. If this procedure were legitimate, it would turn out that we can express Frege's contrastive claims after all. The problem with this proposal, however, is that we want '$\acute{\pi}(\pi \cdot 3 + 4)$' to deny exactly what we affirm of an object like the number 4 when we say that 4 is an object. Otherwise, we are not really saying of a function that it is not an *object*, i.e. that it is not what 4 *is*. But '$\acute{\pi}(\pi \cdot 3 + 4)$' cannot fulfil this role, because 'ξ is an object' and '$\acute{\pi}\varphi(\pi)$' are of different levels, and therefore cannot share their *Bedeutung*.[13] Similarly, 'ξ is not a function' and '$\acute{\varepsilon}\varphi(\varepsilon)$' are of different levels, so that

[13] One may be tempted to refer to Frege's remark, discussed in the previous chapter, that the statements 'There is at least one square root of 4' and 'The concept *square root of* 4 is realized' express the same thought (CP, 188), the idea being that this shows that the first-level concept ξ *is realized* and the second-level concept $(\exists x)(Fx)$ have the same *Sinn*, so that there is no obstacle to taking 'ξ is an object' and '$\acute{\pi}\varphi(\pi)$' to have the same *Bedeutung*.

whatever the latter affirms cannot be the same as what the former denies. We can put the point as follows: As long as we remain on the level of judgment, there is no way to bring the right sort of logico-philosophical unity to statements such as 'Grimes is an object' and '*ξ is the mother of X Æ A-XII* is not an object', since the level of judgment necessitates exactly the sort of disambiguation of the term 'object' that was argued against in the previous chapter. To bring the right sort of unity to those statements – the sort of unity that is at issue in Frege's conception of the logical categories – requires conceiving of such use of the term 'object' as a logically stratified elucidatory use.

What is true, is that we can introduce *Begriffsschrift* terms such as '$\acute{\varepsilon}\varphi(\varepsilon)$' and '$\acute{\pi}\varphi(\pi)$' to satisfy certain extensional requirements, such as being true (or false) of all first-level functions of one argument. This is the connection that exists between such expressions and the logical category of first-level functions of one argument. It is this connection that Frege wishes to point out to Russell, because Russell does not even believe that *such* a thing is possible, due to his being wedded to the amalgamation of logical subjects – and thus the notion of *being true of* – and objects that is characteristic of the old subject/predicate logic. What is confused, however, is to take this connection to reveal that the *Bedeutung* of such expressions give us to the logical categories themselves. The reason why '$\acute{\varepsilon}\varphi(\varepsilon)$' is true of all first-level functions of one argument is not that it *says* that something is such a function. Rather, the reason is that '$\acute{\varepsilon}\varphi(\varepsilon)$' takes first-level functions of one argument as arguments, and what it says can be anything that is true of all such functions, such as '$\varphi(0) = \varphi(0)$' or '$(\forall x)\big(\varphi(x) \vee \neg\varphi(x)\big)$'. It is the logical stratification of the realm of *Bedeutungen* that makes it possible for a *Begriffsschrift* term such as '$\acute{\varepsilon}\varphi(\varepsilon)$' to *have* a *Bedeutung* such that it subsumes all first-level functions of one argument in the first place. What does the work, is not the *Bedeutung* of '$\acute{\varepsilon}\varphi(\varepsilon)$', but the logically stratified realm of *Bedeutungen* wherein, we could say, that *Bedeutung* has its life.

This discussion has served to deepen a point that was already made in the first chapter: If we look at the realm of *Bedeutungen* and try to find a *Bedeutung* that fits Frege's contrastive elucidatory use of the logical category terms, we will see that no suitable *Bedeutung* is to be found. In this way, a logical category is not, for Frege, a *highest genus*, a

---

In that case, however, one would forget that Frege explicitly adds that what is said about the concept-subject in 'There is at least one square root of 4', and what is said about the object-subject in 'The concept *square root of* 4 is realized', is not the same (CP, 188), so that the first-level concept *ξ is realized* and the second-level concept $(\exists x)(Fx)$ cannot be taken to have the same *Bedeutung*.

concept that applies with maximal generality.[14] In this way, as well, the logical categories differ fundamentally from the *Bedeutungen* of primitive scientific terms, which *can* be designated by a *Begriffsschrift* term, but which are primitive because they stand at the beginning of the chain of definitions of a scientific system. Logical categories, on the other hand, are primitive because no *Begriffsschrift* terms can have them as their *Bedeutung* to begin with, since any term that would seem to do the trick would refer to a concept of a certain level, and the logical categories are not concepts of certain levels.

3.2.3 Reply to Part III

Next up is Frege's reply to Part III, i.e. to the single sentence: "It therefore seems to me doubtful whether the φ in φx can be regarded as anything at all" (PMC, 134; BW, 216). Frege writes:

> You are correct in thinking that a function cannot properly be treated as something; for, as I said before, the word 'something' stands for a proper name (PMC, 136-137; BW, 218-219).

This gives us another statement which is to be scrutinized:

> (D3) Functions cannot properly be treated as something.

There is indeed a sense in which Frege's function/object distinction commits him to treating functions as something, and there is also a sense in which it is correct to say that functions cannot be treated as something. The task is to see how these senses are compatible. Given what has been said so far, it will come as no surprise that (D3) must be disambiguated as follows:

> (D3O) Functions cannot properly be treated as objects.
> (D3S) Functions cannot properly be treated as logical subjects.

Because of his adherence to a version of the old subject/predicate logic, Russell sees no difference here. When Frege says that Russell is correct in thinking that a function cannot properly be treated as something, he is talking about (D3O), not (D3S). The purpose of Frege's reply is to show how he can accept (D3O) while rejecting (D3S).

[14] Compare (Diamond, 1984, 358) (Conant, 2020, 839).

To see this, however, a further ambiguity in (D3O) must be addressed. Before, we saw how (D2) could be construed either as a judgment – a generalized first-level subsumption – or as an elucidation. On the first construal, (D2) was self-undermining. The same is true of (D3O). Construed as a judgment, it is a generalized first-level subsumption in which it is said of 'functions' that they cannot properly be treated as objects. Since the notion of a function at issue here is a first-level concept, the 'functions' in question *are* objects. This forces the conclusion that there are no such 'functions'. So far, this is old news. What seems new, however, is that Russell – instead of pointing to the fact that (D3O) threatens to be self-undermining in a way similar to (D2) – in fact seems to endorse (D3O) as a genuine judgment about 'functions'. In endorsing the claim that the φ in φx cannot be regarded as anything at all, Russell seems to be guilty of the very sort of self-contradiction with which he aimed to saddle Frege: He is saying of something that it cannot be regarded as anything.

This is no accident. Already in the *Principles of Mathematics*, Russell struggled to fully carry through his principle that every constituent of every proposition must be capable of being made a logical subject. He admitted that he seemed to be forced to acknowledge that it does not, after all, apply universally:

> There appears to be an ultimate notion of assertion, given by the verb, which is lost as soon as we substitute a verbal noun, and is lost when the proposition in question is made the subject of some other proposition. [...] Thus the contradiction which was to have been avoided, of an entity which cannot be made a logical subject, appears to have here become inevitable. This difficulty, which seems to be inherent in the very nature of truth and falsehood, is one with which I do not know how to deal satisfactorily (PoM, §52).

In a nutshell, Russell's idea is this. In an assertion such as 'Kim loves Kanye'[15], we have *love* used as a verb, i.e. used to make an assertion. We can, however, convert this assertion into the verbal noun 'Kim's love for Kanye', in which *love* is not used as a verb, so that the assertion is lost: such a verbal noun does not assert anything. Thus, it seems that the relevant notion of *assertion*, which is tied to the notion of the verb, cannot be made into a logical subject, undermining Russell's principle. This is, I believe, the background against which to understand Russell's claim that the φ in φx cannot be regarded as anything at

[15] Whether this assertion is true is – it has unfortunately become clear – a different matter.

all. It cannot be regarded as anything because no *thing* comes with the assertion that is internal to the occurrence of 'φ' as a verb in 'φx'. Thus, there is a deep tension in Russell's own position. On the one hand, Russell has no room for anything that is not an object.[16] On the other hand, he admits that his own reflections on the nature of judgment commit him to recognizing something that is not an object: *assertion*. It is this tension that manifests itself in the fact that Russell regards (D2) as self-contradictory, on the one hand, while he seems to want to regard (D3O) as correct, on the other hand. From Russell's own point of view, however, such a position is incoherent, which is what underlies his puzzlement in *Principles of Mathematics*.

From Frege's point of view, Russell's puzzlement arises from his failure to distinguish between a construal of (D3O) as a judgment and a construal as an elucidation. When Russell sees (D3O) as something with which he is forced to agree, it is *qua elucidation* that he is forced to agree with it, and this is perfectly compatible with (D3O) being self-contradictory *qua* judgment. Thus, when Frege expresses his agreement with Russell that a function cannot properly be treated as something, it is with (D3O) construed as an elucidation that Frege expresses agreement.

What Frege proceeds to do, is elucidate how Russell's own admission that he regards (D3O) as correct leads to an acknowledgment of the function/object distinction. This requires addressing a second ambiguity in Russell's formulation of (D3O), which is again an ambiguity between sign and *Bedeutung*. That is to say, while (D3O) is true (construed as an elucidation), the following is not true:

> (D3N) We cannot treat a function name as an object.

That is what Frege makes clear by saying:

> Instead of using the imprecise expression 'ξ is a function', we can say: ''$( ) \cdot 3 + 4$' is a function name' (PMC, 136).

Russell's own statement is ambiguous between (D3O) and (D3N). Insofar as Russell's expression 'the φ in φx' is taken to designate a function name, we can perfectly well treat the φ in φx as something, for instance by saying something like 'the φ in φx is a function name', in a way analogous to Frege's statement that '$( ) \cdot 3 + 4$' is a function name.[17] It is

[16] In his terminology: anything that is not a term.
[17] As noted before, some will balk at this. We will come back to this issue in the final chapter.

only if we take 'the φ in φx' to designate a function, that it becomes correct to say that the φ in φx cannot properly be treated as something, if this is construed as an elucidation.

Thus, although we can properly treat function names as something, it remains true that we cannot treat their *Bedeutung* as something, i.e as an object. Rather than undermining the function/object distinction, however, this serves to reinforce it:

> We cannot properly say of a concept name that it means [*Bedeutet*] something; but we can say that it is not meaningless [*Bedeutungslos*]. It is clear that function signs or concept names are indispensable. But if we admit this, we must also admit that there are some that are not meaningless, even though, strictly speaking, the expression 'the meaning of a function name' must not be used (PMC, 136-137).

Russell, it seems safe to assume, agrees that function names – the sort of name that the 'φ' in 'φx' is – are indispensable. But this, Frege points out, means that such names must have *Bedeutung*.[18] As Russell sees it, however, this last step issues in self-contradiction. This is because Russell has two seemingly incompatible commitments:

> (1) The statement '*A* has *Bedeutung*' entails 'There is something which is the *Bedeutung* of *A*'.
> (2) The *Bedeutung* of a function name cannot properly be treated as something.

Russell's adherence to (1) is once again due to his adherence to a version of the old subject/predicate logic, which leads him to amalgamate – not only logical subjects and objects – but also objects and *Bedeutungen*. His adherence to (2) is an expression of his commitment to (D3). The way to resolve the contradiction, is to locate an ambiguity in the term 'something', so that we can resolve the seemingly contradictory pair (1) and (2) into:

> (1') The statement 'A has *Bedeutung*' entails 'The *Bedeutung* of A can be made into a logical subject'.
> (2') The *Bedeutung* of a function name cannot properly be treated as an object.

[18] How exactly does the fact that function names have *Bedeutung* follow from their indispensability? The crucial point, I think, is that the truth of a judgment turns on the identity of the function names that occur in its expression in a way that is comparable to how the truth of a judgment turns on the identity of the proper names (if any) that occur in its expression. Indeed, it is functions that map objects to truth-values. If function names had no *Bedeutung*, Frege's whole picture of how functionality is interrelated with truth would have to be relinquished. In short, we could say that function names have *Bedeutung* because they bring truth into play. I am grateful to professor Martin Gustafsson for pressing me on the need for further clarification here.

It is this disambiguation that allows Frege to say that, although a function name cannot be said to mean something, it is nevertheless not meaningless. And this yields the function/object distinction as the only way to remain faithful to the fact that function names are indeed indispensable. We must overcome the confusions of the old subject/predicate logic through a proper understanding of the logically stratified notion of *Bedeutung* of the *Begriffsschrift*. Russell himself, Frege suggests, cannot avoid this, given his own admission that he is committed to some version of (D3).

3.3 Russell's objection (second formulation)

Russell's response to Frege's first reply is comparatively brief:

> Concerning function names, there still seems to me to be a difficulty. If we leave aside names altogether and speak merely of what they mean, then we must admit that there is no proposition in which a function takes the place of a subject. But the proposition 'A function never takes the place of a subject', is self-contradictory; and it seems to me that this contradiction does not rest on a confusion of a name with what it means (PMC, 138; BW, 220).

What Russell has remembered from Frege's reply, is that Frege charges him with confusing sign and *Bedeutung*. This brings Russell to focus on yet another statement that can be read as purportedly stating Frege's function/object distinction:

> (D4) A function never takes the place of a subject.

As with (D1) and (D2), Russell believes that (D4) is self-contradictory: in saying of a function that it never takes the place of a subject, we *are* having that function take the place of a subject. Because Russell also believes that Frege's function/object distinction commits him to (D4), he once again concludes that the distinction must be rejected.

3.4 Frege's second reply

It must be noted that, in answering Frege in this way, Russell does not give the impression of having fully taken in Frege's first reply. Indeed, it seems as if Russell has really only read the first part. There, Frege does charge Russell with confusing sign and *Bedeutung*. But in his subsequent discussion of (D2), he does not, and although Frege's discussion of (D3) also involves clearing up an ambiguity between sign and *Bedeutung*, it also provides additional reflections which Russell does not address. In fact, being quite similar to (D3), (D4) does not seem to add much to the dialectic. For starters, it must be disambiguated in a similar way:

> (D4O) A function never takes the place of an object.
> (D4S) A function never takes the place of a logical subject.

Frege's immediate task, as before, is to elucidate how he can accept (D4O) while rejecting (D4S). As before, this involves a recognition of the distinction between first-level subsumptions and second-level subsumptions, which Frege proceeds to rehearse in his reply:

> The difficulty in the proposition 'A function never takes the place of a subject' is only an apparent one, occasioned by the inexactness of the linguistic expression; for the words 'function' and 'concept' should properly speaking be rejected. Logically, they should be names of second-level functions; but they present themselves linguistically as names of first-level functions. I have, I believe, dealt with this in my essay 'On Concept and Object' (PMC, 141; BW, 224).

As before, we should distinguish between a construal of (D4O) as a judgment – which is what Russell intends – and a construal of (D4O) as an elucidation. Russell's idea is that (D4O) – construed as a judgment – is self-contradictory. The reasons exactly mirror those why (D3O) is self-contradictory. Frege would agree that (D4O) is self-contradictory on this construal. As he did with (D2), however, he points out that this does not undermine the function/object distinction. Again, Russell's idea that it does undermine that distinction arises from his adherence to a version of the old subject/predicate logic, which makes him oblivious to the way in which second-level subsumptions allow us to make judgments about functions without falling into the sort of self-contradictory structure of (D4O). Thus, the fact that (D4O) – construed as a judgment – is self-contradictory, does not tell

us anything about the possibility of making judgments about functions. When Frege refers to *Über Begriff und Gegenstand*, he is referring exactly to his treatment of the problem of making judgments about concepts that was discussed in the previous chapter.

Still, as with (D2), it remains true that the device of second-level subsumptions does not allow us to capture the sort of contrastive elucidatory statements that are essential to Frege's logico-philosophical elucidations, and it also remains true that (D4O) – construed as an elucidation – is one of those contrastive elucidatory statements to which Frege is committed. Frege's reply to Russell brings us to see that Frege's elucidations cannot be construed as judgments. We could say that part of what Frege is trying to do in his reply to Russell is to elucidate the distinction between judgment and elucidation itself, making it clear that the function/object distinction does not constitute a subject matter for judgment, and that Russell's worries rest on a neglect of this point.

Having elucidated the distinction between elucidation and judgment, Frege proceeds to discuss his conception of the logical categories and their elucidation itself. He is prompted to do so by Russell's objection, because – unlike Kerry – Russell takes aim at Frege's elucidations themselves. In this way, Russell is Frege's interlocutor who advances the most penetrating challenge to Frege's conception of the logical categories, forcing Frege to face the predicament that was discussed in the Introduction. What follows is a rare moment in Frege's *oeuvre* where he – even if it is just with one sentence – explicitly says something about his conception of the activity of elucidating the logical categories as such:

> If we want to express ourselves precisely, our only option is to talk about words or signs (PMC, 141; BW, 220).

Whatever Frege means exactly by 'expressing ourselves precisely' – a question that will briefly resurface in the final chapter – Frege is pointing out that logico-philosophical elucidation, whatever its exact nature, must in some way or other involve turning our attention to signs, to language. He continues by matching word with deed, showing us straightaway what sort of discussion of signs he has in mind:

> We can analyse the proposition '3 is a prime number' into '3' and 'is a prime number'. These parts are essentially different: the former complete in itself, the latter in need of completion. Likewise, we can analyse the proposition '4 is a square number' into '4' and 'is a square number'. Now it makes sense to fit together the complete part of the first proposition with that part of the second proposition

> which is in need of completion (that the proposition is false is a different matter); but it makes no sense to fit together the two complete parts; they will not hold together; and it makes just as little sense to put 'is a square number' in place of '3' in the first proposition (PMC, 141-142; BW, 220).

Russell objected against Frege that the statement 'A function never takes the place of a subject' is self-undermining. In this way, he is challenging Frege's account of the logical stratification of the realm of *Bedeutungen*. Rather than discussing the statement 'A function never takes the place of a subject' itself, however, Frege converts that statement into the statement 'A function name never takes the place of a proper name', and proceeds to elucidate this statement, providing us with an example of the 'talking about signs' that he has in mind.[19]

In his reply to Russell, then, we find Frege juxtaposing two statements:

> (A) A function never takes the place of an object.
> (B) A function name never takes the place of a proper name.

Russell objected that (A) is self-contradictory, from which he infers that Frege's function/object distinction is self-contradictory, and must be abandoned. In response to Russell's objection, Frege says that we must turn our attention to (B). I believe that this elucidatory turn from (A) to (B) – what I will more generally call Frege's *elucidatory turn to signs* or to *language* – lies at the core of Frege's conception of the logical categories. It is a turn that we could already discern in Frege's first reply to Russell's objection. In responding to Part I of Russell's first formulation of his objection, for instance, we found Frege clearing up a confusion between sign and *Bedeutung* in Russell's treatment of the statement 'ξ can never take the place of a proper name', which led Frege to advance an exact instance of (B):

> '$(\,) \cdot 3 + 4$' can never take the place of a proper name (PMC, 135; BW, 218).

[19] Note that Frege's original statement from the *Grundgesetze* that prompted Russell's objection was exactly this: "A function-name can never take the place of a proper name" (GGA, I, §23).

Against Russell, Frege argues that such an instance of (B) is not self-contradictory in the way (A) is. When Russell replies by asking, in effect, 'but what about (A)?'[20], Frege responds by reiterating that we should turn our attention to (B).

We can also discern this turn to signs in Frege's replies to parts II and III of Russell's first formulation. Having introduced the second-level concept name '$\acute{\varepsilon}\varphi(\varepsilon)$', Frege turns Russell's attention – not to the *Bedeutung* of that sign – but to the sign '$\acute{\varepsilon}\varphi(\varepsilon)$' itself:

> In a conceptual notation we can introduce a precise expression for what we mean when we call something a function (of the first level of one argument), e.g.: '$\acute{\varepsilon}\varphi(\varepsilon)$'. Accordingly, '$\acute{\varepsilon}(\varepsilon \cdot 3 + 4)$' would express precisely what is expressed imprecisely in the proposition '$\xi \cdot 3 + 4$ is a function'. Whatever we now put in place of '$\varphi(\,)$', we always get a true proposition because we can only put in names of functions of the first level of one argument, for the argument place here is of the second kind (PMC, 136; BW, 218).

Frege is not saying: only first-level functions of one argument fall under $\acute{\varepsilon}\varphi(\varepsilon)$. Rather, he is saying: only first-level function names of one argument fit the argument place of '$\acute{\varepsilon}\varphi(\varepsilon)$'. He is not discussing the *Bedeutung* of the sign '$\acute{\varepsilon}\varphi(\varepsilon)$', but rather how it functions in the *Begriffsschrift* notation. Similarly, in his reply to part III, the turn to signs is made explicit: "Instead of using the imprecise expression 'ξ is a function', we can say: '"$(\,) \cdot 3 + 4$" is a function name'" (PMC, 136). In his subsequent argument, moreover, Frege invokes the indispensability of function names. Thus, we find Frege doing in his first reply to Russell exactly what he says we should be doing in his second reply: turn our attention to signs. With regards to the specific example of '$\acute{\varepsilon}\varphi(\varepsilon)$', we can say: a proper understanding of the logico-philosophical phenomenon of second-level subsumptions must go through an understanding of the *Begriffsschrift* sign '$\acute{\varepsilon}\varphi(\varepsilon)$', as Frege presented it. More generally, a proper logico-philosophical understanding of the logical stratification of the realm of *Bedeutungen* must go through a logico-philosophical understanding of signs, so that the latter figures as an essential moment in the former. The aim of the next chapter is to develop an account of this understanding of signs.

[20] Recall Russell's phrase "If we leave aside names altogether and speak merely of what they mean" (PMC, 138). The 'merely' here is indicative of an approach to philosophical logic that is utterly alien to Frege's. Insofar as we are engaged in philosophical logic, there is no such thing, for Frege, as *merely* speaking of what names mean. This will become clearer in the course of the subsequent chapters.

Frege's turn to signs is not, however, the whole story. There is a way in which we cannot but sympathize with Russell's insistence that it is statements such as (A) that are at issue, not (B). It is the stratification of the *realm of Bedeutungen* into functions and objects that must be elucidated, not the stratification of the *Begriffsschrift* into function names and proper names. That sympathy will only be strengthened once we remember that Frege himself does not hesitate, throughout his *oeuvre*, to advance statements such as (A) as purported commitments of his conception of the logical categories. In telling us not to worry about (A), but instead to turn to (B), is Frege not simply evading the crucial issue? Addressing this worry requires showing that there is a suitable logico-philosophical relation between (A) and (B) which is such that an understanding of (A) must indeed go through an understanding of (B). That such a relation is indeed at issue, is confirmed in the final piece of text from Frege's correspondence with Russell that remains to be discussed:

> This difference between the signs corresponds to a difference in the realm of meanings; although it is not possible to speak of it without turning what is in need of completion into something complete and thus falsifying the real situation. We already do this when we speak of 'the meaning of 'is a square number''. Yet the words 'is a square number' are not meaningless. The analysis of the proposition corresponds to an analysis of the thought, and this in turn to something in the realm of meanings, and I should like to call this a primitive logical fact. This is precisely why no proper definition is possible here (PMC, 142; BW, 220).

We will discuss this passage in detail in the final chapter. Frege talks about a correspondence between the distinction between function names and proper names, on the one hand, and the distinction between functions and objects, on the other hands. Similarly, he talks about a correspondence between the analysis of the proposition and – via an analysis of the thought – something in the realm of meanings. Frege's elucidatory turn to language, then, is accompanied by a conception of how the logical stratification of the *Begriffsschrift* corresponds to the logical stratification of the realm of *Bedeutungen*. It is this elucidatory notion of correspondence that must account for the way in which an understanding of the logical stratification of the realm of *Bedeutungen* – its functionality – must go through an understanding of the logical stratification of the *Begriffsschrift*. To develop an understanding of this elucidatory notion of correspondence that accompanies Frege's elucidatory turn to language will be the final task that awaits us, a task that will accordingly be taken up in the final chapter of this dissertation.

Let me end by noting that – although their correspondence continued – Russell never responded to Frege's second reply to his objection. Given what has already been said, this should not come as a surprise: Russell's approach to philosophical logic simply has no room for something like Fregean elucidation, let alone for the idea that this essentially involves a turn to language. As Kremer puts it in the course of explaining why Russell never came to fully appreciate the radically innovative character of Frege's *Begriffsschrift:*[21]

> What prevented Russell from doing this was his conception of 'symbolic logic' as only, insofar as it is *logic*, 'accidentally' *symbolic*. 'Logic' in *Principles* is essentially independent of any particular system of notation. Frege's 'logic' on the other hand, 'is almost all tied up with the concept-script' (Kremer, 2006, 184).

In the next chapter, we will begin our exploration of the way in which Frege's logic is almost all tied up with the *Begriffsschrift*.

[21] Kremer's discussion issues from passages such as the following, where Russell is discussing the term 'symbolic logic': "The word *symbolic* designates the subject by an accidental characteristic, for the employment of mathematical symbols, here as elsewhere, is merely a theoretically irrelevant convenience" (PoM, §11). Such a sentiment is utterly alien to Frege's philosophical logic. See also my (Vanrie, 2020).

Chapter 4
Frege's linguistic turn

Towards the end of the previous chapter, we saw that Frege takes the elucidation of the logical categories to involve an elucidatory turn to language. Frege believes that our understanding of the logical stratification of the realm of *Bedeutungen* must go through an understanding of the corresponding logical stratification of the *Begriffsschrift*. This left us with at two-fold task. First, we must come to an understanding of the logical stratification of the *Begriffsschrift*. Second, we must come to an understanding of the sense in which this stratification 'corresponds' to the logical stratification of the realm of *Bedeutungen* and – relatedly – of the sense in which an understanding of the latter must 'go through' an understanding of the former. This chapter focuses on the first task, the next and final chapter on the second task.[1]

In its most basic formulation, the logical stratification of the *Begriffsschrift* consists in the fact that *Begriffsschrift* signs divide into proper names, first-level function names, second-level function names, and so on, depending on their argument places. To understand this logico-philosophical phenomenon of the logical stratification of the *Begriffsschrift*, then, we must come to understand the nature of such distinctions between *Begriffsschrift* signs, distinctions that – in some sense to be specified – correspond to the logical category distinctions. It may be thought that nothing could be more straightforward: what we have here, are distinctions between syntactic categories. The rules of syntax of the *Begriffsschrift* are such that the *Begriffsschrift* signs divide accross distinct syntactic categories that give rise to the specific ways in which they have to be

[1] As could be expected, it will turn out that these two tasks are deeply intertwined.

combined so as to generate well-formed formulas. Moreover, these syntactic categories are set up by Frege in such a way as to match his ontological categories of objects, first-level functions, etc. The requisite 'correspondence' then, simply consists in this sort of match between the syntax of the *Begriffsschrift* and its semantics.

On such a view, the logical stratification of the *Begriffsschrift* is understood in terms of the contemporary distinction between syntax and semantics. I will seek to argue, however, that such a reading is unfaithful to Frege's philosophical logic. Frege's elucidatory turn to language is not a turn to either syntax or semantics, so that Frege's conception of the logical stratification of the *Begriffsschrift* cannot be understood in terms of the contemporary syntax/semantics distinction.[2] To show this, we must undertake a sustained investigation of Frege's conception of the *sign*, a topic that is mostly neglected in the literature. One reason for this neglect, I suspect, is that readers have tended to implicity assume that Frege's conception is indeed not markedly different from what is yielded by the logico-philosophical framework of the contemporary syntax/semantics distinction. Although explicit discussions of Frege's conception of the sign are sparse, I suspect that practically all readers of Frege, when prompted, would construe Frege's conception of the sign in a way that reads back into Frege the contemporary syntax/semantics distinction. Indeed, in discussing one of the rare explicit discussions of Frege's conception of the sign – that of Michael Kremer[3] – we will see that it is harder to avoid doing so than one may at first suppose.

I will begin by discussing two conceptions of the sign which I will call the *syntactic conception of the sign* and the *semantic conception of the sign*. On the syntactic conception, signs are uninterpreted syntactic units, governed by rules of syntax. On the semantic conception, signs are interpreted syntactic units that have been assigned a *Bedeutung*. Both conceptions are squarely situated within the logico-philosophical framework of the contemporary syntax/semantics distinction and the notion of interpretation that comes with it. Both, I will argue, are inadequate as readings of Frege, because they fail to do justice to a commitment of Frege that comes out, *inter alia*, in his discussions of formalism,

[2] I add the qualifier 'contemporary' because – although Frege himself never uses the term 'syntax' – it does appear, for instance, in Wittgenstein's *Tractatus* but possesses there a wholly different meaning than the contemporary notion. See e.g. (TLP, 3.325). As Wittgenstein uses the term, it *can* be said to be intimately related to Frege's philosophical logic, but this does not bring Frege any closer to the contemporary framework.

[3] That Kremer is sensitive to the importance of the sign in Frege's philosophical logic, I suspect, relates to his comparison of the roles of *logical symbolism* in Frege's and Russell's conceptions of logic that was referred to at the end of the previous chapter (Kremer, 2006, §V).

which I will call the *internality of its use to the sign*. In trying to bring out how the formalist completely loses sight of what is properly called a sign, Frege says that a sign must be taken to have what he calls the *purpose of signifying*. Nothing which does not have the purpose of signifying, is properly called a sign. Formalist *figures* – as Frege call them – fail this criterion.

Frege's commitment to the internality of its use to the sign has far-reaching philosophical implications that, I believe, have yet to be fully grasped by Frege scholars. In more general terms, we can say that the importance of Frege's conception of the sign for his philosophical logic remains to be properly recognized. Even scholars who are anxious to distinguish Frege's philosophical logic from contemporary approaches, seem not to have fully realized how this also requires distinguishing Frege's conception of the sign from the sort of conceptions of the sign that arise within a contemporary logico-philosophical framework. If nothing else, I hope that this chapter succeeds in putting on the agenda Frege's conception of the sign as an important topic for historical-philosophical investigation.

Properly doing justice to the internality of its use to the sign – and thereby to the sign's purpose of signifying – means to see that its use in the expression of judgment is internal to that which is used in the expression of judgment, in such a way that there is *no* order in the sign that can be taken to be constituted logically independently of its use in the expression of judgment, as the order of the syntactic unit and the order of the figure are. It also means to see that the logically fundamental case – which I will call the *full-blooded case* – is that of the occurrence of a sign that signifies in the expression of judgment, relative to which all other occurrences of signs must be seen as what I will call *privative cases*. To properly understand the internality of its use to the sign, is to understand the logical priority of the full-blooded case, which is in turn to be faithful to Frege's judgment-first approach to philosophical logic.

Once Frege's conception of the sign has been brought into view, I will turn to a discussion of the logical stratification of the *Begriffsschrift*. I will show how it is captured by a notion of *fit* that Frege employs in the *Grundgesetze*, which is such that – for instance – proper names *fit* the argument places of first-level function names. In line with the internality of its use to the sign, this notion of fit must be understood as pertaining – in its logically fundamental use – to full-blooded occurrences of signs in the expression of judgment. It characterizes the way in which *Begriffsschrift* judgments ramify into full-blooded occurrences of signs that fit together. It is in so fitting together, that signs realize

their purpose of signifying. In this way, we arrive at what I will call the *compositionality* of the *Begriffsschrift*, which consists in these manifold ways in which *Begriffsschrift* signs fit together in the expression of judgment, thereby realizing their compositional nature. It is this logico-philosophical phenomenon of compositionality, moreover, that properly accounts for the logico-philosophical unity of the *Begriffsschrift*. Compositionality *is* the form of logico-philosophical unity of the *Begriffsschrift*, which *is* its logical stratification.

Before we begin, I should point out the profound indebtedness of this chapter to Section XIII of (Conant, 2020), where Conant presents a thoroughgoing discussion of the notion of the sign as it figures in Wittgenstein's *Tractatus*. My account in this chapter may be read as an attempt to adapt Conant's approach to Frege. To those who are familiar with Conant's discussion, its influence will be in evidence throughout this chapter.[4] It is Conant who has brought to life for me the possibility of thinking about Frege's conception of the sign in the way that I, in turn, seek to bring to life for my own readers.

4.1 Frege's turn in action

Let us start by acquiring an initial feel for how Frege executes his turn to signs by going over some characteristic passages. We already encountered this passage from his correspondence with Russell:

> We can analyse the proposition '3 is a prime number' into '3' and 'is a prime number'. These parts are essentially different: the former complete in itself, the latter in need of completion. Likewise, we can analyse the proposition '4 is a square number' into '4' and 'is a square number'. Now it makes sense to fit together the complete part of the first proposition with that part of the second proposition which is in need of completion (that the proposition is false is a different matter);

[4] Other precedents that are a bit further removed from my account may be found in (Bronzo, 2017) (Moore, 2019, Chapter 4) (Gustafsson, 2020). Although the sign is not its main focus, (Narboux, 2014) also contains some highly pertinent reflections. Having pointed out Conant's influence, I should perhaps add the *caveat* that one would be mistaken to think that there are no substantial philosophical differences between Frege's conception of the sign and that of Wittgenstein in the *Tractatus*. Although there can be no doubt that Wittgenstein was deeply influenced by Frege's conception of the sign, he sought to transform that conception in order to overcome what he regarded as its lingering defects, some of which will be touched upon below.

> but it makes no sense to fit together the two complete parts; they will not hold together; and it makes just as little sense to put 'is a square number' in place of '3' in the first proposition (PMC, 141-142; BW, 220).

Instead of discussing the sentence 'A function never takes the place of an object' – as Russell requested – Frege turns to discuss how a function *name* can never take the place of a proper *name*. Similar sorts of passages are found throughout Frege's *oeuvre* whenever he discusses the function/object distinction. Here are some examples:

> Even so, the two parts of the proposition [i.e. 'Two is a prime number'] are still essentially different; and it is important to realize that this difference cuts very deep and must not be blurred. The first constituent, 'two', is a proper name of a certain number; it designates an object, a whole that no longer requires completion. The predicative constituent 'is a prime number', on the other hand, does require completion and does not designate an object. I also call the first constituent saturated; the second, unsaturated (CP, 281; KS, 269; 371).

> The sign for a function is 'unsaturated'; it needs to be completed with a numeral, which we then call the argument-sign. We see this also with the root-sign, with the logarithm-sign. A functional sign cannot occur on one side of an equation by itself, but only when completed by a sign that designates or indicates a number. Now what does such a complex stand for consisting of a functional sign and a numeral, e.g. 'sin 1', '$\sqrt{1}$', 'log 1'? A number each time. We thus get numerical signs composed of two dissimilar parts, an 'unsaturated' part being completed by the other one (CP, 290-291; KS, 278; 663-664).

> If we call the parts of the sentence that show gaps unsaturated and the other parts complete, then we can think of a sentence as arising from saturating an unsaturated[5] part with a complete part. *The complete part of the sentence I call a proper name, the unsaturated part a concept-name* [Frege's emphasis] (PW, 201; NS, 217).

> The sign 'sin' which we say designates the sine function, only occurs in mathematics in close combination with other signs, as in 'sin 10°', 'sin 10° 11′',

[5] The translation erroneously has 'saturated' here.

> '$\sin \alpha$', and so is a sign in need of supplementation and therein is different from a proper name (PW, 271; NS, 290).

In each case, in the course of elucidating the function/object distinction, Frege invites us to consider the distinction between complete and unsaturated *signs*, between proper names and function names.

It could be thought that such discussions of signs must be taken as merely optional additions that may be of some assistance, but which ultimately play no essential role in elucidating the logical categories. What is essential to that activity, the idea goes, are statements such as 'Functions are unsaturated' and 'Objects are complete', statements that do not concern signs. Such statements may, of course, be supplemented with a consideration of signs, but such supplementation is not mandatory.

This is belied, however, by the striking way in which discussions of signs consistently appear wherever Frege is elucidating the function/object distinction. This includes – not only his more informal presentations – but also the *Grundgesetze*.[6] Here are two examples from that book:[7]

> The nature of the *function* lies therefore in that part of the expression [i.e. the expression '$(2 + 3 \cdot x^2) \cdot x$'] that is present without the 'x'. The expression of a *function* is *in need of completion*, *unsaturated*. The letter 'x' serves only to hold open places for a number-sign that is to complete the expression, and so marks the special kind of need for completion that constitutes the peculiar nature of the function just designated" [Frege's emphases] (GGA, I, §1).[8]

---

[6] In the *Grundgesetze*, Frege explicitly states that he uses quotation marks to talk about signs (GGA, I, 4). Investigating the sort of understanding of quotation that comes with Frege's conception of the sign – and seeing how it relates to contemporary accounts such as the seminal (Davidson, 1979) – is a task I cannot take up here. I can only blankly state my conviction that the sort of overcoming of the contemporary syntax/semantics distinction that I will present as essential to a proper understanding of Frege's conception of the sign comes with a similar overcoming of certain presumptions at work in most contemporary accounts of quotation. This is not surprising, of course, given that such accounts are usually wedded to the syntax/semantics distinction.

[7] Note that Frege's turn to signs is not limited to the logical categories, but also appears in connection with, for instance, the judgment-stroke and the horizontal, as in (GGA, I, §5).

[8] The first sentence from this passage and the general order of Frege's presentation could not make it more clear that Frege takes a consideration of signs to be essential to our grasp of the nature of the function. That Frege does not simply start by saying straight up that functions are unsaturated – but instead introduced this claim through a consideration of function names – runs counter to what many readings of Frege would lead one to expect. Indeed, we will see below how a certain kind of approach can only see Frege's focus on signs as *besides the point*.

> A function-name can never take the place of a proper name, because it will involve empty places corresponding to the unsaturatedness of the function. If we say 'the function $\Phi(\xi)$', then we should never forget that 'ξ' belongs with the function-name only by way of making the unsaturatedness recognizable (GGA, I, §21).

We will encounter further examples below. Since the *Grundgesetze* is meant to constitute a rigorous and considered presentation of Frege's philosophical logic, we should either not expect to find such merely optional discussions intermixed with the core presentation of Frege's logico-philosophical views, or we should expect to find them explicitly marked as optional. If Frege did not believe that the understanding of the logical stratification of the realm of *Bedeutungen* that we seek in philosophical logic essentially involves an understanding of the logical stratification of the *Begriffsschrift*, his exposition in *Grundgesetze* (and in many other places) would be misleading. If he did believe this, however, what we find is exactly what we would expect to find: a presentation of Frege's philosophical logic that is permeated with discussions of signs.

It could be objected that I am making too much of a fuss here. What Frege is doing in discussing *Begriffsschrift* signs, it will be said, is nothing else than laying out the syntax of the *Begriffsschrift*.[9] Rather than the prevalence of such discussions of signs being surprising, then, it is exactly what we should expect from a presentation of his logical system. The problem, however, is that such presentations of syntax are irrelevant to the elucidation of the logical categories, as I will argue below. On such a view, we would need to distinguish between the presentation of the syntax of the *Begriffsschrift*, on the one hand, and the elucidation of the logical category distinctions, on the other hand. But it is Frege's invocation of signs with regards to the latter that we wish to understand, and this cannot be understood in terms of discussions of mere syntax, as we will see.

[9] Compare: "Frege was treating his theory as an uninterpreted set of syntactic operations on strings of symbols" (Stanley, 1996, 63).

4.2 *Begriffsschrift* vis-à-vis ordinary language

Another preliminary question must be addressed. Among the passages considered above, there is a salient difference between those that occur in the *Grundgesetze* and those that occur elsewhere in Frege's *oeuvre*. In the *Grundgesetze*, Frege is concerned with *Begriffsschrift* signs specifically. In the latter, Frege mostly discusses ordinary language expressions. This raises the question whether we are supposed to turn our attention to *Begriffsschrift* signs, or whether ordinary language suits Frege's elucidatory purposes just as well. Ultimately, I claim, the former. It is the logical stratification of the *Begriffsschrift* that must ultimately be understood, for reasons that will become apparent. At the same time, however, this does not mean that ordinary language does not have its own role to play in Frege's elucidatory project. When Frege is trying to bring his readers to an initial grasp of the distinction between complete and unsaturated signs, he tries to show them how we can already find what may be called a crude version of that distinction in ordinary language itself. This is what is happening in some of the passages above, including parts of Frege's correspondence with Russell, and what we saw happening in the stretch of text from *Über Begriff und Gegenstand* that was discussed in Chapter 2. We have here what may be called a *maieutic* exercise: Frege aims to bring us to realize that our mastery of ordinary language already involves an incipient recognition of the distinction between complete and unsaturated signs. In some cases, this will proceed through a consideration of mathematical language, wherein the distinction can be more clearly discerned.[10] At the same time, it is essential to Frege's philosophical logic that a consideration of mathematical language is not mandatory, because it is essential to Frege's philosophical logic that the distinction between complete and unsaturated signs is not specific to mathematical language, but governs statements about *any* subject matter, no matter how

[10] If asked, this is undoubtedly where Frege would place the autobiographical germ of his own grasp of the function/object distinction. Whether it is also the place where we can find the germ for Frege's idea that a fully developed logico-philosophical understanding essentially involves a turn to what we may call a logically purified notation such as the *Begriffsschrift*, is a more subtle question that I cannot address here. Compare: "The advantages of perspicuity and precision are so great that many investigations could not even have been made without a mathematical sign language" (PMC, 33; BW, 58). As professor Jean-Philippe Narboux has emphasized to me, the development of the mathematical notion of a function was itself a complex historical process that was still fully ongoing when Frege entered the scene. A thoroughgoing investigation of that mathematical-historical background – and the way in which Frege's generalized notion of a function engages with that background – would undoubtedly shed further light on Frege's views. I flag this here as a task for further research.

humdrum. It is also essential to Frege's philosophical logic, however, that we cannot rest content with the inchoate grasp of that distinction that is implicit in our mastery of ordinary language, but that this grasp must be further purified into a fully developed understanding of the logical stratification of the *Begriffsschrift*, wherein the distinction between complete and unsaturated signs comes to be fully unfolded.[11]

4.3 The syntactic conception of the sign

Let us now embark on an investigation of Frege's conception of the sign. As announced, I will start by discussing the sort of conceptions that one arrives at if one brings the framework of the contemporary syntax/semantics distinction to one's reading of Frege. One conception of the sign that naturally arises from that framework, is this: A sign is a *syntactic unit*, a rule-governed constituent of well-formed formulas.[12] Syntax consists in the rules that specify how such signs are to be combined into well-formed formulae. As the subject matter of syntax, signs are uninterpreted, they have no meaning. Meaning is instead conferred upon signs through interpretation, which is the business of semantics. Such interpretations are external to the signs themselves, which are purely syntactical in nature. A sign is the syntactical unit that it is independently of how it is (or is not) interpreted. In this way, the order of the sign is conceived as an order that is constituted logically independently of the order of interpretation, independently of the conferral of *Bedeutungen* upon those signs. Call this the *syntactic conception of the sign*.

The syntactic conception of the sign runs counter to how Frege himself thinks we should conceive of the sign. A good place to start is Frege's critique of formalism, in which the notion of the sign takes centre stage. Formalists believe that signs – rather than the numbers themselves – form the subject matter of arithmetic, a view that can also be

[11] What I have said here may be compared to Conant's description of what he calls two *circles* that are involved in Frege's logico-philosophical elucidations (Conant, 2020, 450-451). Roughly, the first puts us in a position to enter the *Begriffsschrift* as such, the second pertains to acquiring a fully developed logico-philosophical understanding of the *Begriffsschrift*. Frege's discussions of ordinary language may be taken as pertaining primarily to the first circle.

[12] I leave it to the reader to account for the usual type/token intricacies that pertain to these standard conceptions of the sign. To what extent something like a type/token distinction is adequate to Frege's own conception of the sign as I will articulate it, is a question that I must leave as a topic for further investigation.

characterized as taking numbers to *be* signs. Arithmetical statements – having those signs as their subject matter – are then conceived as articulating rules for the manipulation of such signs. Such rules constitute what Frege calls the "theory of formal arithmetic" (GGA, II, §93), which is to be distinguished from the game of formal arithmetic itself, i.e. from the actual manipulation of the signs in accordance with the rules. The hackneyed equation '$5 + 7 = 12$', for instance – if taken as belonging to the theory of formal arithmetic – states something along the lines that we may, in the course of doing formal arithmetic, replace any occurrence of the sign '$5 + 7$' with the sign '12'. Such a replacement constitutes a valid move in the game of formal arithmetic. According to the formalist, the fact that the signs of arithmetic may also be used to express judgments, is neither here nor there. Formalist arithmetic is meant to proceed entirely independently of the possibility of using arithmetical signs in the expression of judgment. The only 'judgments' that are necessary are the rules that specify how arithmetical expressions are to be formed and manipulated, but these are only judgments in an attenuated sense at best, since they merely stipulate the rules of the game of formal arithmetic.

We see that the formalist account of arithmetic is closely similar to the syntactic conception of the sign. As is well known, Frege relentlessly criticized formalism throughout his career, in a mode that can be best described as exasperated disbelief that formalism was even taken seriously as a possible view. Here is Frege:

> Despite the emphatic assertion that the signs are empty and that it is they themselves that are the numbers, in the background there always hovers the thought that they do signify [*bedeuten*] something and that it is these contents of signs that really are the numbers. This is indicated by the use of the word 'sign'; for is something that does not designate [*bezeichnen*] anything and does not even have the purpose [*Zweck*] of designating anything a sign? In such a case I shall use the word 'figure' [*Figur*], so as to avoid confusing anyone by means of a wrong expression (CP, 114-115; KS, 105; 97).

Frege's use of the German term '*bezeichnen*' is important, since it contains the word '*Zeichen*' as a root. To '*bezeichnen*' just is to do what signs do, and Frege charges the formalist with losing sight of what signs are supposed to do. What is it that signs are supposed to do? In many cases, the answer is: *Bedeuten*. That is, for instance, what number signs are supposed to do. *Bedeuten*, however, is not what all signs are supposed to do. The notion of '*bezeichnen*' – understood as what signs are supposed to do – is broader than this. I will use the term 'signify' for the broader notion, which nicely takes over the

etymological connection between '*bezeichnen*' and '*Zeichen*'. I thereby depart from the translation of '*bezeichnen*' as 'designate', which I take to be too closely associated with the notion of *bedeuten*, so that it is less apt to capture the more general notion of signifying that is at issue here.

In the case of names – proper names and function names – to signify is to *bedeuten*. That is what names do. As is well known, however, Frege also employs signs such as the judgment stroke, German letters, Roman letters, Greek letters, brackets, and so on. Such signs do not *bedeuten*, as Frege himself makes clear:

> I will not call the German, Roman and Greek letters occurring in concept-script *names* since they are not to refer to anything. [...] Thus, I call a *proper name* or *name* of an object any sign, be it simple or complex, that is to refer to an object, but not such a sign which merely indicates an object (GGA, I, §26).

Frege describes his letters as indicating an object. Indicating an object, then, is a mode of signifying that is distinct from *bedeuten*. The way in which the judgment-stroke signifies, constitutes another such mode. These modes have to be analysed on a case-by-case basis.[13] What is true about all such modes, however, is that they consist in the signs in question being used in the expression of judgment.[14] However it is that a sign signifies, it is through its being used in the expression of judgment that it signifies. In committing himself to the claim that a sign must at least have the purpose of signifying, Frege is forging an essential connection between the sign and its use in the expression of judgment. It is this connection that I will call *the internality of its use to the sign*. In the case of names – which will be most important in what follows – we can talk about *the internality of its Bedeutung to the name*. To come to an understanding of Frege's conception of the sign – and thereby to an understanding of Frege's conception of the logical stratification of the *Begriffsschrift* – we must come to an understanding of this internality.

As Frege sees the formalist, they purport to conceive of signs as constituted independently of their being used to signify in the expression of judgment. They purport to conceive of names of numbers, for instance, as constituted independently of their having *Bedeutung*. In this way, however, the formalist has not conceived of signs at all, but rather of what Frege calls mere 'figures'. It is important to emphasize that Frege does not

[13] For some insightful reflections on Frege's earlier and later use of Greek letters, see (Kremer, 2019).
[14] For current purposes, definitions would have to be seen as a species of judgment. The case of definitions is, in any case, less germane to my concerns.

distinguish figures from signs that signify as such, but from signs that have the *purpose* of signifying. In this way, Frege makes it clear that there is a distinction between signs that signify, on the one hand, and signs that merely have the purpose of signifying without actually signifying, on the other hand, and that it would equally be a mistake to think that formalist figures can be construed as merely having the purpose of signifying without actually signifying. To properly appreciate Frege's critique of formalism, is to appreciate why he thinks that the distinction between a mere figure and what at least has the purpose of signifying is a distinction that matters, in such a way that to arrive at the formalist notion of a figure does not even bring us to the latter notion, and thereby constitutes no progress at all in the development of a logically proper conception of the sign.

It should be noted that this notion of having the purpose of signifying reoccurs in Frege's later *Grundgesetze* discussion of formalism:[15]

> In formal arithmetic, their role [i.e. of signs] is different: they are not supposed to designate something else; rather they are themselves the object of concern. [...] Accordingly, we had better term them figures here, since the purpose[16] [*Zweck*] of designating [*bezeichnen*] something does not come into consideration at all (GGA, II, §100).

The fact that Frege twice invokes this notion of having the purpose of signifying in his discussions of formalism gives us good reason to believe that Frege took it to be essential to an adequate understanding of the sign.

A defender of the syntactic conception of the sign as a reading of Frege could try to capitalize on Frege's distinction between a sign that merely has the purpose of signifying and a sign that actually signifies by claiming that it captures exactly the distinction between an uninterpreted syntactic unit and an interpreted syntactic unit. Whereas the latter signifies, the former does not signify but can be said to have the purpose of signifying because it is susceptible to interpretation. Such a reading, however, obliterates precisely Frege's distinction between formalist figures and signs: Since formalist figures are at least *susceptible* to interpretation, they too would have to be counted as signs, on the proposed reading. This shows that – rather than running parallel to a distinction

[15] We will consider a third passage in which it occurs below.
[16] I have altered the translation – which has 'aim' instead of 'purpose' – in order to preserve the uniformity with the previous passage. Frege uses the German term '*Zweck*' in both cases.

between uninterpreted and interpreted signs – Frege's distinction between a sign that actually signifies and a sign that merely has the purpose of signifying precludes it, which already reveals how Frege's conception of the sign sits uneasily with the syntax/semantics distinction and the notion of interpretation that comes with it, a point that will receive further corroboration throughout our discussion.

Frege's discussions of formalism are not the only place where his commitment to the internality of its use to the sign becomes apparent. In the *Grundlagen*, Frege responds to a suggestion by Hankel that we can introduce negative whole numbers by stipulating that signs of the form '$c - b$' (where $b > c$) are to satisfy equations of the form '$x + b = c$'. Frege writes:[17]

> Nevertheless, there is something to prevent us from regarding $(2 - 3)$ without more ado as a sign[18] [*Zeichen*] which solves the problem; for an empty sign is precisely no solution; without some content it is merely ink or print on paper, as which it possesses physical properties but not that of making 2 when increased by 3. Really, it would not be a sign at all, and to use it as one would be a mistake in logic (GL, §95).

Again, Frege refuses to acknowledge as a sign that which is 'merely ink or print on paper', i.e. that which is taken to be constituted independently of its being used in the expression of judgment. Frege is pointing out that the merely physical properties of ink or print on paper do not give us what is required of a properly logical notion of the sign: that of making 2 when increased by 3. In other words: it is essential to the sign '$(2 - 3)$' that it can be used in the expression of judgments such as '$3 + (2 - 3) = 2$'. Any conception on which its being the sign that it is in no way depends on such use would not be a conception of a *sign* at all, but that of a mere figure, and it is a mistake to think that such figures can be properly used in logic. This is not to say that signs are not material, that they are *not* printed on paper. The signs we use in the expression of judgment are indeed material. Rather, Frege is specifying the proper contours within which to think about such signs – contours that are such that its being used in the expression of judgment is internal to the

[17] It will be noticed that this passage, and the first discussion of formalism above, predate Frege's distinction between *Sinn* and *Bedeutung*. Below, we will discuss how that distinction figures into Frege's mature conception of the sign (of the name). These earlier passages, however, reveal that there is an important continuity in Frege's conception of the sign. For a further confirmation of this continuity, see Bar-Elli's pertinent reflections on Frege's purportedly 'metalinguistic' *Begriffsschrift* definition of identity (Bar-Elli, 2006).

[18] I have amended Austin's translation – which has 'symbol' – to preserve continuity with the other passages.

material sign. As soon as we think that there is an order of materiality in the sign which is constituted logically independently of its use, we have lost our grip on the sign. Instead, we will have on our hands something like the formalist figure, which Frege takes to be irrelevant to logic proper.

Another pertinent passage occurs in Frege's critical discussion of Hilbert's views, as presented and defended by Korselt:

> Mr. Korselt does not always appear to distinguish a proposition as what is sensibly perceptible, from the thought which is its sense. What I call a proposition *tout court* or a real proposition [*eigentliche Satz*] is a group of signs that expresses a thought; however, whatever has only the grammatical form of a proposition I call a pseudo-proposition [*uneigentliche Satz*]. Examples of the latter are often to be found as antecedent and consequent propositions of conditional propositional complexes (CP, 308; KS, 295; 377).

Here, Frege is discussing propositional signs. On the syntactic conception of the sign, a propositional sign is a syntactically well-formed formula, a concatenation of syntactic units in accordance with the rules of syntax. This notion of a well-formed formula is purely syntactical, and is thereby logically independent of the orders of *Sinn* and *Bedeutung*. Frege, however, makes it clear that his conception of the propositional sign is such that it essentially involves the order of *Sinn*, in a way that the syntactic conception cannot account for. Below, we will discuss how the *Sinn*/*Bedeutung* distinction figures in Frege's conception of the sign.

For now, let us consider this passage – and the notion of a *pseudo-proposition* that figures in it – in a bit more detail. One of Frege's main concerns is to show that what Hilbert presents as self-standing axioms – and thereby as self-standing propositional signs – must be construed as pseudo-propositions. Using a toy example, the idea goes as follows. Hilbert presents as a purported axiom an expression such as '$x > 1$', without specifying a *Bedeutung* (or *Sinn*) for 'x'. He then purports to infer from this axiom other expressions such as '$x > 0$', again without specifying a *Bedeutung* for 'x'. According to Hilbert, what we have here are syntactically well-formed formula that are susceptible to multiple interpretations. According to Frege, however, this is just a confused way of getting at what is properly dealt with by inferences from the general to the particular. In this case, the relevant general proposition is '$(\forall x)\big((x > 1) \supset (x > 0)\big)$', and here we see a logically proper use of the pseudo-propositions '$x > 1$' and '$x > 0$' under the scope of generality. In this way, pseudo-propositions constitute another example of signs that

signify in a way that is not *bedeuten*. What Hilbert calls an 'interpretation' in which, for instance, the number 3 is assigned as the meaning of 'x' is properly dealt with, according to Frege, by an inference from '$(\forall x)\big((x > 1) \supset (x > 0)\big)$' to '$(3 > 1) \supset (3 > 0)$'.

Here is Frege again:

> The word 'interpretation' is objectionable, for when properly expressed, a thought leaves no room for different interpretations. We have seen that ambiguity simply has to be rejected and how it may appear to be necessary because of insufficient logical insight. [...] On the basis of our understanding of the nature of Mr. Korselt's purely formal system it is easy to guess what Mr. Korselt means by 'interpretation'. When we proceed from the general theorem 'If $a > 1$, then $a^2 > 1$' to the particular one 'If $2 > 1$, then $2^2 > 1$' by means of an inference, then the pseudo-proposition '$a > 1$' corresponds to the proper proposition '$2 > 1$'. According to Mr. Korselt's usage, '$2 > 1$' or the thought of this proposition will be an interpretation of '$a > 1$'. As if the general proposition were a wax nose which we could turn now this way, now that. In reality, we have not an interpretation but an inference (CP, 315-316; KS, 301-302; 384-385).

Pseudo-propositions are especially apt to confuse because, as Frege put it in the first passage above, they have the grammatical form of a proposition, even though they are not used to express thoughts. Hilbert, seeking to exploit this grammatical form, nevertheless presents them as self-standing axioms, while at the same time wishing to leave them open to various interpretations. In this way, however, he loses the right to regard them even as pseudo-propositions in Frege's sense, and thereby as signs. According to Hilbert, we can present as an axiom a self-standing uninterpreted syntactic string that is subsequently given *Bedeutung* through an interpretation, in line with the syntactic conception. What confuses Hilbert, according to Frege, is that inferences from the general to the particular do feature pseudo-propositions that have the grammatical form of a proposition – and are genuine signs – without expressing thoughts. Since such pseudo-propositions do not express thoughts anyway, Hilbert wrongly believes that he can extract them from the contexts in which they signify, and replace inferences from the general to the particular with his notion of the interpretation of mere syntactic strings, while nevertheless preserving all that matters to their nature as signs. In this way, however, he has severed the internal connection between the sign and its use, so that he ends up with mere figures instead, which have no proper role to play in logic.

The syntactic conception of the sign also prevents a proper account of Frege's elucidatory turn to language. On the syntactic conception, the logical stratification of the *Begriffsschrift* must be understood as a stratification into syntactical categories determined by the rules of syntax that is entirely independent of the use of *Begriffsschrift* signs in the expression of judgment. More specifically, this yields a conception of the distinctions between *Begriffsschrift* names as being logically independent of the logical category distinctions.[19] It is this independence that undermines the idea that an understanding of the logical stratification of the realm of *Bedeutungen* must go through an understanding of the corresponding logical stratification of the *Begriffsschrift*. Because interpretation is logically independent of syntax, there is nothing internal to the order of syntax that has any bearing on the logical stratification of the realm of *Bedeutungen*. As such, rules of syntax can be set up entirely arbitrarily. In practice, of course, we will not set up just any system of rules whatsoever, but we will seek to do so in ways that suit our purposes. Such pragmatic constraints, however, are external to the order of syntax itself, and thereby do not flow from the nature of the sign, conceived syntactically.[20] To think that a grasp of the syntactical distinction between function names and proper names is essential to a grasp of the function/object distinction would constitute a straightforward case of illegitimately hypostasizing what is merely a contingent aspect of an arbitrarily set up syntactical system. Indeed, it is precisely this arbitrariness that has made possible – in contemporary logic – the proliferation of logical systems with wildly divergent rules of syntax, leading to the standard set of metalogical questions concerning soundness, completeness, categoricity, and so on.[21]

---

[19] Compare: "the distinction between proper names and expressions of other types must be one that can be drawn in wholly linguistic terms, without the necessity for any scrutiny of the things for which the respective expressions stand" (Dummett, 1981, 57). Dummett is a good example of an interpreter who reads Frege in terms of the contemporary syntax/semantics distinction, and who thereby fails to get into view what I take to be a properly Fregean conception of the sign.

[20] Compare: "According to Thomae, the question, 'what does arithmetic require of the number', is likely to be answered like this: in arithmetic we require of the numbers only their signs, which are however not treated as such but rather as figures; and we require rules according to which these figures are manipulated. Here, we do not extract these rules from the reference of the signs, rather we lay them down on our own absolute authority, reserving complete freedom in principle and acknowledging no necessity to justify these rules, while admittedly we exercise this freedom with an eye to possible applications, for without them arithmetic would be a game and nothing more" (GGA, II, §94).

[21] This is what I would call the primordial *Carnapian* heritage of contemporary analytical philosophy, which is, I think, more deeply indebted to Carnap in its way of approaching such questions than many have realized. In this way, the commonplace idea that Quine 'won' the debate with Carnap for posterity is a misconception that arises from a narrowly epistemological construal of their debate, to the neglect of the matters of philosophical

There is much debate about the extent to which Frege himself was engaged in metalogical reflections along contemporary lines. Some claim that we should indeed read Frege as taking an incipiently metalogical approach at certain crucial junctures in his work (Dummett, 1981) (Stanley, 1996) (Heck, 2010).[22] Others demur, and take it that there are fundamental historical-philosophical reasons for regarding such an approach as incompatible with Frege's philosophical logic (Ricketts, 1986) (Goldfarb, 2001) (Weiner, 2005).[23] What neither party to the debate seems to have fully realized, however, is how the issue crucially depends on one's construal of Frege's conception of the sign. If one wishes to regard Frege as engaged in metalogical reflections along contemporary lines, one must ascribe to him a conception of the sign that fits the syntax/semantics distinction. Moreover, as soon as one ascribes to Frege such a conception of the sign, it becomes difficult to argue that Frege did not in fact engage in such reflections.[24] On the other hand, if one can show that Frege's conception of the sign is incompatible with the syntax/semantics distinction – as I seek to do – one has cut the ground from under the metalogical reader's feet, confirming that there is a much wider philosophical gap separating Frege's conception of logic from contemporary approaches than many have realized.

---

logic that were equally at stake. Note that what I have in mind here cuts deeper than the standard construals of Quine's opposition to intensionality, construals that I would claim remain too superficial to do justice to the real underlying issues in philosophical logic. As those familiar with my postdoctoral research proposals – currently under review at the relevant bureaucratic-scientific institutions – know, I hope to explore these themes in future work. Note, finally, that reading Wittgenstein's notion of the arbitrariness of grammar as a version of the Carnapian arbitrariness of syntax – which gave rise to Carnap's notorious *Principle of Tolerance* – is to completely misunderstand Wittgenstein's views, who was concerned to *resist* such an approach as much as Frege was.

[22] Here is Dummett, for instance: "Frege would therefore have had within his grasp the concepts necessary to frame the notion of the completeness of a formalization of logic as well as its soundness [...] but he did not do so" (Dummett, 1981, 82). See also pp. 89-90.

[23] Here is Ricketts, for instance: "The notion of a logical schema that admits of multiple interpretations is foreign to Frege's thought. Nor is it possible, through reasonable emendations, to read the contemporary view back into Frege. For the contemporary view requires the ineliminable use of a truth predicate. Such a use is antithetical to Frege's conception of judgment. This conception of judgment precludes any serious metalogical perspective and hence anything properly labelled a semantic theory" (Ricketts, 1986, 76).

[24] At least, this is how it seems to me. A defence of this claim would lead us too far astray.

4.4 The semantic conception of the sign

I have argued that the internality of its use to the sign invalidates the syntactic conception of the sign as a reading of Frege. It may seem, however, that there is an easy way to fix the issue. Let us focus on names and the internality of its *Bedeutung* to the name. Although all will agree that names conceived as mere syntactic units do not have *Bedeutung*, interpreted names do. Accordingly, it may seem that we can read Frege as holding the view that names are *interpreted* syntactic units. This is, in effect, what Stanley does in claiming that Frege's critique of formalism comes down to the complaint that "no interpretation is given to the signs of a formal system" (Stanley, 1996, 60).[25] Correctly formulated, Stanley's suggestion would be that Frege complains that no interpretation is given to the *figures* of a formal system. On this conception, uninterpreted syntactic units do not qualify as signs – they are mere figures – only interpreted syntactic units do. Call this the *semantic conception of the sign*. It is still squarely situated within the framework of the contemporary syntax/semantics distinction. The semantic conception may seem to do a better job at accounting for the internality of its *Bedeutung* to the name, since it takes it to be essential to a name that it is interpreted, which means that it is essential to a name that it has *Bedeutung*.

What sort of conception of the logical stratification of the *Begriffsschrift* do we get on the semantic conception? Let us take the distinction between proper names and function names as an example. We are no longer conceiving of that distinction as a distinction between syntactic categories, but instead as a distinction that pertains to interpreted syntactic units. As such, it is the distinction between signs that are interpreted as having functions as *Bedeutung*, on the one hand, and those that are interpreted as having objects as *Bedeutung*, on the other hand. A proper grasp of this distinction must yield the fact that function names and proper names cannot replace each other in *Begriffsschrift* judgments. On the syntactic conception, this was to be construed as a merely arbitrary matter of syntactical stipulation. On the current conception, however, this is construed as issuing

[25] Stanley believes that we can more or less find the contemporary notion of interpretation in Frege: "Frege certainly seems to use the term ['*Bedeutung*'] in a classically semantical manner" (Stanley, 1996, 46). There is a nice irony here, since Stanley is exactly the sort of reader who would complain that Frege uses the term '*Bedeutung*' in categorically ambiguous ways – as we saw Wright, Jolley, and Black complain in Chapter 2. To take seriously Frege's logically stratified use of the term '*Bedeutung*' is already to see that he cannot be using it in a classically semantical manner.

from the *interpretation* of function names and proper names. In other words: it is *because* function names designate functions and proper names designate objects, that they cannot replace one another. Whatever is interpreted as designating a function, cannot be replaced by what is interpreted as designating an object.

In this case, the problem is not that the resulting conception of the logical stratification of the *Begriffsschrift* has no bearing on the logical categories. It is instead, that the logical stratification of the *Begriffsschrift* is rendered entirely parasitic on the logical stratification of the realm of *Bedeutungen*. On the semantic conception of the sign, *Begriffsschrift* signs are seen as inheriting the logically prior logical category distinctions of the realm of *Bedeutungen* through their interpretation. Whatever understanding we have of the logical stratification of the *Begriffsschrift* becomes entirely parasitic on a logically prior understanding of the logical stratification of the realm of *Bedeutungen*, so that the latter cannot, in fact, go through the former in the way that Frege's elucidatory turn to language requires.

This problem with the semantic conception is unwittingly made manifest in discussions of Frege that take their departure from the contemporary syntax/semantics distinction. A good example is Wright, who has claimed that Frege is simply wrong in how he thinks about his interpretations of his *Begriffsschrift* signs. On Wright's view, there is no sharp distinction between functions and objects as Frege conceives it. Instead, functions must be conceived as constituting a kind of objects (Wright, 1998, 258ff.). It remains true, of course, that we cannot replace a proper name with a predicate in the expression of judgment, but this is not, on Wright's view, because there is a sharp concept/object distinction of the kind Frege defended. It is rather because concepts constitute a specific kind of objects that are *ascribed* by predicates, which distinguishes them from objects that cannot be ascribed, but only referred to by proper names.[26]

Wright's discussion clearly shows how, on the semantic conception, our understanding of the distinctions between *Begriffsschrift* signs comes to be seen as parasitic on a logically prior understanding of the logical categories themselves. It is the latter that informs our understanding of the distinction between proper names and function names, and it is this order of priority that makes possible the sort of disagreement that Wright takes himself to have with Frege concerning the nature of the distinction between proper names and

[26] In distinguishing only between ascription and reference, Wright does not tell us what it is that higher-level concepts are to be construed as doing, on his account.

function names. We could say that, on this approach, there is nothing *in* the *Begriffsschrift* that serves to distinguish between a Fregean and a Wrightian conception of the distinction between proper names and function names. Rather, it is the nature of the function/object distinction itself that determines the nature of the distinction between function names and proper names.

In this way, the semantic conception of the sign fails to do justice to the internality of its use to the sign. At this point, we must remind ourselves of a central commitment of Frege's philosophical logic:

> I do not begin with concepts and put them together to form a thought or judgment;
> I come by the parts of a thought by analyzing the thought (PW, 253).[27]

Frege takes what we could call a *judgment-first* approach to philosophical logic.[28] He considers the act of judgment to be logically fundamental. This is, in turn, connected to Frege's well known context principle: "never to ask for the meaning of a word in isolation, but only in the context of a proposition" (GL, x). That the act of judgment is logically fundamental, means that *Bedeutung* must be conceived as wedded to the notion of judgment, which means that the notion of a name's having *Bedeutung* is wedded to its being used in the expression of judgment.[29] *Bedeuten is* what a name does in the expression of judgment.[30]

On the semantic conception, syntactic units acquire *Bedeutung* through interpretation. Moreover, it is essential to this notion of interpretation that the semantic relation between a name and its *Bedeutung* is logically prior to that name's appearing in the expression of judgment. On the semantic conception, we put interpreted syntactic units

---

[27] Compare: "And so instead of putting a judgment together out of an individual as subject and an already previously formed concept as predicate, we do the opposite and arrive at a concept by splitting up the content of possible judgment" (PW, 17; NS, 18).

[28] For a canonical discussion of Frege's judgment-first approach to philosophical logic, see (Ricketts, 1986).

[29] Note that Frege formulated the context principle before he introduced the *Sinn*/*Bedeutung* distinction. I am here helping myself to the assumption that we can – within the context of his later thought – make sense of the context principle as applying to the *Bedeutung* of signs. For an indication of some of the complexities that surround this issue, see (Conant, 2020, 967-968n).

[30] I take my use of the notion 'expression of judgment' to accord with what Frege has in mind when he talks – in (GGA, I, §5) – about *concept-script propositions*. Such propositions are preceded by the judgment-stroke, so that they do not merely express thoughts, but *assert*, as Frege also emphasizes elsewhere (CP, 149n). An expression of judgment is that which we write down or utter *in* judging; In the paradigmatic case, to write down or utter an expression of judgment *is* to judge. To be clear: none of this is meant to deny that there are complex philosophical issues to be addressed here, which I have unfortunately not been able to adequately investigate within the scope of this dissertation. I am grateful to professor Maria van der Schaar for pressing me on the need for some clarification on how I mean to be using the notion of an expression of judgment.

(names) together to arrive at the proposition. This notion of interpretation, then, constitutes a prime example of the sort of 'asking for the meaning of a word in isolation' that Frege's context-principle is meant to rule out. On Frege's judgment-first approach, we arrive at the name by analysing the proposition, in such a way that its being used in the expression of judgment is internal to the name. Adapting the context principle, we could say that we are not to look for the name in isolation, but only in the context of the expression of judgment.

4.5 Full-blooded occurrences and privative occurrences of the sign

If the act of judgment is logically fundamental, this should be reflected in a logically proper conception of the sign. In the same way that *Bedeutung* comes – through the context principle – to be wedded to judgment, so the sign should be wedded to judgment. We saw how both the syntactic and semantic conception fail to do this. If the act of judgment is logically fundamental, the logically fundamental notion of the sign is that of a sign used in the expression of judgment, and this must be taken as the basis for our reflections on the sign. As was already said above: what a sign does – to signify – *is* to be used in the expression of judgment.

That a sign's being used in the expression of judgment constitutes the logically fundamental case, does not mean, however, that there are no other cases to which the notion of a sign also has application. One notorious such case is that of Frege's *Grundgesetze*. In formulating his *Basic Law V*[31], Frege seems to be using signs to express a judgment. As it turns out, however, he has not succeeded in expressing a judgment at all. Nevertheless, such a case must still be coherently describable as a case in which Frege has used *signs* rather than mere figures. A failure to express judgment cannot be sufficient to label a case as mere formalist babble.

[31] In contemporary notation: $(\Phi)(\Psi)\left(\left(\grave{\varepsilon}\Phi(\varepsilon) = \grave{\alpha}\Psi(\alpha)\right) = \left((x)\left(\Phi(x) = \Psi(x)\right)\right)\right)$.

It is exactly such considerations, I claim, that motivate Frege's distinction between signs that signify and signs that merely have the purpose of signifying without actually signifying. What we have in the latter sort of case, are signs that occur in a way that falls short of being used in the expression of judgment. Their mode of occurrence can be said to constitute a privation in relation to the fundamental case of being used in the expression of judgment. I will call such a case a *privative occurrence of the sign*, or a *privative case* for short. When a sign actually signifies, I call this a *full-blooded occurrence of the sign*, or a *full-blooded case*.[32] Crucially, both sort of cases must be construed as occurrences of *signs*.

To understand Frege's conception of the sign, it is crucial to get correctly in view the logical order of priority between the full-blooded and the privative case. In accordance with Frege's judgment-first approach, it is the full-blooded case – the use of signs in the expression of judgment – that constitutes the logically fundamental case. That Frege characterizes the privative case in terms of 'having the purpose of signifying', reveals that the privative case is to be conceived as logically dependent on the case in which a sign actually signifies, on the full-blooded case.[33]

There are many ways in which the occurrence of a sign may constitute a privative case, many ways in which its purpose to signify may end up not being realized. We will consider some examples shortly. It is important to emphasize, however, that such cases must always be seen as logically dependent on the full-blooded case. Indeed, this is what gives the notion of the privative case its unity. Any conception of the sign on which the privative case – or a certain kind of privative case – comes to be understood as logically prior to the full-blooded case fails to properly account for the way in which Frege seeks to align his conception of the sign with his judgment-first approach to philosophical logic, an alignment that gives rise to the internality of its use to the sign.

[32] My use of the term 'full-blooded' is inspired by the ways in which it is also used in (Conant, 2020).

[33] Compare: "The sign-in-use is the logically prior notion. Cases of the employment of a sign that are characterizable in such a way that we are entitled to speak of the occurrence as involving a 'mere sign' are always ones that indicate something attenuated or defective in the employment of the sign. A sign in its logically full-blooded employment is a sign-in-use" (Conant, 2020, 891n). It is important to realize, however, that Conant's notion of a sign-in-use need not be the same as Frege's notion of a sign that is used in the expression of judgment. That is to say: a shared commitment to the logical priority of the sign-in-use may nevertheless come with different conceptions of the relevant notion of use.

4.6 *Sinn* and *Bedeutung*

Let us now consider in some more detail what I take to be a crucial sort of privative occurrence of the sign, which will also allow me to make good on my promise that we would come to see how the *Sinn*/*Bedeutung* distinction is involved in Frege's conception of the sign. Consider the following passage:

> A proper name must at least have a sense [*Sinn*] (as I use the word); otherwise it would be an empty sequence of sounds and it would be wrong to call it a name. But if it is to have a use in science we must require that it have a meaning [*Bedeutung*] too, that it designates or names an object (PW, 124; NS, 135).

Frege indicates that he wishes to limit the notion of a name to what at least has *Sinn*, although it need not have *Bedeutung*.[34] What I wish to argue, is that the case of such a name with *Sinn* but no *Bedeutung* gives us one important case of privative occurrences of signs, in this case specifically of names. I will call such occurrences *lapsed occurrences of names* or *lapsed cases*.

In the full-blooded case, a name is used with *Bedeutung* in the expression of judgment. In some cases, however, a name may fall short of having *Bedeutung*, in such a way that we have a lapsed case.[35] Still, as argued above, a lapsed occurrence remains the occurrence of a *sign*, not of a mere formalist figure. What distinguishes such a case from that of a mere figure, is that the name must still have *Sinn*. To have *Sinn* but no *Bedeutung*, then, is one way for a name to *merely* have the purpose of signifying, i.e. to occur without successfully signifying. Frege presents examples of such cases in *Über Sinn und Bedeutung*:

> It may perhaps be granted that every grammatically well-formed expression figuring as a proper name always has a sense. But this is not to say that to the sense there also corresponds a thing meant. The words 'the celestial body most distant from the Earth' have a sense, but it is very doubtful if there is also a thing they mean. The expression 'the least rapidly convergent series' has a sense but demonstrably there is nothing it means, since for every given convergent series,

[34] I am assuming that Frege's remark applies to names in general, not just proper names.

[35] My formulation is meant to bring out that I am not committing myself to the claim that all privative occurrences of names are lapsed cases in my sense. Whereas the notion of the lapsed case is internally related to that of *Sinn*, it remains an open question whether this obtains for the notion of a privative occurrence of a name as such. If it does not, Frege's formulation above cannot be upheld in its full generality.

> another convergent, but less rapidly converging, series can be found. In grasping a sense, one is not certainly assured of meaning anything (CP, 159; KS, 145; 28).

In claiming that cases of names with *Sinn* but no *Bedeutung* constitute examples of privative occurrences of signs, I seem to be committing myself to a reading of Frege according to which there can be such a thing as *Sinn* without *Bedeutung*. I do not, however, wish to incur such a commitment upon myself. As is well known, this is a controversial matter, and readers such as Evans and McDowell have argued that what we have in such cases is only an *illusion* of *Sinn* (Evans, 1982, 22ff.) (McDowell, 1988, 212-213). They base their account on passages where Frege characterizes proper names without *Bedeutung* as 'mock proper names' and thoughts without *Bedeutung* as 'mock thoughts', writing things like:

> The logician does not have to bother with mock thoughts, just as a physicist, who sets out to investigate thunder, will not pay any attention to stage-thunder. When we speak of thoughts in what follows we mean thoughts proper, thoughts that are either true or false (PW, 130; NS, 142).

The analogy with stage-thunder does suggest that the idea of *Sinn* without *Bedeutung* gives us some sort of simulacrum of *Sinn* at best, and not *Sinn* proper. In my terms, the question is this: must we construe a lapsed occurrence of a name as an occurrence where it has *Sinn* but no *Bedeutung*, or where there is merely an illusion of *Sinn*? There are considerations to be advanced in favour of both options, and this is one of the many issues which I am, unfortunately, unable to fully explore in this dissertation. More generally, I cannot fully explore how Frege's *Sinn*/*Bedeutung* distinction must be integrated into my reading of Frege's conception of the logical categories. A thoroughgoing exploration of that question, I suspect, would at least double the length of this dissertation.

Let me, nevertheless, say a little bit more. Frege characterizes *Sinn* as a mode of presentation of *Bedeutung*, or as a mode of designating (CP, 157-158). This makes clear that the notion of *Bedeutung* is logically prior to that of *Sinn*, and that the logically fundamental case of *Sinn* is that of successfully presenting a *Bedeutung*. This is in accordance with the logical priority of the full-blooded case of a name with both *Sinn* and *Bedeutung*, as I have

emphasized it.[36] In the full-blooded case, to have *Sinn* just is a way to have a *Bedeutung*, and thereby just is a way of signifying. We can now reformulate the question above as follows: Is the lapsed case a case where there is what we could call a failed presentation of *Bedeutung* which is nevertheless a genuine case of *Sinn*, or should we say that there is only the illusion of such a presentation, and thereby only an illusion of *Sinn*. Put differently: is Frege's notion of a mode of presentation of a *Bedeutung* such that anything that genuinely counts as such a mode of presentation must *ipso facto* yield a *Bedeutung*, or not? As before, however, I can only formulate this question, not answer it.

What I do wish to commit myself to in this dissertation, is the claim that the lapsed case – however it is construed – is a genuine case of a *sign*, and must thereby be distinguished from that of a mere figure. Even if we are to construe lapsed occurrences of names as involving us in mere illusions of *Sinn*, this notion of an illusion of *Sinn* must still allow us to distinguish between signs that have the purpose of signifying, on the one hand, and mere figures, on the other hand, so that it must turn out that the latter cannot even be properly said to present an *illusion* of *Sinn*. If one believes that no notion of a mere illusion can deliver on this score, then this constitutes an important objection to the sort of readings advanced by Evans and McDowell. Whatever it is we are to say exactly about cases such as *Basic Law V*, we must be able to make sense of the fact that such cases involve genuine *signs* and not just mere figures.[37]

[36] In some cases – such as the case of storytelling, which is often invoked by Frege as an example – the fact that we have lapsed cases may be essential to the way in which names are employed, so that its falling short of the purpose of *bedeuten is* its purpose in this case. Such cases require separate treatment, but they remain parasitic on the full-blooded case.

[37] Compare: "If [Frege] had rejected the possibility of sense without reference, it would have followed that in [*Grundgesetze*] he had neither expressed any thoughts nor carried out any reasoning. Clearly, he may have found this evaluation of his life's work unattractive" (Kremer, 2010, 259n). If Kremer is right, we need a notion of genuine *Sinn* without *Bedeutung*. What Frege would *certainly* have found – not only unattractive – but deeply confused, is the idea that in *Grundgesetze* he had not even written down *signs*, so that the failure of his project has the result that there is no distinguishing Frege from the formalist.

4.7 Elucidatory use

There is another sort of occurrence of signs that deserves special discussion: that of the occurrence of the logical category terms in his elucidations.[38] Here, the notion of *Sinn* is certainly inappropriate, since it belongs to the very essence of Frege's elucidations that the notion of *Sinn* does not even *seem* to have application to such cases, since elucidations – properly understood – do not even *resemble* judgment. As soon as it is clear that what is at issue is a logico-philosophical elucidation, any seeming connection to judgment is obliterated.[39] At least, that is what follows from the sort of account I have presented, an account that I take – at least in its broad outline – to be true to Frege's considered conception of the logical categories. Nevertheless, we run into trouble at this point, as I now wish to explain.

In saying that Frege's elucidations involve a logically stratified elucidatory use of the logical category terms, I have of course been meaning to describe Frege as using genuine *signs*. It would be as disastrous to think that Frege's elucidations constitute mere formalist babble as it would be to conclude that Frege's arithmetical judgments themselves constitute such babble. Both are essential to his philosophical logic, and neither are on a par with what the formalist does. There arise, however, serious issues about how to understand Frege's elucidatory use of signs. The problem is that Frege's notion of the sign is such that the nature of the sign is fundamentally wedded to its use in the expression of judgment, which constitutes the one standard against which to understand all privative occurrences of signs *as* privative.

In this way, the occurrences of the logical category terms in Frege's elucidations would have to be understood as privative because they fall short of the full-blooded case in which such signs are used in the expression of judgment. But this gives us a conception of Frege's elucidatory use of the logical category terms as falling short of *judgment*, whereas I have been at pains to argue that we should *not* understand elucidations as failed attempts to express judgments. That elucidating the logical categories is a *sui generis* activity that is not to be modelled on judgment, means that Frege's elucidatory use of the

[38] I will focus on the logical category terms, but what I say can be extended to other occurrences of signs in Frege's elucidations.
[39] That a naïve ineffabilist may confusedly take Frege's elucidations as attempts at judgment, is as irrelevant as the fact that a formalist may confusedly take formalist figures as signs.

logical category terms is a *sui generis* use that is not to be modelled on the use of signs in the expression of judgment. In focusing on the use of signs in the expression of judgment in the way that he does, Frege seems to leave no room for a proper conception of his *elucidatory use* of the logical category terms.

To overcome this problem, the internality of its use to the sign has to be conceived in a way that accounts both for the use of signs in the expression of judgment *and* the elucidatory use of signs in elucidations, where neither of them comes to be seen as merely privative with regards to the other. This, however, introduces a deep bifurcation into Frege's conception of the sign that is entirely absent from Frege's own discussions and which it is unclear how to deal with in Frege's own terms. Here, then, I think we run into a crucial limitation of Frege's philosophical logic, a limitation which it is unclear how to overcome in Frege's own terms, and which I can only flag here as a topic for further research.[40]

4.8 The internality of its use to the sign

To do justice to Frege's judgment-first approach to philosophical logic, one must take the full-blooded case as the logically fundamental case, and the privative case as the logically derivative case. It is important to realize, however, that not any characterization of the sign in terms of a notion of use succeeds in doing this. An instructive example is found, I think, in Michael Kremer, who has presented one of the rare explicit discussions of Frege's conception of the sign. Quoting several passages in which the internality of its use to the sign is at issue, Kremer concludes with the following statement:

[40] That overcoming this limitation requires a far-reaching transformation of Frege's conception of the sign and thereby of his philosophical logic is, I would argue, a conclusion that underlies Wittgenstein's reflections in the *Tractatus* that resulted in his sign/symbol distinction (TLP, 3.3ff.). As I understand this distinction, it has no genuine ancestor in Frege, but is instead meant to relieve Frege's conception of the sign of a crippling ambiguity *between* sign and symbol. In a nutshell: To use a sign in the expression of judgment is to use it as a symbol, but its being used *as a symbol* is not internal to the sign in the way Frege thought, precisely because signs can also be used in other ways that are not mere privations of their use as symbols. I hope to explore these matters in future work.

> We can say that the same figure can be used in different ways, yielding different signs; we have the *same* sign if and only if we have the same figure, put to the same use (Kremer, 2010, 255).

It may be thought obvious that this account does justice to the internality of its use to the sign. Yet, I wish to argue that on a properly Fregean understanding of what that internality amounts to, it does not. The problem lies in Kremer's invocation of the notion of the figure. It is clear from the context that Kremer uses the term 'figure' to coincide with Frege's notion of a formalist figure. In other words: figures are constituted logically independently of their being used in the expression of judgment. Because of this – although Kremer does not invoke anything like the semantic notion of interpretation – his notion of use nevertheless inherits its crucial feature: use is taken to be externally imposed on what is constituted independently of it, i.e. the figure. The result is that Kremer's full-blooded case – the case of a figure in use – is conceived as logically posterior to the notion of the figure. This is further confirmed in the following passage:

> Adapting a suggestion of Wilfrid Sellars [...] to this context, let us use 'asterisk quotes' to form names of the 'figures' enclosed in them; and let us reserve ordinary quotation for naming 'signs', figures-in-use. So, for example, *and* occurs twice in *Sand and water make mud*, while 'and' occurs only once in 'Sand and water make mud' (Kremer, 2010, 255).

The primary mode of occurrence of the figure is the way in which *and* is said to occur twice in *Sand and water make mud*, a mode of occurrence that is logically independent of any use to which the figure may be put in the expression of judgment. The figure also occurs in 'Sand and water make mud', and there occurs as *used*, but such occurrence is logically dependent on its occurring in the figure *Sand and water make mud*. Indeed, the figure *and* occurs twice in 'Sand and water make mud' as well, once as used in the expression of judgment, and once as not so used. As far as the *figure* is concerned, both occurrences are logically on a par. It is only when we bring in considerations of use – considerations that are external to the nature of the figure – that these occurrences can be distinguished.

At first sight, Kremer's account seems to include the following claim: its being used lies in the nature of the sign. Yet, I have argued that Kremer's account is not faithful to Frege's conception of the internality of its use to the sign. How can this be? In fact, statements such as 'its being used lies in the nature of the sign' can be understood in different ways,

and Kremer's understanding is not Frege's. Kremer's conception of the sign essentially involves the notion of the figure, which is conceived as being constituted logically independently of its use in the expression of judgment. A figure *may* be used, but it need not be used, and its being used is external to its nature as a figure. For Frege, however, a logically proper conception of the sign does not involve such a notion of the figure at all. As we saw, Frege's critique of formalism is not merely that formalists do not interpret or use their figures. It cuts much deeper. Frege's critique is that the very notion of a figure is a logical dead end. The point is not that something must be superadded (interpretation or use) to a figure in order to arrive at signs proper.[41] The point is that we can never arrive at a logically proper conception of the sign to begin with if we take our departure from the formalist notion of a figure.

Here is what is fundamental to Frege's conception of use: the notion of use must not be conceived as having application to what is constituted logically independently of its being used in the expression of judgment. Rather, that to which the notion of use has application must be conceived as itself internally related to its use, as being such that its being used in the expression of judgment lies in its very nature. It is a conception that equally breaks with the syntactic conception, the semantic conception, and Kremer's conception of the sign, all of which are committed to the following principle:

> (EXTERNALITY) That which is used in the expression of judgment – that to which the notion of use has application – is constituted logically independently of its being so used.

It is (EXTERNALITY) that constitutes the philosophically fateful commitment underlying the contemporary syntax/semantics distinction that is incompatible with Frege's conception of the sign. To properly account for the internality of its use to the sign, is to leave behind (EXTERNALITY), and thereby to leave behind the contemporary syntax/semantics distinction. On Frege's conception, an understanding of the nature of what is used in the expression of judgment, *is* an understanding of how it is used in the expression of judgment, since its being so used is constitutive of its nature. In understanding how *Begriffsschrift* signs are used to express judgments, we need no longer look beyond the nature of what is used in the expression of judgment itself – we need no

[41] Compare what Bronzo calls the e*xtra feature account* (Bronzo, 2017, 1346) and what (Conant, 2020b) calls the *additive conception of language*.

longer look beyond the *Begriffsschrift* – since its being used to express judgments lies in its very nature. Therein lies the crucial difference with the syntactic, the semantic, and the Kremerian conception. Whichever way we view things on those conceptions, we will always run up against a form of externality between what is used – the syntactic unit, the figure, etc – and its use. What is decisive, then, is not just whether signs are conceived in a way that essentially involves a notion of use, but also whether that use is essentially a use *of the sign*, and not primarily a use of something else, such as an uninterpreted syntactic unit, or a figure. Any conception of the full-blooded case that involves an order of what is used that is taken to be constituted logically independently of its being used fails to properly account for the internality of its use to the sign.

It could be objected that this is all good for the full-blooded case, but that it raises problems for an understanding of the privative case. After all, if a sign can occur without in fact being used in the expression of judgment, then in what sense can it still be said that its being so used lies in the nature of the sign?[42] The answer is that it is precisely for this reason that the full-blooded case must be conceived as logically prior to the privative case, so that the latter indeed constitutes a *privation*. The point can be put like this: that its being used lies in the nature of the sign, does not mean that the sign realizes its nature in all of its occurrences. That a sign fails to realize its nature of being used in the expression of judgment in a privative occurrence, does not make it any less true that its being so used belongs to its nature. A proper conception of the privative case as logically posterior to the full-blooded case, is as a case in which the sign fails to realize its nature of being used in the expression of judgment, a nature that it does realize in the full-blooded case.

There is what could be called a metaphilosophical point in the offing here. The puzzle above can be formulated as follows: how can that which is present in all occurrences of the sign (the purpose of signifying) be logically posterior to that which is present only in some occurrences (actually signifying)? The answer is: because that which is present only

[42] This is the train of thought, I think, that underlies remarks like the following: "For Frege, the sentences of the *Begriffsschrift* have a meaning right from the start. It makes no sense to speak of a sentence without meaning" (van der Schaar, 2018, 227). Although van der Schaar correctly sees that, for Frege, it belongs to the nature of the sentence to have meaning, she takes this to entail that there can be no such thing as a sentence without meaning, so that there is no such thing as a privative occurrence of a sentence. In this way, Frege is saddled with a position that is too crude to be philosophically viable. The key is to see how the idea that it belongs to the nature of the sentence to have meaning *and* the idea that there is such a thing as a privative occurrence of a sentence can be held together.

in some occurrences is constitutive of the nature of the sign, and its absence in the other cases must be understood as a privation.[43] The metaphilosophical point is the following: The question of which of two notions is logically prior should not be confused with the question of which of them has the larger extension.[44]

The reflex – as we may call it – to think about the sign in ways that involve a version of (EXTERNALITY) is certainly not limited to readers of Frege, but is ubiquitous throughout analytical philosophy. Here are some characteristic passages:

> The question of propositional identity comes to admit of one or another definite answer in terms of geometrical similarity or conventional correspondence of written marks (Quine, 1934, 476).

> Propositional signs are merely 'inscriptions'; only in relation to *my will* do they constitute symbols [original emphasis] (Hacker, 1986, 100).[45]

> Well, a word is just an inscription, a mark on paper. Something has got to be done to it by its user for it to get a meaning. That much is clear (Boghossian 2008, 488).

Consider another passage from Stanley:

> In order to understand semantical results, one must have recourse to a metaperspective from which one can speak of relations between a formal language and its interpretation. But on Ricketts' view, Frege had no perspective from which he could separate a formal language from its interpretation. Hence, metatheoretical questions such as soundness or completeness would make no sense for Frege (Stanley, 1996, 47).

It is exactly this idea of separating a formal language from its interpretation that I have argued to be entirely foreign to Frege's philosophical logic, which therefore precludes the

[43] Such a way of thinking has, of course, a venerable line of ancestry in philosophy, going back at least to Aristotle.

[44] What I have said here is not unrelated to what could be called *Conant's maxim*, as it features throughout (Conant, 2020). One possible formulation runs as follows: Never to take what constitutes an attenuated exercise of a logical capacity to constitute a logically essential, self-standing moment in the full-blooded exercise of that same capacity. Note that I do not mean to commit myself to the claim that we find in Frege what could be called a full-blooded adherence to Conant's maxim. Conant's own discussions of Frege aim to show that we find in Frege a privative adherence to Conant's maxim at best.

[45] Hacker is here discussing the *Tractatus*. From what has been said, it is already clear that I consider such a reading of the *Tractatus* to be equally mistaken.

sort of metalogical investigations that Stanley wishes to say were perfectly admissible to Frege.

Once again, however, we must be careful to understand this point correctly. Ricketts argues for the negative conclusion that "[Frege's] conception of judgment precludes any serious metalogical perspective and hence anything properly labelled a semantic theory" (Ricketts, 1986, 76). He summarizes his discussion by saying that "Frege has no nonsyntactical metalogical vocabulary" (Ricketts, 1986, 83). If my interpretation is on the right track, however, this is at best a most unfortunate way of putting the matter. It makes it sound as if Frege shares and accepts a notion of metalogical vocabulary, only to reject the nonsyntactical variety of it, leaving him with a remainder of merely syntactical metalogical vocabulary. There *is* a metalogical perspective in Frege, it seems, only it is not 'serious'. On my reading, to properly grasp Frege's conception of the sign is to grasp how Frege must regard the very idea of such metalogical vocabulary – semantic *or* syntactic – as inherently confused, relying as it does on (EXTERNALITY). It is true, of course, that the contemporary framework makes it appear as if everyone must accept *some* metalogical vocabulary – on pains of not being able to talk about signs at all – and perhaps this is the impulse that is driving Ricketts to formulate his point in this way. But to view things in this way is to have already taken the fateful step of situating Frege's philosophical logic within that contemporary framework, which is precisely what one should not do.

4.9 The compositionality of the *Begriffsschrift*

The main task of this chapter was to come to an understanding of the logical stratification of the *Begriffsschrift*. Having attained an initial understanding of Frege's conception of the sign, we are now ready to take up this task. What we need to come to understand, is the nature of the distinction between the different categories of *Begriffsschrift* names, such as proper names, first-level function names with one argument place, second-level function names with one argument place for a first-level function name with one argument place,

and so on. There is a crucial passage in the *Grundgesetze* in which Frege discusses these distinctions:[46]

> We recognize in these examples the great multiplicity of functions. We can also see that there are fundamentally different functions, since the argument places are fundamentally different. In particular, those which are suited to take proper names cannot receive names of functions, and conversely. Moreover, argument places that can take names of first-level functions with one argument are unable to take names of first-level functions with two arguments. [...] We distinguish:
>
> > *argument places of the first kind*, that are suitable to take proper names;
> > *argument places of the second kind*, that are suitable to take names of first-level functions with one argument;
> > *argument places of the third kind*, that are suitable to take names of first-level functions with two arguments.
>
> Proper names and object-letters *fit* [*passen*] the argument places of the first kind; names of first-level functions with one argument *fit* the argument places of the second kind; names of first-level functions with two arguments *fit* argument places of the third kind (GGA, I, §23).

Frege invokes a notion of *fit* to elucidate the logical stratification of the *Begriffsschrift*. When Frege says, for instance, that proper names *fit* the argument place of first-level function names, he is characterizing the nature of proper names and first-level function names. It is *in* being fitted together in the expression of judgment that proper names and first-level function names realize their nature as signs. Fitting together *is* how they are used in the expression of judgment.

If we have a first-level function name that is a concept, then a proper name fitting the argument place of that function name yields a first-level subsumption sign. Thus, the first-level subsumption sign gives us a nexus wherein proper name and first-level function name can realize their nature. Of course, the first-level subsumption sign is itself a sign the nature of which is to be used in the expression of judgment. The simplest case of such use is that where the first-level subsumption sign is self-standingly used to

[46] For now, I am concerned with the logical stratification of the *Begriffsschrift*, so I have left out Frege's discussions of the realm of *Bedeutungen* in this passage. In the next chapter, we will return to this passage and discuss it in full.

express a judgment.[47] This yields a *Begriffsschrift* judgment that ramifies into a full-blooded occurrence of a proper name and a full-blooded occurrence of a first-level function name that fit together. In so forming this nexus of full-blooded occurrences, both these signs realize their nature of being used in the expression of judgment.

In line with Frege's judgment-first approach, it is in the expression of judgment that we have the logically fundamental case of names fitting together. Cases such as complex proper names give us logically derivative cases of fit. The expression of a judgment always involves the full-blooded occurrence of a sign that ramifies into full-blooded occurrences of names fitting together. The logical stratification of the *Begriffsschrift*, then, consists in these mutual relations of fit that govern the ways in which *Begriffsschrift* judgments ramify into full-blooded occurrences of names that fit together. This gives us what I will call the *compositionality* of the *Begriffsschrift*, which consists in nothing else than these nexuses of fit wherein *Begriffsschrift* signs realize their nature of being used in the expression of judgment. Its compositionality, then, *is* the logical stratification of the *Begriffsschrift*.

The compositionality of the *Begriffsschrift* does not consist in a division of *Begriffsschrift* signs into self-standing categories. Rather, it reveals how all *Begriffsschrift* signs are essentially interconnected through their fitting together in the *Begriffsschrift* judgments. Proper names fit the argument places of first-level function names, but first-level function names also fit the argument places of second-level function names, and so on. In this way, the compositionality of the *Begriffsschrift* gives the *Begriffsschrift* its unity. Its compositionality *is* the logico-philosophical form of unity of the *Begriffsschrift*, which, as said, is nothing else than its logical stratification. The categories of proper name, first-level function name, second-level function name, etc, are not constituted logically independently of each other. None is logically prior to the other. Rather, they are essentially intertwined through their mutual relations of fit. To think that one can understand one nexus of fit – for instance, the way in which a proper name fits a first-level function name – independently of an understanding of any other, would constitute a complete misunderstanding of Frege's philosophical logic. One cannot, for instance,

[47] Once again, I abstract from the role that Frege ascribes to the judgment-stroke and the horizontal in the expression of judgment. Taking these into account would further complicate matters in ways that I have not been able to investigate within the scope of my work on this dissertation, and therefore am not yet in a position to address, although I hope to do so in future work. I do not claim that such an investigation would not substantively impact my own account.

understand how 'Fa' involves a proper name fitting a first-level function name without *ipso facto* understanding how '$(\forall x)(Fx)$' involves the same first-level function name fitting a second-level function name. To think that one can, gives rise to a sort of misunderstanding that is akin to the sort of misunderstanding – discussed in Chapter 2 – according to which the unsaturatedness of first-level functions, correctly conceived, has no logical bearing at all on the unsaturatedness of second-level functions. This parallel will be unfolded in more detail in the next chapter.

This also reveals that the completeness of proper names should not be taken as indicating that proper names are constituted independently of their occurrence in the expression of judgment. The contrast between complete and incomplete names is not a contrast between that which requires a connection with other signs to fulfil its logical function, and that which does not, or something similar. *All* names are *essentially* such as to fit together in the expression of judgment, and that is how they realize their nature. Rather, the contrast is between names that do and names that do not have argument places.

To further our understanding of these matters, it is helpful to examine Black's investigation of Frege's claim that function names are unsaturated. Here is a selection of some things he says: [48]

> A function sign, unlike a designation (*Eigenname*) must have gaps—places intended to be filled by designations; it is logically impossible to refer to a function except by means of a sign having such gaps or hiatuses. [...] [Frege] is thinking of the function sign as a kind of container—something, as he often says, that can be 'filled up.' From this standpoint, the function sign is like a frame without a picture, a glove without a hand, or an empty mold; and the free variable is like a lay figure, a mere dummy used to draw attention to the way in which the gap might be filled. [...] Now is it really necessary for the function sign to include some indication of empty spaces that are intended to be filled up? We must notice, first, that the use of such a sign can be demanded only when the function sign is used separately, and is not attached to a designation to form a complex designation of an object. [...] It would be too farfetched to regard an expression like 'log 2' as containing a gap or space in addition to the signs 'log' and '2.' This would be like saying that a bottle of wine

[48] Although I will discuss Black's reading in quite some detail in this and the next chapter, there remains much more to be said about it. Black's paper constitutes a most wonderful example of the sort of reading of Frege to which I wish to present an alternative in this dissertation.

> consists of the empty bottle, the wine, and the space filled by the wine (Black, 1968, 233-234).

Black thinks that some function names are unsaturated – which means that they contain a gap – while others are not. In the sign '$\log(\,)$' there is a gap. In the sign '$\log \xi$', we can say that there is a gap that is filled by the Greek letter. Since the sign '$\log 2$' is composed of a function name and a proper name, we may be tempted, Black thinks, to say that the function name in '$\log 2$' is unsaturated as well, so that there must be a gap present in '$\log 2$' as well. But this is absurd, due to the absurdity of saying certain things about bottles of wine. The conclusion is that the function name that occurs in '$\log 2$' should not be taken to have a gap, and thereby should not be taken to be unsaturated at all. Hence, it is only where we have function names that are not combined with proper names, that the issue of the unsaturatedness of the function name comes up at all.

For Frege, the idea that the question of unsaturatedness only comes up outside contexts where function names fit together with proper names (or other signs) constitutes a momentous misunderstanding. Rather, it is *in* such contexts that the unsaturatedness of the function name is in its prime, so to speak, *because* this is where the proper name fits the function name. To say that a function name is incomplete, is to characterize its compositional nature as a sign, is to characterize how it fits together with other names in the expression of judgment. So understood, the idea that the sign '$\log \xi$' retains its unsaturatedness in '$\log 2$' is nothing else than the idea that it retains its nature of being such that a proper name fits its argument place – a nature which it realizes in '$\log 2$'[49] – an idea that has none of the strangeness that Black seeks to bring out. To go back to his analogy: in being filled with wine, the bottle does not thereby lose its capacity to contain wine. Rather, it realizes it.

Black's way of thinking about the issue stems from his operating in terms of (EXTERNALITY). He is thinking about signs as something like Fregean figures, so that he cannot but conceive of the unsaturatedness of the function name in merely geometrical terms, as signaling the necessary presence of a gap on the page, or something similar. As if replying directly to Black's himself, Frege writes the following in a discussion of the sign 'and' as it combines sentences:

[49] For ease of exposition, I am momentarily abstracting from the point that this is only a privative realization, since a complex name is not an expression of judgment. Taking this into account would not substantively impact my argument here.

> The 'and' [...] seems doubly unsaturated: to saturate it we require both a sentence preceding and another following. And what corresponds to 'and' in the realm of sense [*Sinn*] must also be doubly unsaturated: inasmuch as it is saturated by thoughts, it combines them together. As a mere thing, of course, the group of letters 'and' is no more unsaturated than any other thing. It may be called unsaturated in respect of its employment as a sign[50] [*Zeichen*] meant to express a sense, for here it can have the intended sense only when situated between two sentences: its purpose as a sign [*Zweck als Zeichen*] requires completion by a preceding and a succeeding sentence (CP, 393; KS, 381; 39).

This passage wonderfully brings out some of the central points of my discussion. To think that the unsaturatedness of the function name pertains to it as a 'mere thing' in the way Black does, is confused. If we look for the unsaturatedness of the function name on the level of the mere figure, we will never find the properly logical notion of unsaturatedness that Frege has in mind. What is unsaturated, is not a figure, but a *sign* that is used in the expression of judgment, and its unsaturatedness consists in the way in which, in being so used, it fits together with other signs. Note how Frege again uses the notion of purpose (*Zweck*) to characterize the nature of the sign: the sign fulfils its purpose as a sign – realizes its compositional nature – in fitting together with other signs in the expression of judgment.[51]

From Black's vantage point, it must become a complete mystery why Frege insisted that a function name *must* be unsaturated. He writes:

> That a symbol is a function sign can be shown in all sorts of ways—by the use of special type, numerical subscripts, and so on; and the same is true of information about the number of variables involved. [...] To use the sign 'log' correctly, we must of course know its 'logical grammar,' that is to say, the rules governing its correct use. But the demand that this grammar shall always be explicitly symbolized in the case of every sign used obviously leads to an infinite regress (Black, 1968, 236).

The regress Black has in mind runs, I think, along the following lines: if we symbolize the logical grammar of 'log' by using a Greek letter to indicate an argument place – as in '$\log \xi$'

[50] I amend the translation, which has 'symbol'.
[51] Frege's formulation in German is worth quoting: "Sein Zweck als Zeichen verlangt eine Ergänzung durch einen vorhergehenden und einen nachfolgenden Satz" (KS, 381). To try to make sense of such statements in merely geometrical terms, is to miss completely what is at stake in Frege's conception of the sign.

– then we should also indicate the logical grammar of that Greek letter itself, and the logical grammar of whichever signs we use in doing that, and so on. On Black's understanding, we can have a function name 'log' that is not unsaturated, and we can use a Greek letter to *add* an argument place to it and thereby make it unsaturated, as in '$\log \xi$'. We have the logical grammar of the function name, on the one hand, and the question whether it is unsaturated, on the other hand.

It may now be tempting to put the point as follows: Black has not realized that the notion of unsaturatedness is taken by Frege to characterize logical grammar of the function name rather than its geometrical shape. This, however, again involves the sort of externality that I have argued to be incompatible with Frege's conception of the sign. There is no such 'rather than'. This dichotomy between logical grammar and geometrical shape rests on a conception of the sign as something like Kremer's figure in use, so that the 'logical grammar' of the figure is taken to be external to its nature, which is conceived in merely geometrical terms. According to Frege, however, the notion of unsaturatedness characterizes the nature of what is used in the expression of judgment, where its being so used lies in the very nature of what is used. Black will hear 'the nature of what is used' as ambiguous between 'the geometrical nature of the mere figure' and 'the logical grammar that is externally imposed on the figure'.[52] To properly grasp Frege's conception of the sign, is to see that there is no such ambiguity, and that to characterize what is used in the expression of judgment as being unsaturated is *ipso facto* to characterize its use in the expression of judgment – the way it fits together with other names in the expression of judgment – since its being so used lies in its very nature. We do not *add* argument places to function names by writing down Greek letters. Rather than constituting argument places, Greek letters *occupy* them, as is clear from the following passages:

> Every function sign must always carry with it one or more places which are to be taken by argument signs; and these argument places – not the argument signs themselves – are a necessary component part of the function sign (PMC, 116; BW, 184).

[52] This comes out most wonderfully in passages such as: "What gives the series of sounds used in a genuine assertion its 'unity' is simply the fact that they are used by a speaker in accordance with the rules for correct speech of the language in question" (Black, 1968, 239).

> When in what follows an expression like 'the function $\Phi(\xi)$' is used, it is always to be borne in mind that 'ξ' contributes to the designation of the function only insofar as it marks its argument places, and that the nature of the function would be unchanged if any other sign were put for 'ζ' (GGA, I, §1).

> If we say 'the function $\Phi(\xi)$', then we should never forget that 'ξ' belongs with the function-name only by way of making the unsaturatedness recognizable. [...] The function-name is really just a part of '$\Phi(\xi)$' (GGA, I, §21).

> Consequently one can always speak of the name of a function as having empty places, since what fills them does not, strictly speaking, belong to them (PW, 119; NS, 129).

This serves to bring home the point that it is a complete misunderstanding of Frege to think that the unsaturatedness of function names lies in the order of the mere geometrical features of the figure, which can be added or removed at will. Whether the unsaturatedness of a function name is made recognizable by Greek letters or not, it is there all the same, and this is because its unsaturatedness consists in nothing else than the way in which it fits together with other signs in the expression of judgment, an aspect of the sign that is not geometrical, but *logical*. This also helps to better understand Frege's repeated insistence that function-names must have argument places. Here is a characteristic passage:

> Only very few mathematicians would think twice about expressing by '$f = g$' the circumstance that $f(\xi)$ always has the same value for the same argument as the function $g(\xi)$. However, this embodies a mistake which springs from a faulty conception of the nature of a function. An isolated function-letter without an argument-place is surely a complete aberration, just as an isolated function-sign like 'sin' is an aberration. For what characterizes a function, in contrast to an object, is just this unsaturatedness, that it requires completion by an argument, *and this must feature in the notation too* [my emphasis] (GGA, II, §147n).

Frege calls 'an isolated function-letter without an argument-place' a complete aberration. Black understands this along the following lines: 'We may only use as a function name figures that have a spatial gap or a similar feature'. As Black sees it, however, its having such a spatial gap or not is logically independent of the logical grammar of the function name, which is imposed externally on the merely geometric figure. For Frege, however,

to say that the unsaturatedness of (first-level) functions (of one argument) *must* feature in the notation is not to set up certain norms about what geometrical figures we may or may not proceed to use in certain ways. Rather, it is to say something about the compositional nature of the *sign*, i.e. that which has the purpose of signifying. That the unsaturatedness of first-level functions of one argument *must* feature in the notation, just *means* that such function names are intrinsically such as to fit together with proper names in the expression of judgment, something that *cannot* fail to come out in the expression of judgment, since it is constitutive of a first-level function name of one argument's occurring in an expression of judgment at all. The idea of a function name that is used in the expression of judgment without having argument places is a complete aberration because there is no such thing, and the idea that there is can only lead to logical confusions, such as a blurring of the distinction between function and object.

Here is how Frege continues the passage above:[53]

> The impermissibility of a notation like '$f = g$' emerges from the fact that in particular cases it immediately falters. If we take, for instance, $\xi^2 - 1$ for $f(\xi)$, and $(\xi - 1) \cdot (\xi + 1)$ for $g(\xi)$, then it is striking that nothing can be written to correspond to the equation '$f = g$' (GGA, II, §147n).

What does Frege mean when he says that nothing can be written to correspond to the equation '$f = g$', where 'f' and 'g' are construed as explained in the first passage above? It could be thought that one can straightforwardly write down '$\xi^2 - 1 = (\zeta - 1) \cdot (\zeta + 1)$' to correspond to '$f = g$'. The problem comes out if we examine the relation between '$f = g$' and an expression such as '$f(3) = g(3)$'. If '$\xi^2 - 1 = (\zeta - 1) \cdot (\zeta + 1)$' gives us a way of writing down '$f = g$', how are we to write down '$f(3) = g(3)$'? The answer seems obvious: we write down '$3^2 - 1 = (3 - 1) \cdot (3 + 1)$'. This does not work, however, because it depends on taking the Greek letters in '$\xi^2 - 1 = (\zeta - 1) \cdot (\zeta + 1)$' as occupying argument places, and this would mean that '$\xi^2 - 1 = (\zeta - 1) \cdot (\zeta + 1)$' is actually a way of writing down '$f(\xi) = g(\zeta)$', contrary to what was intended. In other words: the Greek letters in '$\xi^2 - 1 = (\zeta - 1) \cdot (\zeta + 1)$' cannot be construed as Greek letters in Frege's sense at all, since they do not occupy argument places. In writing down '$\xi^2 - 1 = (\zeta - 1) \cdot (\zeta + 1)$', we have not really written down anything that *corresponds*

[53] Compare also (PMC, 161-162).

to '$f = g$' in the proper sense, since the expressions '$\xi^2 - 1$' and '$(\zeta - 1) \cdot (\zeta + 1)$' as they appear in '$\xi^2 - 1 = (\zeta - 1) \cdot (\zeta + 1)$' have no logical connection to the *function names* '$\xi^2 - 1$' and '$(\xi - 1) \cdot (\xi + 1)$' as they appear in an expression such as '$(\forall x)\big(x^2 - 1 = (x - 1) \cdot (x + 1)\big)$'. If the latter is written as '$(\forall x)\big(f(x) = g(x)\big)$', then there is indeed nothing to write that corresponds to '$f = g$', since there is no way to render '$(\forall x)\big(x^2 - 1 = (x - 1) \cdot (x + 1)\big)$' without making the argument places recognizable. There is nothing that stands to '$f = g$' as '$(\forall x)\big(x^2 - 1 = (x - 1) \cdot (x + 1)\big)$' stands to '$(\forall x)\big(f(x) = g(x)\big)$', and this reveals that there is no such thing as using isolated function letters without argument places to express judgments about functions.

The upshot is that, insofar as '$f = g$' is used to express the same judgment as '$(\forall x)\big(f(x) = g(x)\big)$', it can only do so by going proxy for it, not by containing 'function names without argument places'. In a lecture on logic, for instance, it could be stipulated that – for ease of presentation – judgments of the form '$(\forall x)\big(f(x) = g(x)\big)$' will be written down '$f = g$'. Frege would not claim that *this* is somehow logically impossible, that we cannot properly make sense of such a use of '$f = g$'. The real point is that this use of '$f = g$' is entirely parasitic on its going proxy for '$(\forall x)\big(f(x) = g(x)\big)$'. Insofar as we describe '$f = g$' as itself expressing a judgment, we have here an attenuated sense of 'expressing a judgment' at best that is logically parasitic on the sense in which '$(\forall x)\big(f(x) = g(x)\big)$' expresses a judgment. Only in the latter do we have full-blooded occurrences of *Begriffsschrift* signs that realize their compositional nature, and this use of '$f = g$' is parasitic on such full-blooded cases. How we are to describe, exactly, the use of the signs 'f', '$=$', and 'g' in such a case, remains open. What is important is that – however we describe that use – the sense of 'use' at issue will remain a logically attenuated sense that is parasitic on the full-blooded occurrence of *Begriffsschrift* names that fit together in the expression of judgment.[54]

[54] In this way, there will be a whole swath of phenomena that may be described through all sorts of attenuated uses of terms such as 'sign', 'designate', 'judgment', 'use', and so forth. It is only on a case-by-case basis that such phenomena can be investigated, and such an investigation requires keeping in mind at all times the logical priority of the full-blooded case.

In the *Grundlagen*, Frege wrote:[55]

> It is possible, of course, to operate with figures mechanically, just as it is possible to speak like a parrot: but that hardly deserves the name of thought. It only becomes possible at all after the mathematical notation has, as a result of genuine thought, been so developed that it does the thinking for us, so to speak (GL, IV).

A notation that does the thinking for us, is a notation which is such that its mode of expression issues in full-blooded occurrences of signs through which they realize their compositional nature of fitting together in the expression of judgment in the way that *Begriffsschrift* signs do. In short, it is a notation that is, in the sense I have sought to articulate, logically stratified.

[55] This may be compared to a passage from the *Tractatus* which is best, I think, read in the original German: "Wir sagten, manches an den Symbolen, die wir gebrauchen, wäre willkürlich, manches nicht. In der Logik drückt nur dieses aus: Das heißt aber, in der Logik drücken nicht *wir* mit Hilfe der Zeichen aus, was wir wollen, sondern in der Logik sagt die Natur der naturnotwendigen Zeichen selbst aus" (TLP, 6.124). Here is the Ogden/Ramsey translation: "We said that in the symbols which we use something is arbitrary, something not. In logic only this expresses: but this means that in logic it is not *we* who express, by means of signs, what we want, but in logic the nature of the essentially necessary sign itself asserts" (TLP, 6.124).

Chapter 5
The internality of the *Begriffsschrift* and the realm of *Bedeutungen*

Frege's turn to signs involves the idea that our understanding of the logical stratification of the realm of *Bedeutungen* must go through an understanding of the corresponding logical stratification of the *Begriffsschrift*. In the previous chapter, I developed an account of Frege's conception of the logical stratification of the *Begriffsschrift*. Essential to that account was Frege's commitment to *the internality of its use to the sign*, so that its being used in the expression of judgment lies in the very nature of what is so used. We saw that the logical stratification of the *Begriffsschrift* consists in the manifold ways in which *Begriffsschrift* judgments ramify into full-blooded occurrences of *Begriffsschrift* names that fit together, giving rise to the compositionality of the *Begriffsschrift*. To fit together in *Begriffsschrift* judgments, *is* how *Begriffsschrift* names realize their nature, *is* how they are used in the expression of judgment.

The task awaiting us in this final chapter is to develop an account of the 'correspondence' that obtains between the respective logical stratifications of the *Begriffsschrift* and the realm of *Bedeutungen*. Such an account must provide us with a way of making sense of Frege's idea that an understanding of the latter 'goes through' an understanding of the former. That is to say: it must yield a logico-philosophical relation between functionality and compositionality that vindicates Frege's turn to signs as essential to the elucidation of the logical categories, rather than rendering it a merely optional gesture.

As can be expected, the internality of its use to the sign will be crucial to this correspondence. I will start by discussing two accounts that I will describe as *realist* and

idealist. We will see that, in failing to do justice to the internality of its use to the sign, they equally fail to give us an adequate conception of the logico-philosophical relation between the *Begriffsschrift* and the realm of *Bedeutungen*. In a way that will become apparent, a proper understanding of this relation consists in nothing more than a further unfolding of a proper understanding of the internality of its use to the sign. Because of this, my discussion in this chapter will manifest clear parallels to my discussion in the previous chapter, parallels that arise from this further unfolding and that are thereby themselves to be seen as revelatory of the nature of the relation in question.

The result is what I will describe as a form of internality of the *Begriffsschrift* and the realm of *Bedeutungen*. Rather than bringing in a matter that is external to the realm of *Bedeutungen* itself, Frege's turn to notation consists precisely in an unfolding of the internality that obtains between the realm of *Bedeutungen* and the *Begriffsschrift*. To develop an understanding of that internality, is the main task of this chapter. Frege's conception of the logical categories *is* his conception of that internality, and to elucidate the logical categories is nothing else than to elucidate that form of internality.

This internality, I will argue, consists in what I will call the *internal heterogeneity* of the logical stratification of the *Begriffsschrift* and logical stratification of the realm of *Bedeutungen*. I will criticize Geach's reading of Frege, according to which Frege's distinction between function names and proper names just is the distinction between functions and objects, so that function names *are* functions. This, I will show, brings us back to a form of naïve ineffabilism. What is needed instead, is a conception of the internality of the *Begriffsschrift* and the realm of *Bedeutungen* that does not render their respective logical stratifications identical, but rather heterogeneous. Indeed, we already have such a conception: whereas the logical stratification of the former consists in its compositionality, the logical stratification of the latter consists in its functionality. Although these are internally related, they are not the same, and it is only through their internal heterogeneity that each can give us the unity of their respective domains. There is, then, no self-standing logico-philosophical unity of either the *Begriffsschrift* or the realm of *Bedeutungen*. There is only the unity that they possess *in and through* their internal heterogeneity. I will end the chapter by discussing how this may be taken to give rise to a Fregean notion of *showing* on which the functionality of the realm of *Bedeutungen* shows itself in the compositionality of the *Begriffsschrift*.

In the introduction to the previous chapter, I noted its indebtedness to Section XIII of (Conant, 2020). Although this chapter is equally so indebted, I also wish to note its

indebtedness to (Diamond, 1984). The depth of Diamond's reading of Frege was, I think, unprecedented at the time, and I repeatedly went through the experience of further developing my own ideas and subsequently returning to Diamond's piece, only to come to the conclusion that her ideas seemed to have developed alongside mine. To adapt a quip of Putnam's that Jim Conant has shared with me: the smarter I got, the smarter Diamond seemed to get.[1] If one wanted, one could approach this chapter – and perhaps this dissertation as a whole – as my attempt to unfold what I take to be the main ideas underlying her marvellous paper.[2]

5.1 Frege's elucidatory notion of correspondence

My use of a notion of 'correspondence' to characterize the logico-philosophical relation that obtains between the logical stratification of the *Begriffsschrift* and the logical stratification of the realm of *Bedeutungen* – although it may be considered somewhat infelicitous, as we will see – is no whim, but taken from Frege himself. This notion of correspondence constitutes one salient way in which the issue of this relation comes up in Frege's own writings. We have already encountered the following passage from his correspondence with Russell:

> This difference between the signs must correspond [*entsprechen*][3] to a difference in the realm of meanings [*Gebiet der Bedeutungen*]; although it is not possible to speak of it without turning what is in need of completion into something complete and thus falsifying the real situation. We already do this when we speak of 'the meaning of 'is a square number''. Yet the words 'is a square number' are not meaningless. The analysis of the proposition corresponds [*entsprechen*] to an analysis of the thought, and this in turn to something in the realm of meanings, and I should like

[1] Putnam's own quip had 'Kant' instead of 'Diamond'.
[2] Another, more recent, precedent that deserves mention is (Johnston, 2017), whose reading of Frege I also take to exhibit crucial points of agreement with my account in this chapter.
[3] I follow the translators in translating '*entsprechen*' as 'to correspond to'. An alternative translation would be: 'to accord with'.

> to call this a primitive logical fact [*logische Urthatsache*]. This is precisely why no proper definition is possible here (PMC, 142; BW, 220).

Recall how there were two sorts of elucidatory statement at issue in Frege's discussion with Russell:

> (A) A function never takes the place of an object.
> (B) A function name never takes the place of a proper name.

When Russell complained about the incoherence of (A), Frege responded by presenting a discussion of (B). After that discussion, Frege continues with the passage above, wherein he turns his attention back to (A) – back to the realm of *Bedeutungen*. What Frege says is that the distinction between function names and proper names that he has just articulated must correspond to something in the realm of *Bedeutungen*. What he has in mind, of course, is the function/object distinction. In this way, Frege makes it clear that a proper understanding of the latter distinction – and thereby of (A) – must go through an understanding of the former distinction – an understanding of (B). Frege's notion of 'correspondence', then, is meant to capture precisely the sort of logico-philosophical relation between the *Begriffsschrift* and the realm of *Bedeutungen* that must obtain to vindicate his elucidatory turn to (B) as providing us with what is necessary for a proper understanding of (A).

Frege further motivates the way in which he approaches (A) through a discussion of (B) by reminding us that any purported attempt to directly state the function/object distinction must falter, because it turns what is in need of completion into something complete. This refers back to Frege's prior discussion with Russell. As was discussed in Chapter 3, that discussion involves the observation that a statement such as 'There is something that is not an object' can be construed in two ways – as a judgment or as an elucidation. Likewise, (A) can be construed both as a judgment and as an elucidation, and as long as (A) is construed in the former way – which is how Russell wants to construe it – it will end up talking about 'functions' as a first-level category of objects, thereby turning what is in need of completion into something complete. The conclusion is not that such a construal of (A) is incoherent, but rather that it has no bearing on the logical stratification of the realm of *Bedeutungen*, which is not a matter for judgment, but for elucidation.

Frege also briefly repeats the point – which he discussed in more detail in his first reply to Russell – that we must recognize function names such as 'ξ is a square number' as

having *Bedeutung*, given their indispensable role in the expression of judgment. At the same time, their *Bedeutung* cannot be an object, a point that Russell ultimately could not escape himself – as we saw in Chapter 3 – leading to his expression of puzzlement in (PoM, §52). Thus, we are forced to acknowledge the function/object distinction as a genuine distinction in the realm of *Bedeutungen*.

After having hinted back at some of his previous arguments in this way, Frege again invokes his notion of correspondence to bring home the fact that what is at issue here is a certain sort of relation between the *Begriffsschrift* and the realm of *Bedeutungen*. It will be noted that Frege invokes a double correspondence between the *Begriffsschrift* and the realm of thoughts, on the one hand, and between the realm of thoughts and the realm of *Bedeutungen*, on the other hand. This could be taken to suggest that the correspondence between the *Begriffsschrift* and the realm of *Bedeutungen* is logically parasitic on these two relations of correspondence. I believe, however, that this would be a mistake. As we saw in the previous chapter, the notion of *Sinn* must be understood as being logically dependent on the notion of *Bedeutung*. In its logically basic formulation, the purpose of a name is to have *Bedeutung*, not to have *Sinn*. Although names also have *Sinn* in the full-blooded case, their *Sinn* constitutes a mode of presentation of their *Bedeutung*, so that their having *Bedeutung* is the logically prior phenomenon. Even though it may be possible for a name to have *Sinn* without *Bedeutung* – depending, as we saw, on one's reading of Frege – it is the full-blooded occurrence of the name with *Bedeutung* that constitutes the logically prior phenomenon. For these reasons, the correspondence between the *Begriffsschrift* and the realm of *Bedeutungen* should equally be seen as the logically prior phenomenon. In any case, it is on this supposition that I will proceed. In this vein, it is important to note that Frege himself – as we will shortly see – often talks directly in terms of a correspondence between the *Begriffsschrift* and the realm of *Bedeutungen*, without invoking the realm of *Sinn* at all. To investigate how the intervening realm of *Sinn*, and *its* correspondence both to the *Begriffsschrift*[4] and the realm of *Bedeutungen*, are to be understood, is a task I cannot take up here. For one, it would require a thoroughgoing investigation of what sort of conception of *Sinn* is yielded by my reading. As already indicated, this is an investigation that I cannot take up here.[5]

[4] Here, the relevant notion is that of expressing: names *express* their *Sinn*, and sentences *express* thoughts (CP, 161, 354).
[5] Below, I will ever so briefly touch upon one possible shape that the integration of the realm of *Sinn* into my account may take.

In characterizing the *Begriffsschrift*-pole of the correspondence, Frege talks about 'the analysis of the proposition'. In accordance with my account from the previous chapter, we should understand this notion of analysis as pertaining to the compositionality of the *Begriffsschrift*, consisting in the manifold ways in which *Begriffsschrift* judgments ramify into full-blooded occurrences of names that fit together.

Frege ends by characterizing the correspondence in question as a logically primitive fact that cannot be provided with a definition. As we know, this is Frege's way of marking off a phenomenon as a topic for elucidation. Note that we have here a topic for *logico-philosophical elucidation*, not for scientific elucidation, in the sense of Chapter 1. The correspondence between the logical stratification of the *Begriffsschrift* and the logical stratification of the realm of *Bedeutungen* is – it should be clear by now – not a matter for judgment, but for elucidation.[6]

It is worthwhile to point out that Frege's elucidatory notion of correspondence does not only appear in the correspondence with Russell, but at several other places where Frege is elucidating the logical categories. Here is one salient example, which again occurs after Frege has discussed the distinction between complete and unsaturated signs:

> To this difference in the signs there of course corresponds [*entsprechen*] an analogous one in the realm of meanings [*Reich der Bedeutungen*]: to the proper name there corresponds the object; to the predicative part, something I call a concept. This is not supposed to be a definition; for the decomposition into a saturated and an unsaturated part must be considered a logically primitive phenomenon which must simply be accepted and cannot be reduced to something simpler (CP; 281; KS, 269; 371).

Again, Frege's presentation makes it clear that he believes that an understanding of the logical stratification of the realm of *Bedeutungen* must go through an understanding of the logical stratification of the *Begriffsschrift*. Here are three additional passages:

> The peculiarity of functional signs, which we here call 'unsaturatedness', naturally has something answering [*entsprechen*] to it in the functions themselves. They too may be called 'unsaturated', and in this way we mark them out as fundamentally different from numbers (CP, 292; KS, 279; 665).

[6] In Tractarian terms, the point can be put by saying that the logico-philosophical relation that obtains between language and the world cannot be pictured.

> This predicative component part of our sentence [i.e. the part 'is a prime' of the sentence 'Two is a prime'] which we have described in this way [i.e. as being unsaturated], is also meaningful. We call it a concept-word or *nomen appelativum*, even though it is not customary to include the copula in this. Just as it itself appears unsaturated, there is also something unsaturated in the realm of meanings [*Bereich der Bedeutungen*] corresponding [*entsprechen*] to it: we call this a concept (PW, 177; NS, 192).

> The sign 'sin' which we say designates the sine function, only occurs in mathematics in close combination with other signs, as in '$\sin 10^\circ$', '$\sin 0^\circ 11''$', '$\sin \alpha$', and so is a sign in need of supplementation and therein is different from a proper name. Its content [*Inhalt*] is correspondingly [*dementsprechend*] also in need of supplementation and therein is different from any object (PW, 271; NS, 290).

In each of these passages, Frege invokes his elucidatory notion of correspondence to bring his discussions of signs to bear on a proper understanding of the logical stratification of the realm of *Bedeutungen*.

In addition to such passages, there are many passages where Frege does not explicitly employ this notion of correspondence, but where the relation between the *Begriffsschrift* and the realm of *Bedeutungen* is equally at issue. Here is one example:

> Consequently one can always speak of the name of a function as having empty places, since what fills them does not, strictly speaking, belong to them. *Accordingly* [*demgemäss*] I call the function itself unsaturated, or in need of supplementation, because its name has first to be completed with the sign of an argument if we are to obtain a meaning that is complete in itself [my emphasis] (PW, 119; NS, 192).

Frege's use of the term 'accordingly' serves to bring out the requisite correspondence of the respective logical stratifications of the *Begriffsschrift* and the realm of *Bedeutungen*. Indeed, Frege goes so far as to say that it is *because* its name has to be completed, that the function itself is unsaturated.

The question, then, is how to understand this elucidatory notion of the correspondence of the respective logical stratifications of the *Begriffsschrift* and the realm of *Bedeutungen*. In a way that corresponds to my discussions of the syntactic and semantic conceptions of the sign from the previous chapter, I will first consider and reject what I will call a *realist* and an *idealist* conception of this correspondence. Seeing how they fail –

and how their failure is intertwined with their failure to do justice to the internality of its use to the sign – will help us to advance to what I consider a more adequate account.

5.2 Realism

Consider what good old Max has to say on the relation between the *Begriffsschrift* and the realm of *Bedeutungen*:

> Wittgenstein's conception of the philosophy of language (the search for its essence) required a stand on ontological issues: anybody who hopes to delineate a *Begriffsschrift* that adequately manifests the grain of reality must have at least some schematic view concerning the true structure of thought and its true, if hidden, connexion with reality; for how is one to distinguish the 'accidental' from the 'essential' features of language except in terms of *prior* notions about what 'reality' is like? [my emphasis] (Black, 1964, 7).

Although Black is talking about the *Tractatus*, it is clear that we can transpose his remarks to Frege. As Black sees it, the fundamental question is whether the *Begriffsschrift* adequately manifests the grain of reality. Put like this, there may still be a way of understanding this question that need not be unfaithful to Frege. Trouble begins, however, when Black adds that to answer this question, one must possess a prior notion of what reality is like. In our terms: one must possess a prior understanding of the logical stratification of the realm of *Bedeutungen*. Thus, the logical stratification of the realm of *Bedeutungen* is seen as logically prior to the logical stratification of the *Begriffsschrift*.[7]

In this way, we arrive at what may be called a *realist* construal of the correspondence of the *Begriffsschrift* and the realm of *Bedeutungen*: The realm of *Bedeutungen* is constituted logically independently of the *Begriffsschrift*, and the question is whether the latter

[7] Compare: "The logical (metaphysical) forms of states of affairs are language independent – *de re* possibilities do not depend upon our descriptions of them" (Hacker, 1999, 120). Many will be baffled by the very suggestion that such a statement could possibly be incorrect. The reason is that they are thinking of the operative notions of 'language' and 'description' in the sort of externalist terms yielded by the contemporary syntax/semantics distinction. It is in leaving behind the logico-philosophical framework of that distinction that one can come to see how it may not only be possible but also philosophically appropriate to reject a statement such as Hacker's.

succeeds in capturing the logically prior stratification of the former.[8] In setting up the *Begriffsschrift* in such a way that it adequately manifests the grain of reality, we bring to bear upon the *Begriffsschrift* a logically prior understanding of the realm of *Bedeutungen*. The result is similar to what we found in the previous chapter for the semantic conception of the sign: Frege's elucidatory turn to signs is rendered spurious because our understanding of the logical stratification of the *Begriffsschrift* turns out to be parasitic on a logically prior understanding of the realm of *Bedeutungen*.

Such a realist approach fails to do justice to the internality of its use to the sign. The very question whether the *Begriffsschrift* succeeds in capturing the logically prior stratification of the realm of *Bedeutungen* relies on there being an order *in* the *Begriffsschrift* that is constituted logically independently of its use in the expression of judgment, since it is only of such an order that this question makes sense. There must be a sense in which the logical stratification of the *Begriffsschrift* can be made to vary independently of the logical stratification of the realm of *Bedeutungen*, and this requires a conception of *Begriffsschrift* signs as something along the line of mere syntactic units or mere figures that are constituted logically independently of their use in the expression of judgment. In this way, the realist conception relies on exactly the sort of externality between what is used and its being used that is yielded by the contemporary syntax/semantics distinction, and that was rejected in the previous chapter.

It may be objected that my criticism of the realist conception entails that there can be no such thing as a 'logically inadequate' notation, a notation that in some sense does not succeed in 'manifesting the grain of reality'. But Frege himself believed that there were such notations – such as Boole's and Peano's – so this cannot be right. The question is not, however, whether there are notations that are in some sense inadequate, but rather what that sense must be. On the realist conception, the problem with Boole's and Peano's notation is that their syntactic structure does not fully correspond to the logical stratification of the realm of *Bedeutungen*. Conceived purely as a syntactic system, however, they are on the same logical footing as the *Begriffsschrift*, which is itself just another syntactic system. That the *Begriffsschrift does* succeed in manifesting the grain of

[8] Compare: "it is required of a syntactic system as a whole, if it is to be such as to allow for its elements to be meaningful, that it generate not just any forms but rather *logical* forms—those forms, that is, which are found independently in reality. In order for its units (propositions) to be appropriate for representation, language as a whole, the realistic thought will be, is required to replicate, in its syntax, the logical structure of reality" (Johnston, 2007, 388). Note that Johnston is also discussing the *Tractatus*.

reality, lies in an external relation of correspondence that happens to obtain between its syntactic structure and the logical stratification of the realm of *Bedeutungen* – in its making the right *click*, as it were – a correspondence that does not pertain to the nature of the *Begriffsschrift* itself, conceived as a purely syntactic system.

On the sort of conception I am advocating, the thing to say about notations such as Boole's and Peano's is not that they do not correspond to an independently constituted realm of *Bedeutungen*, but rather that the signs themselves do not fully realize their nature of being used in the expression of judgment. The 'norm of adequacy', we could say, lies not in an external relation of correspondence, but in the nature of the signs themselves.[9] For signs to fully realize their compositional nature, is for them to fit together in the expression of judgment as *Begriffsschrift* signs do. In the logical stratification of the *Begriffsschrift*, we could say, signs come to their own. Insofar as Boole's and Peano's notation can be used to express judgments, their notations go some way towards realizing the compositional nature of the sign. But because those notations fall short of realizing the compositionality of the *Begriffsschrift*, their signs nevertheless fall short of what their *Zweck als Zeichen verlangt* – to use this nice formulation of Frege's that we encountered in the previous chapter – they fall short of fully realizing their purpose of signifying.

Let us continue our discussion of realism. It is important to see that realism also leads us back to naïve ineffabilism. On the realist approach, we possess a self-standing understanding of the logical stratification of the realm of *Bedeutungen* which informs the construction of the *Begriffsschrift*. This self-standing understanding will have to be articulated through statements such as (A)[10] above, conceived as purporting to state the function/object distinction as it obtains in the realm of *Bedeutungen*. On a realist conception, statements such as such as (B)[11] can make no genuine contribution to the elucidation of the logical categories. At the same time, however, (A) cannot succeed in stating the function/object distinction for the familiar reasons, which is why it is an elucidation. But now there is nothing left but to conceive of (A) as somehow ineffably

[9] Compare: "When the design of a concept-script goes wrong, it is not that the structure of the script fails to match something external to it, but that the concept script has structural features which diverge from its own, as it were, inevitable inner structure, which is thus not revealed clearly" (Diamond, 1984, 354). Again: "The distinctions which the concept script must have in it are not fixed by anything external to it. Nothing external to it fixes its logical structure; but it is not arbitrary. *Where* then is the reality that fixes what distinctions must be embodied in it? That reality lies *in it*. There is an order, a logical order, *in* thought and *in* language" (Diamond, 1984, 366).

[10] 'A function never takes the place of an object'.

[11] 'A function name never takes the place of a proper name'.

gesturing at the function/object distinction, which is thereby conceived as an ineffable truth. That is to say: as soon as we think that it is (A) alone that must fulfil the duty of elucidating the logical categories, we cannot but end up thinking of (A) as some kind of failed attempt to state what cannot really be said, an attempt that nevertheless – *in* its failure – succeeds in conveying that which it could not possibly have said, so that we end up with exactly the sort of conception that was criticized in Chapter 1, and which this dissertation seeks to overcome.

5.3 Idealism

Black, to his credit, saw that the realist construal cannot be ascribed to Frege. Since he prefers such a construal himself, he finds Frege's position problematic. In finding it problematic, however, he falls into what we might call the mirror image of the realist misreading of Frege. Here he is again:

> There is, indeed, something absurd in trying to base inferences about the logical structure of reality upon any physical characteristics of the signs we use. Yet Frege comes near to doing so. At times he seems to think of the sign 'log( )' almost as literally containing a hole, and one which is preserved in every symbol by which it could be faithfully translated. But the physical gap between the brackets is of no ontological interest: exactly the same information might be conveyed by allowing the brackets to approach, overlap, or disappear. All that is 'essential' to the function sign is that we recognize it correctly, without confusing it with other signs: a particular pattern of ink traces may help us in this, but no design can guarantee understanding, nor does 'the nature of things' impose any restrictions whatsoever upon the character of the signs that we can successfully use (Black, 1968, 237).

We already encountered similar remarks in the previous chapter. As before, it is clear that Black is thinking about the unsaturatedness of the function name as a merely geometrical feature of something like a figure, which it possesses independently of its use in the expression of judgment. All that this unsaturatedness may do, is serve as a visual reminder of the logical grammar that has been externally imposed upon the figure, but nothing more.

Thus, the above passage further confirms my account of Black's views from the previous chapter. What is important to our current purposes, is how Black's way of thinking about the unsaturatedness of the function name connects to his way of thinking about the relation between the logical stratifications of the *Begriffsschrift* and the realm of *Bedeutungen*. Black recognizes that Frege wishes to assign fundamental logico-philosophical importance to the unsaturatedness of the function name. The only way in which Black can understand this, however, is as giving rise to a way of basing inferences about the logical structure of reality upon the merely physical characteristics of the signs we use, something he regards as patently absurd. On this sort of view, it is the logical stratification of the *Begriffsschrift* that is taken to be logically prior to the logical stratification of the realm of *Bedeutungen*, so that the latter is determined by the former. This gives us a version of what could be called *idealism* or *linguistic idealism*.[12]

Although an idealist construal of the logico-philosophical correspondence between the *Begriffsschrift* and the realm of *Bedeutungen* is, I think, not more problematic than a realist construal, it is more universally recognized as inadequate, because it does not even generate the illusion of the idea that there are genuine constraints on how the *Begriffsschrift* must be set up in order to 'adequately manifest the grain of reality'. It relieves the logical stratification of the *Begriffsschrift* of any objective purport it may be taken to have.[13] What constraints there are, must now be construed as purely psychologistic in nature, grounded in what we happen to think about how the *Begriffsschrift* is to be constructed, where what we so happen to think is not answerable to anything beyond the psychologistic realm of what we happen to think.

[12] I wish to point out that, in construing idealism as this sort of position, I am following a certain tradition. There is, however, a way of construing idealism according to which it amounts to precisely the sort of view that I will try to articulate below, a view that seeks to understand the relation between the *Begriffsschrift* and the realm of *Bedeutungen* as itself a form of internality that does justice to the internality of its use to the sign. This latter construal of idealism would do a much better job of capturing what is at issue in what has come to be known as 'German idealism'. The difference between these two senses of 'idealism' may be compared to the difference between what Kant calls *empirical idealism* and *transcendental idealism* proper. What I have called realism will then correspond to what Kant calls *transcendental realism*.

[13] Compare: "The kind of realist I imagined thought that the logical characteristics of things and the logical structures of our mode of expression *ought* to match but that they might not do so. It is actually part of such a view that the only alternative to it which can be conceived by someone who holds it is: in our modes of expression *anything* goes" (Diamond, 1984, 367).

Such a view is, of course, too blatantly psychologistic to qualify as a reading of Frege.[14] Above, I called it the 'mirror image' of the realist view, since it involves the same form of externality between what is used in the expression of judgment and its being so used. It only reverses the purported order of priority between the *Begriffsschrift* and the realm of *Bedeutungen*.[15] In conceiving of the logical stratification of the *Begriffsschrift* as logically prior to the logical stratification of the realm of *Bedeutungen*, the former must be conceived as governing an order that is constituted logically independently of its use in the expression of judgment. Idealism, then, gives us exactly the sort of hypostasizing into the realm of *Bedeutungen* of what can be nothing more than arbitrary convention that was discussed and rejected in the previous chapter in connection with the syntactic conception of the sign.

5.4 The internality of compositionality and functionality

We need an account of the correspondence between the respective logical stratifications of the *Begriffsschrift* and the realm of *Bedeutungen* that leaves behind the form of externality that underlies both realism and idealism. A good place to start is the passage from the *Grundgesetze* that was already partly discussed in the previous chapter, and which we can now consider in full:

> We recognize in these examples the great multiplicity of functions. We can also see that there are fundamentally different functions, since the argument places are fundamentally different. In particular, those which are suited to take proper names

[14] Note that construing it as blatantly psychologistic in this way, itself depends on a certain way of understanding the relevant notion of 'psychology'. This touches upon the theme from Cavell of how later Wittgenstein may be read as depsychologizing psychology (Cavell, 2002, 91). Of course, if one transforms the operative notion of psychology, the dichotomous contrast with what I have called realism will vanish.

[15] Compare: "Frege's use of the ontological term 'object' is strictly correlative to his use of the linguistic term 'proper name': whatever a proper name stands for is an object, and to speak of something as an object is to say that there is, or at least could be, a proper name which stands for it. The question therefore naturally arises in which realm, the linguistic or the ontological, the primary principle of classification is to be applied" (Dummett, 1981, 55-56). Compare also: "Some commentators tell us that the crucial idea here is that there are two isomorphic 'structures'—there is a structure in the world and a structure in language (or thought)—and in order to say (or think) what is true, one of the two structures must 'reflect' the other" (Conant, 2020, 962). Conant is talking about the *Tractatus*, and wishes to resist such a reading. In passing, he indicates that he would equally resist such a reading of Frege (Conant, 2020, 963).

> cannot receive names of functions, and conversely. Moreover, argument places that can take names of first-level functions with one argument are unable to take names of first-level functions with two arguments. Accordingly, we distinguish:
>
> > *arguments of the first kind*: objects;
> > *arguments of the second kind*: first-level functions with one argument;
> > *arguments of the third kind*: first-level functions with two arguments.
>
> Likewise, we distinguish:
>
> > *argument places of the first kind*, that are suitable to take proper names;
> > *argument places of the second kind*, that are suitable to take names of first-level functions with one argument;
> > *argument places of the third kind*, that are suitable to take names of first-level functions with two arguments.
>
> Proper names and object-letters *fit* [*passen*] the argument places of the first kind; names of first-level functions with one argument *fit* the argument places of the second kind; names of first-level functions with two arguments *fit* argument places of the third kind. The objects and functions whose names fit the argument places of names of functions are *fitting* arguments for these functions. Functions with one argument, for which arguments of the second kind are fitting, we call *second-level functions with an argument of the second kind*; functions with one argument, for which arguments of the third kind are fitting, we call *second-level functions with an argument of the third kind* (GGA, I, §23).

In the previous chapter, we discussed how Frege invokes this notion of *fit* to articulate the logical stratification of the *Begriffsschrift*. What we see now, is that he applies the same notion of fit to the logical stratification of the realm of *Bedeutungen*. If we adhere to the form of externality governing the realist and idealist approach, Frege will be regarded as using the same term 'fit' to discuss what are two logically independent phenomena: the logical stratification of the *Begriffsschrift* and the logical stratification of the realm of *Bedeutungen*. Seen from such a vantage point, Frege's use of that term can only appear as ambiguous and therefore highly misleading. At the very least, we would need to disambiguate between the notion of fit that governs the *Begriffsschrift* and the notion of fit that governs the realm of *Bedeutungen*.

On such an approach, we find Frege being sloppy and careless at a crucial juncture in the *Grundgesetze*, since this elucidation of the different kinds of arguments and argument places goes to the very core of the *Begriffsschrift*. At the very least, we would expect Frege to flag his use of the term 'fit' as misleading, we would expect to find him, perhaps, asking for another grain of salt. But this is not what Frege does. Since we can disregard the hypothesis that he was simply oblivious to the issue as ludicrous, the conclusion must rather be that we find Frege using the term 'fit' exactly as he thinks it *should* be used. Thus, he must have believed that there is indeed a single elucidatory notion of fit that characterizes *both* the logical stratification of the *Begriffsschrift and* the logical stratification of the realm of *Bedeutungen*, and this shows that he believed that these were not logically independent. Instead, what we have here are two logico-philosophical phenomena that are *internally related* to each other. The fact that both phenomena are to be elucidated through this single elucidatory notion of fit, earmarks precisely their internality.

How are we to understand this internality? Consider the statement 'Objects are fitting arguments for first-level functions'. We can rephrase this statement as follows: In being the *Bedeutungen* of proper names and first-level function names respectively, objects and first-level functions are such that they are the *Bedeutungen* of signs that fit together in the expression of judgment as proper names and first-level function names fit together. Thus understood, our understanding of the functionality of the realm of *Bedeutungen* goes through our understanding of the compositionality of the *Begriffsschrift*, in the following way:

(FIT-1) For *Bedeutungen* to fit, is for them to be *Bedeutungen* of names that fit.

It may seem that this brings us back to a form of idealism, according to which the compositionality of the *Begriffsschrift* is logically prior to the functionality of the realm of *Bedeutungen*. This overlooks, however, that we can move in the other direction as well. Take the statement 'Proper names fit the argument places of first-level function names', a statement characterizing the compositionality of the *Begriffsschrift*. We can now rephrase this statement as follows: In having objects and first-level functions as their *Bedeutungen* respectively, proper names and first-level function names are such that they have *Bedeutungen* that fit together as objects and first-level functions fit together. This gives us:

(FIT-2) For names to fit, is for them to be names whose *Bedeutungen* fit.

This time, it will seem as if (FIT-2) commits us to a realist approach.

In order to overcome this realism/idealism dichotomy, we must come to see how neither of (FIT-1) and (FIT-2) give us a self-standing expression of the correspondence between the compositionality of the *Begriffsschrift* and the functionality of the realm of *Bedeutungen*. What we do *not* have here, are *two* notions of fit – one for the *Begriffsschrift* and one for the realm of *Bedeutungen* – which are such that the one is to be explained in terms of the other, as both (FIT-1) and (FIT-2) may be read as doing. Rather, what we have is *one* notion of fit the unity of which is brought out precisely by the elucidatory interplay of (FIT-1) and (FIT-2) – the way in which each is internal to the other – two elucidatory statements that serve to elucidate that notion *in tandem*. We have an internal unfolding of a single notion of fit, such that a proper understanding of that notion *immediately* yields both (FIT-1) and (FIT-2) as constituting two interrelated moments of that internal unfolding. To understand how (FIT-1) and (FIT-2) constitute two interrelated moments in a single elucidatory unfolding of Frege's elucidatory notion of fit that characterizes both the compositionality of the *Begriffsschrift* and the functionality of the realm of *Bedeutungen*, *is* to have a proper understanding of that notion of fit. And so to understand this elucidatory notion of fit, is to understand the form of internality that obtains between the logical stratification of the *Begriffsschrift* and the logical stratification of the realm of *Bedeutungen*, and thereby to understand Frege's elucidatory notion of correspondence. This reveals that the term 'correspondence' is somewhat misleading, insofar as it comes with connotations of externality.[16] It would perhaps be better to talk in terms of something like the *attunement* of the logical stratification of the *Begriffsschrift* and the logical stratification of the realm of *Bedeutungen*.

We are now in a position to better understand Frege's elucidatory turn to signs. Consider again these two elucidatory statements:

(A) A function never takes the place of an object.
(B) A function name sign never takes the place of a proper name.

To Black, the way in which Frege approaches (A) through a consideration of (B) can only appear as manifesting a commitment to a form of idealism. Black's mistake lies in thinking that (A) and (B) each form self-standing articulations of two self-standing phenomena:

[16] As noted, Frege's German term is '*entsprechen*'. I am not in a position to judge to what extent this term comes with similar connotations.

the functionality of the realm of *Bedeutungen* and the compositionality of the *Begriffsschrift*. Understood in this way, Frege's elucidatory turn will be seen as engendering a commitment to the logical priority of (B), so that it furnishes us with an independent ground for (A).

In bringing in (B) in order to attain an understanding of (A), however, Frege is not trying to bring in an order – the order of the sign – that is external to the realm of *Bedeutungen* at all. Rather, Frege is trying to show how a proper unfolding of (A) *itself* involves bringing in (B). He is trying to show how a proper understanding of (A) *is* an understanding of the internality of (A) and (B). Rather than, we could say, moving us away from (A), Frege seeks to guide us deeper into (A), by unfolding the internality of (A) and (B). That is why it is indeed essential to the elucidation of the logical categories that it involves an elucidation of the logical stratification of the *Begriffsschrift*. Only through the latter can the requisite internality be brought out. There is nothing optional about Frege's discussions of *Begriffsschrift* signs, nor could he have presented his discussion of the notion of *fit* solely in terms of the realm of *Bedeutungen*. Rather, that the elucidation of the logical categories must involve *both* the realm of *Bedeutungen* and the *Begriffsschrift*, constitutes the very core of the phenomenon to be elucidated, which is nothing else than the internality of the compositionality of the *Begriffsschrift* and the functionality of the realm of *Bedeutungen* itself. To understand that internality, is to understand Frege's conception of the logical categories.

The logico-philosophical form of unity of the realm of *Bedeutungen* is its functionality. The logico-philosophical form of unity of the *Begriffsschrift* is its compositionality. We can now see that these forms of unity must be taken together as comprising two sides of the same coin. It is only through its internality with the *Begriffsschrift*, that we can grasp the unity of the realm of *Bedeutungen*, and *vice versa*. The unity of the *Begriffsschrift* consists in its compositionality, which goes through the relations of fit that govern *Begriffsschrift* signs. The unity of the realm of *Bedeutungen* consists in its functionality, which goes through the relations of fit that govern *Bedeutungen*. But these relations of fit are themselves internally related, so that a proper understanding of the logical stratification of either *ipso facto* constitutes a proper understanding of the logical stratification of both. The functionality of the realm of *Bedeutungen* must go through the compositionality of the *Begriffsschrift*. It is *in* being the *Bedeutungen* of *Begriffsschrift* signs that fit together in *Begriffsschrift* judgments in the way that they do, that those *Bedeutungen* fit together in the way that they do. This was (FIT-1). Conversely, the compositionality of the *Begriffsschrift*

must go through the functionality of the realm of *Bedeutungen*. It is *in* having *Bedeutungen* that fit together in the way that they do, that *Begriffsschrift* signs fit together in *Begriffsschrift* judgments in the way that they do. This was (FIT-2). There is no self-standing logico-philosophical unity of either the *Begriffsschrift* or the realm of *Bedeutungen*. There is only the logico-philosophical unity that they possess *in* being internally related in the way that they are.

To further our grasp on these matters, it is helpful to look at a critical discussion of Frege's views by Kevin Klement. In discussing Frege's logical category distinctions, Klement advances what he calls the *Bridge Principle* as crucial to Frege's conception (Klement, 2004, 5):

> *Bridge Principle*: No entity can be both the referent of an expression containing an empty spot and the referent of an expression not containing an empty spot.[17]

Only if the bridge principle obtains, Klement says, will Frege's grammatical distinctions between different kinds of expressions such as proper names and function names yield his logical category distinctions. But, Klement adds, we find surprisingly little argument for the Bridge Principle in Frege.

To discuss Klement's arguments in their full detail would take us too far afield here. What I wish to do, is make palpable how his way of approaching the relevant issues differs crucially from my own. We can start from the observation that it may seem that I must also accept the Bridge Principle as articulating a central commitment of Frege. In a sense, this is correct: there is a construal of the Bridge Principle that I must accept as a commitment of Frege – although the terminology of 'empty spots' is quite infelicitous. At the same time, however, the reason why Klement finds a lack of argument for the Bridge Principle in Frege, is that his construal of it is quite different. More specifically: Klement does not understand the notion of an expression containing an empty spot along the lines of Frege's conception of the sign as I have presented it. This comes out, for instance, in the following passage:

> Despite their centrality in his philosophy, Frege's philosophical argumentation and explanation for [the logical category distinctions] are not entirely satisfactory. While he strongly intimates that the notions of *object* and *function* are incapable of

[17] This may be compared to what Hale and Wright call the *Reference Principle* (Hale & Wright, 2012, 93-94).

> definition, his usual explanation is that the hallmark of a function is that the expression for it exhibit an empty spot or spot for completion, whereas the hallmark of an object is that the expression for it not exhibit any such empty spot [...]. *If the distinction is founded in the nature of things, and not in language, this would seem to characterize the difference by inessential features* [last emphasis mine] (Klement, 2004, 4-5).

In contrasting 'the nature of things' with 'language' in the way that he does, Klement betrays an adherence to the sort of externality between the *Begriffsschrift* and the realm of *Bedeutungen* which I have argued to be incompatible with Frege's philosophical logic. In line with Black, he can only see Frege's characterization of function names as unsaturated as pertaining to 'inessential features' of what is constituted logically independently of its use in the expression of judgment. That is why Klement feels that an additional principle is needed to connect in the right way the grammatical distinctions of the *Begriffsschrift* with the logical category distinctions in the realm of *Bedeutungen.* Klement is then struck by the apparent lack of argument for such a principle in Frege. This apparent lack of argument, however, is not due to philosophical oversight on Frege's part, but is rather due to the fact that it is only *if* there is a gap to be bridged of the kind Klement takes there to be, that additional argument will be required to bridge that gap.[18]

But, it could be objected, does not the Bridge Principle on my way of construing it also require argument? In a sense, yes, but not in the way Klement thinks. On a properly Fregean understanding of the matters at hand, the Bridge Principle constitutes an unfolding of the internality of the respective logical stratifications of the *Begriffsschrift* and the realm of *Bedeutungen.* Crucially, this internality is not a logico-philosophical phenomenon that must be superadded to their respective logical stratifications in order to suitably connect them. Rather, the very nature of the compositionality of the *Begriffsschrift* and the functionality of the realm of *Bedeutungen* is such that they are *ipso facto* internally related as the logico-philosophical phenomena that they are.

On Klement's picture, we have three separate phenomena: the grammatical distinctions of the *Begriffsschrift*, the logical category distinctions of the realm of *Bedeutungen*, and the correspondence between them that is stated in the Bridge Principle.

[18] Compare Johnston's claim (about Frege) that "neither side [of the syntax-ontology pair] has substance independent of the other—substance on the basis of which the alignment of their categories might become a philosophical thesis" (Johnston, 2017, 576). Although he does not discuss this explicitly, Johnston's notion of syntax cannot be the one that figures in the contemporary syntax/semantics distinction.

We then also have three separate activities of coming to an understanding of each of these, where an understanding of the last phenomenon will depend on a prior understanding of the first two. On Frege's conception, however, what Klement regards as three separate phenomena come to be regarded as two internally related phenomena: the compositionality of the *Begriffsschrift* and the functionality of the realm of *Bedeutungen*. Thus, there is no separate understanding of the Bridge Principle in play that requires separate argument. Rather, to properly understand the compositionality of the *Begriffsschrift is* to properly understand its internality with the functionality of the realm of *Bedeutungen*, and *vice versa*. Bringing out their internality is not a matter of providing philosophical argument for a self-standing thesis that moves beyond a proper understanding of these two phenomena themselves, but is rather a matter of unfolding what goes into such an understanding to begin with. We can, if we want, label such an unfolding as 'argument' for the Bridge Principle, but then this 'argument' must be found in the place where Klement thinks it cannot possibly be found: In Frege's elucidations of what Klement takes to be 'inessential features' of the *Begriffsschrift* itself. To properly understand the Bridge Principle, we could say, is to understand why Klement is looking in the wrong place for the wrong kind of argument.[19]

5.5 Internal heterogeneity

I now wish to discuss a reading of Frege that is closer to my own: that of Peter Geach. In his discussion of Frege, Geach seeks to do justice to the fundamental importance that Frege attaches to the unsaturatedness of function names without falling into the sort of

[19] It is worth quoting a long passage from Diamond here: "The objector ascribes to Frege a view of a certain kind, a substantial thesis that Frege has (the objector thinks) built into his concept script, but which can be discussed independently. The supposed Fregean thesis is that nothing can be thought of things of more than one logical kind, and the objector believes that it has been built into the concept script in the rules governing what can go into argument places. He thinks that a *better* concept script would allow Socrates and the concept *bald* into the same place in some thoughts (e.g., ascriptions of self-indentity) and would allow Socrates but not the concept *bald* into a place in *other* thoughts, e.g., those ascribing baldness. One could disagree with Frege about what a good concept script would allow without disagreeing in the fundamental way our objector does. He does not so much reject a view of Frege's as not see what kind of view it is. He does not see the possibility of letting the use teach you what you are talking about, where logical analysis is what shows you how a term is being used, what kind of contribution it makes (what kind of contribution logic sees it making) to the sentences in which it occurs" (Diamond, 1984, 364).

idealism discussed above, in a way which it is instructive to investigate in some detail. Geach seeks to provide a response to readers of Frege who are inclined to say something along the following lines with regards to Frege's logical category distinctions:

> Frege's trouble would be diagnosed as his having thought of all signs as *names* – a diagnosis confirmed by his use of the term '*Functionsnamen*' for functors. If all signs have to be names, each one standing for something, then signs that do not name or stand for objects will have to be assigned some strange non-objects, such as concepts and functions, as their *Bedeutungen*. What he ought to have done is to distinguish the different mode of significance of signs; instead he misconstrued these distinctions as difference of ontological category between things named (Geach, 1976, 58).

Although Geach does not explicitly mention him, this is clearly reminiscent of Black's criticism of Frege.[20] Geach then proceeds to discuss Frege's function names – as containing Greek letters – and explains why he thinks this criticism misses the mark:

> The role of these expressions with a Greek letter ['ξ'] in them is to serve as stencils for constructing numerical signs; the numerical signs themselves are values of a certain function *from* numerical expressions *to* numerical expressions; so the function/object distinction reappears at the level of language and language about language cannot be used to charm away the distinction (Geach, 1976, 59).

On Geach's reading, there is a straightforward reason why Frege uses the term 'unsaturatedness' to characterize both the nature of function names and functions: function names *are* functions, so that they are unsaturated in exactly the same sense in which functions are.[21] We could say that, on Geach's reading, the form of internality that obtains between the logical stratification of the *Begriffsschrift* and the logical stratification of the realm of *Bedeutungen* is the strongest form of all: identity.[22] On Geach's view, the distinction between function names and proper names just *is* the distinction between

[20] Recall especially the passage at (Black, 1968, 246) that was quoted in Chapter 2.
[21] To be clear: Geach's view does not amount to the view that sign-types are concepts (and thereby functions), whereas sign-tokens are objects that fall under those concepts. If one feels compelled to characterize Geach's view in type/token terms, the best way to do so would be by saying that Geach thinks that tokens of first-level function names are first-level functions that map proper names to proper names, whereas the associated types are second-level concepts under which the relevant tokens fall.
[22] Similar accounts are advanced in (Long, 2001) (Noonan, 2006).

functions and object, so that the compositionality of the *Begriffsschrift is* the functionality of the realm of *Bedeutungen*.

At least, that is how it seems natural, at first sight, to characterize Geach's view. In the end, we will have to conclude that there is a way in which it is misleading to say that, for Geach, the compositionality of the *Begriffsschrift* is identical to that of the realm of *Bedeutungen*. For now, let us consider another passage:

> In fact, the Fregean *Functionsnamen* [sic] can *never* figure as physical parts in Frege's logical notation, since the Greek consonantal letters never occur in any formula. Of course such a signs as '$\log_{10}$' could appear (if suitably defined); but this is not the *Functionsname* – '$\log_{10}$' is [sic][23]. And the point of insisting on this is that an isolated functional sign is for Frege a monstrosity; what signifies a given function is not the presence in a formula of a given piece of type, but the occurrence of a given pattern – in this case, the prefixing of '$\log_{10}$' to some proper name or some proxy for a proper name (Geach, 1976, 60).

In talking about function names, Frege employs Greek letters. But, Geach points out, those Greek letters never make their way into actual *Begriffsschrift* judgments. This shows that a purely typographical account of the function name – on which the Greek letters would have to accompany the function name in all its occurrences – cannot be what Frege has in mind. It is not – as Geach puts it – the presence of a given piece of type that signifies a function in a formula.

It is true that Frege does not have a typographical account of the unsaturatedness of the function name, and this serves to undermine the sort of reading proposed by Black, as we have seen. The problem, however, is that Geach believes that this means that we cannot conceive of function names as signs that occur in the expression of judgment at all. Here is what he says immediately after the previous passage:

> The same goes for predicates, and for symbolic concept-expressions: the occurrence of a predicate too must be recognized from the occurrence of a pattern, not from the occurrence of a quotable part of a sentence. Even when one is tempted to identify a predicate with such a quotable part, the presence of the predicate is

[23] Given the previous sentence, I am assuming that Geach wants to say here that '$\log_{10}\,\xi$' is the function name rather than '$\log_{10}$', in accordance with what Frege himself would say. I will proceed on this supposition.

> constituted not by this but (say) by this quotable part's being appended to a proper name, or some proxy for a name such as a relative pronoun (Geach, 1976, 60).

Consider a judgment such as '$\log_{10} 100 = 2$', in which the proper name '$\log_{10} 100$' occurs. Geach thinks that it would be a mistake to say what we have here, is the function name '$\log_{10} \xi$' prefixed to '100'. As he sees it, this constitutes exactly the sort of view that is blocked by the observation that no Greek letter appears in '$\log_{10} 100 = 2$'. On Geach's view, what is prefixed to '100' is not aptly called a function name at all, since the fact that a function is designated does not consist in the occurrence of a sign, but rather in the pattern of prefixing itself, which must be conceived in functional terms. What we have, is not the occurrence of a sign that *bedeutet* a function, but instead the result of *applying* a function that maps the proper name '100' to the proper name '$\log_{10} 100$'. I will call these sorts of functions that figure centrally in Geach's conception *sign-functions*.

My claim is not, of course, that sign-functions do not exist. My claim is rather that they cannot afford us with a proper understanding of the compositionality of the *Begriffsschrift*. Let me try to bring out why this is so. On Geach's view, we have a sign-function that maps '100' to '$\log_{10} 100$', so that something is prefixed to '100' in the latter. But *what* is it that is prefixed? It is not a proper name, of course. But neither is it a function name, since the whole point of Geach's account is that the complexity of an expression such as '$\log_{10} 100$' is not to be conceived in terms of a *composition* of function name and proper name, but rather in terms of the *application* of a sign-function. So there seems to be no conceptual room left to account for the logico-philosophical status of '$\log_{10}$', on Geach's conception.

Contrast this with what I take to be Frege's understanding. Frege would agree that we can say that what characterizes the fact that a function is signified in '$\log_{10} 100$' is a certain pattern. For Frege, however, that pattern is perfectly coherently described in terms of the occurrence of the function name '$\log_{10} \xi$'. What we have, is a proper name fitting a first-level function name so as to compose the proper name '$\log_{10} 100$'.[24] There is here no corresponding mystery about the status of '$\log_{10}$', which *is* a function name, although we are better to use "$\log_{10} \xi$" (double quotation marks) to designate that function name, so as to make recognizable its unsaturatedness.

[24] I am again abstracting from the point that we have here only the privative composition of a complex proper name, and not composition in the expression of judgment.

A second problem I have with Geach's view is that I do not see how it can account for higher-level subsumptions. Take, for instance, the second-level judgment '$(\forall x)\big((x > 10) \supset (\log_{10} x > 1)\big)$'. On Geach's construal, what we must have is – not two function names fitting together – but a proper name that results from the application of a second-level sign-function to a first-level sign-function. What signifies a second-level function in this expression, is the occurrence of the pattern that results from applying the second-level sign-function to the first-level sign-function. But how are we to understand this pattern? The case of a first-level sign-function is comparatively straightforward: the pattern consists in something's being prefixed to a proper name – or a similar notational operation – and it is this proper name itself to which the first-level sign-function is applied. No analogous account is possible for a second-level sign-function. It is applied to a first-level sign-function, but we cannot say that something is prefixed to *that* function, or that some other notational operation is carried out on *it*, since a first-level sign-function is a function, and therefore does not appear on the page at all. What appears on the page, are signs that exhibit the pattern that results from the *application* of the first-level sign-function, not the sign-function itself. This is, to repeat, the whole the point of Geach's account. Perhaps we could say something like the following: The proper name '$(\forall x)\big((x > 10) \supset (\log_{10} x > 1)\big)$' results from taking '$x$' as a dummy proper name, applying the first-level sign-function to it to generate another dummy proper name '$\log_{10} x$', and then performing the requisite 'second-level' notational operation on that dummy proper name to yield '$(\forall x)\big((x > 10) \supset (\log_{10} x > 1)\big)$'. But now we need to account for this peculiar notion of a dummy proper name, which turns out to be essential to a grasp of how second-level sign-functions give rise to notational patterns. Apparently, first-level sign-functions can be applied to dummy proper names such as 'x' to yield '$\log_{10} x$', which is not a proper name at all. What we have in this kind of application, is a mere piece of type 'x' that is mapped onto another mere piece of type '$\log_{10} x$', and it is this mere piece of type that is subsequently operated on to yield the proper name '$(\forall x)\big((x > 10) \supset (\log_{10} x > 1)\big)$'. Similarly, we can also get the proper name '$(\forall x)\big((x > 10) \supset (\sin x > 1)\big)$'. But now we can ask: what makes it the case that the former concerns the function $\log_{10} \xi$, whereas the latter concerns the function $\sin \xi$? There seems to be only one possible answer: the occurrence of the expressions '$\log_{10} x$' and '$\sin x$', respectively. But we just saw how those expressions turn out to be, on Geach's view, mere pieces of type. Thus, we end up with exactly the sort of view that Geach wished

to avoid, according to which the fact that a certain function is signified in an expression consists in the presence of a mere piece of type.

I need not go into detail anymore to explain why no such issues arise on Frege's view, as I have construed it. According to Frege, what we have in '$(\forall x)\big((x > 10) \supset (\log_{10} x > 1)\big)$' is quite simply the first-level function name '$\log_{10} \xi$' fitting the argument place of the second-level function name '$(\forall x)(\forall y)\big((x > 10) \supset (\varphi x > 1)\big)$', and it is the occurrence of both function names that accounts for the fact that a first-level function and a second-level function – the respective *Bedeutungen* of those function names – are signified.

Geach's correct insight is that no account of what is used in the expression of judgment that proceeds in purely typographical terms will give us an adequate account of the compositionality of the *Begriffsschrift*. Geach overshoots, however, in thinking that this renders logically inadequate a conception of this compositionality in terms of signs *as such*, where a sign is something that can properly be said to occur and fit together with other signs in the expression of judgment. This reveals how Geach himself is still caught in an externalist way of thinking about these matters, since he cannot see how a conception of signs as occurring and fitting together in the expression of judgment could be anything other than merely typographical. This leads him to believe that the only way to avoid such a typographical account is to relinquish a conception of function names as signs altogether, and to replace it with an account in terms of sign-functions.

Geach's view can only seem to work as long as we limit ourselves to first-level subsumptions. Making sense of higher-level subsumptions requires a conception of function names as signs that occur and fit together with other signs in the expression of judgment. Conceiving of function names *as* functions, so that we instead have notational patterns that result from the application of such sign-functions, renders impossible a coherent account of how higher-level *Begriffsschrift* judgments are composed.

That Geach's conception is not Frege's, is confirmed by the way in which Frege presents his own elucidations. At one place, after having elucidated the distinction between proper names and function names, Frege writes:

> To this difference in the signs there of course corresponds an analogous one in the realm of meanings: to the proper name there corresponds the object; to the predicative part, something I call a concept (CP, 281; KS, 269, 371)

There are at least two reasons why Geach's reading has trouble accounting for such passages. First, Frege's talk of predicative parts of expressions to which concepts *correspond* is already problematic. On Geach's view, this is at best a misleading way of putting the matter, since concepts do not correspond to *any* part of an expression. Rather, they correspond to *patterns*. Second, insofar as it makes sense, on Geach's view, to speak of a difference in the signs, this would have to be the difference between a proper name and a sign-function. But then it becomes strange to say that what corresponds to that difference in the realm of *Bedeutungen* is an *analogous* difference, since the difference between object and function is supposed to *be* this very difference. Elsewhere, after having elucidated the unsaturatedness of function names, Frege adds: "We can, *metaphorically* [*bildlich*] *speaking*, call the concept unsaturated too; alternatively, we can say that it is predicative in character" [my emphasis] (PW, 193). Again, Geach has no way of making sense of Frege marking his characterization of the concept as unsaturated as metaphorical, since Geach takes it to be the exact same notion of unsaturatedness that is at issue both with regards to sign-functions and functions. Frege himself, however, clearly wishes to distinguish between the unsaturatedness of the concept and the unsaturatedness of the concept name, so that these are not identical.[25]

The compositionality of the *Begriffsschrift*, then, is not a mere copy of the functionality of the realm of *Bedeutungen*. This can be brought out through a couple of observations. If we fit a proper name into the argument place of a first-level function name to yield a *Begriffsschrift* judgment 'Fa', both are constituents of this judgment, which ramifies into full-blooded occurrences of these signs fitting together in this judgment. That both signs *occur* in the judgment, is essential to the compositionality of the *Begriffsschrift*, as we have seen. But if we fit an object into the argument place of a first-level function, the result is an object in which neither the first-level function, nor – in most cases – the original object can be said to 'occur' in any proper sense of the term. These two modes of logical stratification – compositionality and functionality – do not have what we might call the same logical behavior. A second crucial difference is that there are privative occurrences of signs, but no privative occurrences of *Bedeutungen*.[26] In composing *Begriffsschrift* judgments, signs realize their nature in a sense in which they *need* not do so. In functional

[25] Compare also (CP, 292).
[26] This is why Wittgenstein said that "The limit [to thinking] can, therefore, only be drawn in language and what lies on the other side of the limit will be simply nonsense" (TLP, p. 27). It is also why he said that "We cannot give a sign the wrong sense" (TLP, 5.4732).

application, objects and functions may perhaps be said to realize their nature, but there is no sense in which they might fail to do so. The compositionality of the *Begriffsschrift* leaves room for cases that fall short of such compositionality, the functionality of the realm of *Bedeutungen* does not. These two observations suffice to show that the compositionality of the *Begriffsschrift* cannot be identical to the functionality of the realm of *Bedeutungen*, contrary to what Geach thinks. To take them to be identical, is to gainsay the nature of both.

Above, I discussed how (FIT-1) and (FIT-2) together bring out how the unity of the *Begriffsschrift* and of the realm of *Bedeutungen* are two sides of the same coin. We can now understand this a bit better. To account for the unity of either, one needs the unity of both, and these are not identical, contrary to what Geach thinks. The relations of fit that give the *Begriffsschrift* its unity, yield a notion of compositionality that is constituted by the way in which *Begriffsschrift* judgments ramify into full-blooded occurrences of *Begriffsschrift* signs that fit together. The relations of fit that give the realm of *Bedeutungen* its unity, yield a notion of functionality that is constituted by the ways in which functions can be applied to arguments to yield objects as values. To understand Frege's conception of the logical categories, is to understand how these two modes of logical stratification – compositionality and functionality – that constitute the respective logico-philosophical forms of unity of the *Begriffsschrift* and the realm of *Bedeutungen* are internally related, in such a way that neither is logically prior to the other and neither is logically independent of the other. But, crucially, we must add to this that they are also not *the same*, that there is a genuine heterogeneity between them.[27] Contrary to what Geach thinks, to compose a judgment is *not* to apply a function. Although their modes of fit are internally related, they are different, and both are essential to a proper understanding of the logical categories. To account for the unity of either the *Begriffsschrift* or the realm of *Bedeutungen*, we need *both* the compositionality of the *Begriffsschrift and* the functionality of the realm of *Bedeutungen in* their heterogeneous internality.[28]

---

[27] My notion of heterogeneity is inspired by Gertrudis Van de Vijver.

[28] Perhaps it is in the holding together of these two heterogeneous registers that we can find a fundamental logico-philosophical role reserved for Frege's notion of *Sinn*, which would make it fulfil a function that is similar to that of Kant's notion of a *schema*. *Sinn* would have to be what, in some sense to be specified, brings together compositionality and functionality, where this bringing together is of course not to be conceived as the bridging of a gap between two logically independent phenomena – along the lines of Klement's Bridge Principle – but as a further essential moment in the unfolding of their internality. Frege certainly says things that indicate that

5.6 Interlude: Signs as objects?

Beyond what I have said already, there are further exegetical reasons to think that Frege did not believe that the notion of a function name must be refurbished into Geach's notion of a sign-function. These further reasons, however, come with philosophical problems of their own. Several things Frege says (or does not say) indicate that he took function names to be *objects*.[29] Such a view is incompatible with Geach's view, which replaces the notion of a function name with that of a sign-function. If we still wish to talk about function names, the proper thing to say would be that they are themselves functions, not objects.

The clearest case comes from a passage from Frege's first reply to Russell discussed in Chapter 3, where we find him blankly stating: "While '$(\,) \cdot 3 + 4$' is a function name, '"$(\,) \cdot 3 + 4$"[30] is a proper name, and its meaning [*Bedeutung*] is the function name '$(\,) \cdot 3 + 4$'" (PMC, 136). This notion of a function name being the *Bedeutung* of a proper name is incompatible with Geach's reading, since it implies both that function names are objects, and that they are quotable, whereas Geach repeatedly insists that functions or predicates "cannot be identified with any separately quotable and displayable part of [...] sentences" (Geach, 1976, 60). More generally, Frege uses expressions of the form 'The function name '$\Phi(\xi)$'' throughout his *oeuvre*, and *at no point* does he exhibit the sort of qualms that he constantly expresses when using locutions such as 'The function $\Phi(\xi)$'. On Geach's conception, this makes no sense, since the former runs into exactly the same sorts of problems as the latter. In the *Grundgesetze*, to raise a final point, we find the following passage:

> In what follows, we propose to understand by 'equal-shaped signs' [*gleichgestalteten Zeichen*] those that, according to the intention of the writer, are meant to be equal-shaped in order to designate [*bezeichnen*] the same thing. Now, it is common to express oneself loosely and to speak of equal-shaped signs as one and the same sign,

---

he regards *Sinn* as, in some sense, inheriting compositionality: "The structure of the sentence can serve as a picture of the structure of the thought" (CP, 390). Whether he also takes *Sinn* to inherit functionality is a more difficult question. Perhaps this is one of the upshots of his conception of indirect discourse, on which the *Sinn* of names comes to be *bedeutet*. In any case, I am at the present moment incapable of fleshing out any further the suggestion under consideration.

[29] This view is also taken by (Ricketts, 2010, 184) (Moore, 2012, 218) (Heck & May, 2013, 845-846).

[30] Note the double quotation marks.

> although in fact I create another object every time I write down an equality-sign (GGA, II, §99).

Frege's notion of equal-shapedness itself raises a host of philosophical questions – to be touched upon briefly below – but my main interest lies in Frege's claim that another object is created every time an equality-sign is written down. These objects are, of course, equal-shaped, so that – according to Frege's characterization – they are meant to designate the same thing. On the view that function names such as the equality-sign are objects, these are natural things to say.[31] On Geach's view, however, Frege is expressing himself in a way that is highly misleading at best. What Frege is here calling 'an equality-sign' cannot be conceived as designating equality. Rather, what signifies equality is the presence of a pattern that results from applying a sign-function to two proper names, an application that – presumably – results in having them flank '=', where different occurrences of this '=' will be characterized as 'equal-shaped'. As explained before, it is unclear how to make sense of the status of this '=' on Geach's view, but whatever it turns out to be, it does not designate equality.

Given the negative point that function names are not themselves functions, Frege's philosophical logic seems to leave no other option than to take them to be objects, at least if we suppose that they must fit *somewhere* into the realm of *Bedeutungen*. As is clear by now, there are many passages that suggest that Frege did indeed take this view, which would reveal that Frege did make this supposition. This view, however, is not itself without issues, which I will now briefly try to bring out. It results in a conception of signs as constituting a species of objects, so that there are no *categorical* differences between signs at all. When we say, for instance, that proper names are complete and that function names have argument places, this is logically on a par with saying, for instance, that whales are mammals and that sharks are fishes. Such a reading, it must be admitted, would make good sense of Frege saying to Russell: "If we want to express ourselves precisely, our only option is to talk about words or signs" (PMC, 141). The mode of precision at issue here would be that of scientific judgment as such, since a discussion of

[31] It could be objected that it is incorrect, on Frege's view, to speak of equal-shaped equality-signs as 'designating the same thing', since what is designated is not a thing, but a function. The German is '*dasselbe zu bezeichnen*', which does not have the same objectual connotations. Still, Frege would agree, of course, that the terms '*dasselbe*' and '*bezeichnen*' are here employed with their logically stratified elucidatory use. If he wished his characterization of 'equal-shaped signs' to be limited to proper names that designate objects – as it would have to be, on Geach's reading – he would not have chosen the equality-sign as an example.

signs turns out to be a discussion of a *species* of objects that can be studies scientifically as any other species of objects can.[32] Frege's notion of equal-shapedness, for instance, turns out to be a first-level relation between objects, a relation that is to account for the repeatability of the material sign.[33]

I believe, however, that such a view cannot do justice to Frege's better insights regarding the nature of the sign, which revolves around the internality of its use to the sign. I cannot provide a full discussion of this claim here, however, for the interrelated reasons that it would require at least another full chapter, and that I myself do not yet understand these matters well enough. All I can do, then, is to delineate what I take to be the shape of the main problem facing what seems to be Frege's view.[34] This problem, I think, can be put by saying that each of the following must – on a logically proper conception of the sign – turn out to be *sui generis*: the distinctions between different categories of signs, the order of the sign as such, and the logical stratification of the

[32] None of this, to be clear, should be taken as showing that Frege was not, after all, committed to the internality of its use to the sign. In line with that commitment, signs would have to be understood as objects the nature of which it is to be used in the expression of judgment. To be sure, this raises a host of issues, but the upshot is not that we are now free to disregard everything Frege says about the internality of its use to the sign. Rather, it reveals that Frege did not fully think through the philosophical ramifications of that commitment.

[33] Compare: "Signs would have little use if they could not serve to designate the same thing repeatedly and in different contexts, while making it easily recognisable that the same thing is intended. This is achieved by using signs that are as similar as possible on the different occasions" (GGA, II, §99). What it amounts to to talk rigorously in terms of signs that are equal-shaped (*gleichgestaltet*) in the way Frege has in mind, can be gathered from his late unfinished piece on logical generality, where we find passages such as: "'If $\mathfrak{a}$ is a man, $\mathfrak{a}$ is mortal' is a sentence of the object-language in which a general thought is expressed. We move from the general to the particular by substituting for the equiform [*gleichgestaltet*] indefinitely indicating letters equiform proper names" (PW, 261).

[34] I have had extensive conversations on the matter with Jim Conant, and he remains convinced, I think, that Frege cannot *really* have believed that signs are object, because such a view cannot possibly account for the sort of repeatability that is characteristic of the sign, a repeatability to which Frege himself constantly appeals – implicitly and explicitly – in his talk about signs, as indeed anyone must. Conant wishes to say, if I have understood him well, that the whole discussion of equal-shapedness in the *Grundgesetze* – the core paragraphs are §99-§100, for those whose appetite has been whetted – is to be read as conducted in a sort of dialectical mode, in which Frege is trying to bring out how it is the *formalist* who is committed to such a first-level relation of equal-shapedness – because *they* have nothing but mere figures to work with, which *are* objects – and where this is taken to further confirm the absurdity of formalism. From my side, I simply cannot see how to understand Frege's discussion of equal-shapedness as doing anything else than laying out his own conception of the sign. Indeed, I would claim that Frege's notion of equal-shapedness is precisely *meant* to account for the repeatability of the sign, *given* that signs are conceived as objects, and that the fact that it involves the notion of signifying (*bezeichnen*) is meant to distinguish it from anything that the formalist has at their disposal. Note, in this regard, that Frege proceeds to adumbrate a *different* notion of equal-shapedness for formalist figures (GGA, II, §100). That the first notion nevertheless fails to account for the repeatability of the sign, I take to show that Frege failed to fully grasp the philosophical ramifications of resolutely thinking through the internality of its use to the sign. I could discuss more extensively my understanding of my own discussions with Conant on the matter – and my reasons for taking the view I have just adumbrated – but the upshot would remain the same: I do not know how to properly deal with the matter, so I have decided to leave things here, and mark this as an issue for further research.

Begriffsschrift. This is what is required for a proper conception of the internality of its use to the sign, and thereby of the internality of the compositionality of the *Begriffsschrift* and the functionality of the realm of *Bedeutungen*. If signs are merely a subcategory of objects, however, *none* of these is *sui generis*. What we have here, then, is another point at which certain fundamental limitations of Frege's philosophical logic come to the fore.[35] I can do no more than flag this issue as a topic for further investigation. In what follows, I will leave it aside, as I have already done in what came before. Let us return, then, to my attempt to provide what I take to be a Fregean account of the internality of the compositionality of the *Begriffsschrift* and the functionality of the realm of *Bedeutungen*, keeping in mind that a full discussion of the topic would require a proper confrontation with the problem that I have just delineated.

5.7 Fregean showing

In Chapter 2, we discussed Jolley's reading of Frege, according to which there is no such thing as making judgments about functions, and no such thing as designating functions. Jolley criticized Frege's use of the notion of *Bedeutung* as waffly (Jolley, 2015, 124n). In so rejecting that use, Jolley rejects the functionality of the realm of *Bedeutungen* itself. The idea that there is no such thing as making judgments about functions, when resolute thought through, *becomes* the idea that there is no such thing as an application of higher-level functions to lower-level arguments, since for a judgment to be about functions just *is* for there to be such applications. We cannot properly understand Frege's conception of higher-level subsumptions unless we understand how the functionality of the *Begriffsschrift* comes with the logical stratification of such logico-philosophical phenomena as aboutness and *Bedeutung*, a logical stratification that is essentially tied to exactly the sort of logically stratified elucidatory use that Jolley rejects.

There is a sense in which Geach's view stands to the compositionality of the *Begriffsschrift* as Jolley's view stands to the functionality of the realm of *Bedeutungen*. Jolley refuses to acknowledge the idea of a function being the logical subject of a judgment.

[35] Again, I believe that these are limitations that Wittgenstein sought to overcome in the *Tractatus*.

Geach refuses to acknowledge the idea of a function name occurring in the expression of judgment. Jolley's view makes it impossible to properly make sense of the idea of higher-level functions applying to lower-level functions, thereby undermining the very functionality of the realm of *Bedeutungen* itself. Geach's view makes it impossible to make sense of the way in which function names compose higher-level *Begriffsschrift* judgments, thereby undermining the very compositionality of the *Begriffsschrift* itself. Jolley's view cannot properly account for the logico-philosophical unity of the realm of *Bedeutungen*, since it excludes as 'waffly' Frege's logically stratified conception of the function/argument nexus that constitutes the functionality of the realm of *Bedeutungen*. Geach's view cannot account for the logico-philosophical unity of the *Begriffsschrift*, since it excludes as 'waffly' Frege's logically stratified conception of the composition of *Begriffsschrift* judgments that constitutes the compositionality of the *Begriffsschrift*. Although neither discusses very explicitly the matters on which the other focuses, we can expect that Jolley and Geach would each consider the views of the other as nicely complementing their own. Their confusions are two sides of the same coin, in a way that – with a beautiful irony – itself brings out the internality of the compositionality of the *Begriffsschrift* and the functionality of the realm of *Bedeutungen*.

Let us discuss Geach's view a bit further. On that view, the logical stratification of the *Begriffsschrift* cannot be accounted for in terms of how *Begriffsschrift* signs fit together in *Begriffsschrift* judgments. Rather, its logical stratification arises through the ways in which *Begriffsschrift* expressions result from applications of sign-functions. But these sign-functions are not themselves *Begriffsschrift* expressions: they apply externally to such expressions, and are constituted logically independently of them. Considered in themselves, *Begriffsschrift* expressions simply have no unity, on Geach's view. That unity is imposed externally through the application of sign-functions. It is only by seeing signs as resulting from such applications, that we can talk about logical distinctions between signs at all, on Geach's views. Considered in themselves, the signs '$F(a)$' and '$a = b$' are logically on a par. Insofar as they are considered as presenting us with different categories of signs, the source of this difference lies in their resulting from the application of different categories of sign-functions that are external to the signs themselves. This is why, as I said above, it is not really proper to say that, on Geach's view, the logical stratification of the *Begriffsschrift is* the logical stratification of the realm of *Bedeutungen*. It seems better to say that there is, for Geach, *no such thing* as the logical stratification of the *Begriffsschrift*, insofar as there is no such thing as different categories of signs. There

are only different categories of sign-functions, and it is only through their application that signs may come to be seen as different, a difference that is entirely external to the signs themselves.

In this way, insofar as sense can still be made of the idea that the *Begriffsschrift* is logically stratified at all, its logical stratification comes to be entirely parasitic on the logically prior logical stratification of the sign-functions. In this way, Geach's account inevitably falls back into the regarding the logical stratification of the realm of *Bedeutungen* as the logically prior phenomenon, and thereby into a form of realism. For Geach, an understanding of the logical stratification of the *Begriffsschrift* depends on a logically prior understanding of the stratification of the sign-functions.

Since Geach's view amounts to a form of realism, we should expect him to fall back into naive ineffabilism, according to which the logical category distinctions are ineffable truths that are conveyed through failed attempts to state what cannot be said, but which in their failure nevertheless succeed in conveying these ineffable truths after all. And this is indeed what we find. At the very beginning of his essay, Geach writes:

> Wittgenstein holds that various features of reality come out, *sich zeigen*, in our language, but we cannot use this language to *say*, assert, that reality has these features: if we try to frame propositions ascribing these features to reality, then it will be possible to show that strictly speaking these are not propositions, only sentence-like structures which violate the principle of logical syntax and are thus devoid of any sense, true or false. All the same, these nonsensical (*unsinnig*) structures may be useful; they may serve to convey from speaker to hearer an insight that cannot be put into proper propositions (Geach, 1976, 54).

Geach wishes to argue that Frege should be seen as a precursor to Wittgenstein, so that Frege's elucidations are to be conceived as such 'sentence-like structures' that serve to convey the requisite sort of insights that cannot be put into proper propositions (Geach, 1976, 69). In targeting readers such as Black, he writes:

> The category-distinctions in question are features both of verbal expressions and also of the reality our language is describing; in consequence, the manoeuvre of 'semantic ascent' – transformation of talk about things in the world into talk about expressions in a language – is in principle entirely futile as an attempt to [...] remove the difficulty about unsayables (Geach, 1976, 55).

In targeting the idea of transforming talk about the world into talk about language, Geach is targeting Frege's own elucidatory turn to language. On his reading, such a transformation cannot contribute in any way to the elucidation of the logical categories. We already know why this is so: on Geach's conception, the logical stratification of the realm of *Bedeutungen* is logically prior, and the logical stratification of the *Begriffsschrift* arises from the external application of sign-functions. It is the logical stratification of the realm of *Bedeutungen* that we must come to understand, and we come to understand it through elucidations, which ineffably convey the requisite insights. There is now, however, nothing left but to think of such elucidations along the lines of naïve ineffabilism: as failed attempts to state what cannot be said, but which nevertheless succeed in conveying it. In this way, Frege's elucidations again come to be modelled on judgment in the way discussed in the first chapter. They are failed attempts to do what judgments do, and it is *in* their failure that they nevertheless succeed in conveying what they cannot say. Thus, Geach is brought to say things like: "none of these sentences are both syntactically well-formed and semantically supplied with *Bedeutungen* for the expressions employed,[36] but they may *nevertheless* succeed in conveying insights" [my emphasis] (Geach, 1976, 69). The 'nevertheless' says it all: it is a way – as I put it in the first chapter – to take back with one hand what one claims to have given away with the other.

At least two objections could be raised against my criticism of Geach's account. The first is that – although it is true that Geach says things such as that elucidations convey insights that cannot be put into proper propositions – he also says things such as the following:[37]

> All the same, there is a test for these sentences' having conveyed the intended distinctions – namely, that by their aid mastery of the formalized language is attainable (Geach, 1976, 55).

Thus, Geach gives us what may be called an effable operationalization of the conveying of ineffable insights that is the business of elucidations: saying that those insights have

[36] Geach's invocation of syntax/semantics in this way further suggests that he is indeed working within the same sort of externalist framework about which I have been at pains to argue that it cannot provide us with an adequate understanding of Frege's philosophical logic.

[37] Compare: "The business of function and argument and value cannot be shelved by talking about expressions of different category, because it reappears if we consider, as we must, the ways of forming expressions out of expressions. No proper explanation is possible; the matter can only be presented and talked around, in the hope that some people will catch on; that they have caught on is shown by their command of simple logical apparatus" (Geach, 1975, 143).

been conveyed to someone, is in effect to say that they have mastered the *Begriffsschrift*. In this way, it may seem that we need not revert to ineffable truths to account for the elucidatory function of Frege's elucidations.

This sort of idea – that the success of Frege's elucidations lies in them bringing someone to a mastery of the *Begriffsschrift* – is common in the Frege literature.[38] The problem, however, is that we are never given a clear explanation of what this criterion is supposed to amount to. What does it mean, exactly, for someone to have mastered the *Begriffsschrift*?[39] Perhaps it means that one can reliably produce valid proofs in the *Begriffsschrift* notation. In that case, however, philosophers such as Bertrand Russell and Crispin Wright should be acknowledged as having mastered the *Begriffsschrift*, and therefore as having attained the ineffable insights that Frege's elucidations are meant to convey. But Frege himself makes it clear in their correspondence that he does not think that Russell has understood his elucidations, and readers such as Geach would wish to claim that Crispin Wright has not understood them either. So there must be more to an understanding of Frege's elucidations than the capacity to produce valid proofs in the *Begriffsschrift*. But here we immediately run into trouble, since this 'something more' will inevitably involve one's logico-philosophical understanding *of* the *Begriffsschrift*, an understanding that – on the sort of account advanced by Geach – inevitably brings us back to ineffable truths. In this way, the relevant notion of mastery itself turns out to be ineffable, so that it cannot move us away from naïve ineffabilism.

The second objection is that my discussion has not properly taken into account Geach's claim that the requisite features of reality – as he calls them – *come out* in language, show themselves in the *Begriffsschrift*. This, it could be said, gives us a way in which those features are conveyed that does not reduce to failed attempts to do what judgments do. It is true, I think, that *a* notion of showing is crucial to the internality of compositionality and functionality that is under study in this chapter. I also believe, however, that Geach's view does not yield the right notion.

The idea is that the logical stratification of the realm of *Bedeutungen* shows itself in the logical stratification of the *Begriffsschrift*. We must recall, however, that on Geach's view, the logical stratification of the *Begriffsschrift* is externally imposed on it through the

[38] Jolley, for instance, says that I understand Frege's elucidations if I can "properly symbolize" (Jolley, 2015, 120). Ricketts talks about Frege "instructing us in the use of the *Begriffsschrift*" (Ricketts, 1986, 87).
[39] Or for someone to be able to properly symbolize?

application of sign-functions. Insofar, then, as the logical category distinctions are shown in the *Begriffsschrift* at all, they are shown in the distinctions between proper names and different categories of sign-functions. But this just means that we are thrown back on the logical stratification of the realm of *Bedeutungen* itself, since the latter distinctions just *are* the logical category distinctions. Geach's idea of showing comes down to the empty idea that the logical stratification of the realm of *Bedeutungen* is shown in the logical stratification of the realm of *Bedeutungen*. I take it to be no coincidence, then, that Geach nowhere connects this notion of showing to elucidations themselves. Geach never relates the way in which elucidations are meant to convey ineffable insights into unsayable features of reality, on the one hand, and the way in which those features show forth in the *Begriffsschrift*, on the other hand. It's as if we can pick and choose: either we see what is shown in the *Begriffsschrift*, or we grasp what elucidations convey.[40]

I wish to end by indicating how, I believe, my own reading of Frege gives rise to a more fruitful notion of showing.[41] To say that the functionality of the realm of *Bedeutungen* shows itself in the compositionality of the *Begriffsschrift*, just is another way of characterizing their internality. The logico-philosophical phenomenon of showing that is at issue *is* such that what shows itself and that in which it shows itself *must* be internally related in this way. It is only through such a form of internality that the showing itself of what shows itself in that in which it shows itself is possible at all. On Geach's view, the way in which the unsayable features of reality show forth in the *Begriffsschrift* stands completely detached from the way in which elucidations purportedly convey those same features. On my reading, to elucidate the internality of compositionality and

[40] This gives us a sense in which Hacker's sort of reading of the *Tractatus* tries to do better. Hacker says that elucidations are meant to "guide the attentive reader to apprehend what is shown by other propositions" (Hacker, 1986, 18). I say 'tries to do better', because he cannot avoid falling into the same sort of view as Geach in the end. See Chapter 2 of my BPhil dissertation (Vanrie, 2017) for some relevant discussion.

[41] My account is indebted to Narboux's fascinating discussion of the Tractarian notion of showing (Narboux, 2014). An anonymous referee of (Vanrie, 2020) summarized its account of Frege in a way that can be paraphrased as follows: the logical category distinctions are not such as to be 'said' by propositions, but show themselves in the structure of the *Begriffsschrift*. I hope that my discussion brings out how much of the work lies precisely in acquiring a proper understanding of the relevant notion of showing. (To be sure, I was myself wholly unclear about this at the time). It is often taken for granted that we understand what is meant in saying that the logical category distinctions show themselves in language – or something similar – whereas I would argue instead that one main task of the *Tractatus* is precisely to achieve clarity about the relevant notion of showing, a point that I take to be confirmed by Narboux's discussion. Perhaps there is a larger methodological point to be made here: often, a philosopher's use of a notion that figures centrally in his or her philosophy must be read as constituting a sustained attempt to achieve clarity *about* that notion, rather than constituting a mere application of a notion that is already well understood. Note, finally, that I am not committing myself to the claim that the Fregean notion of showing that I am about to delineate coincides with the Tractarian notion of showing. From what has already been said, it follows that it does not.

functionality, *is* to bring out the showing itself of functionality in compositionality, and thereby *is* to elucidate the relevant logico-philosophical phenomenon of showing.

It is crucial that that which shows itself is not identical to that in which it shows itself, on pains of being left with a notion of showing that loses all elucidatory force, which is really to be left without a notion of showing at all. We may put the point by saying that the notion of showing must be ampliative. At the same time, that it must be ampliative does not mean that that which shows itself is logically external to that in which it shows itself. In that case, we would not have a proper notion of showing either. What a properly logico-philosophical notion of showing involves, are two logico-philosophical phenomena that are internally related, but nevertheless not the same. In other words, it involves exactly what I have called the internal heterogeneity of compositionality and functionality. It is precisely *in and through* the internal heterogeneity of compositionality and functionality that the latter shows itself in the former. We have here, in the end, one and the same logico-philosophical phenomenon. The internal heterogeneity of compositionality and functionality, the showing itself of functionality in compositionality, this is Frege's conception of the logical categories.

Conclusion

In the course of this dissertation, we have encountered three principal interlocutors of Frege: Kerry, Russell, and Hilbert. Each is seen, from the vantage point of Frege, as embroiled in a confusion that betrays a misunderstanding of the nature of the logical categories. Kerry believed that Frege's concept/object distinction precludes the possibility of making judgments about concepts. In this way, Kerry betrays that he is wedded to a version of the old subject/predicate logic, in which the notions of a logical subject – that which a judgment can be about – and an object are amalgamated. But Frege's conception of the logical categories is such that this amalgamation is broken through the device of higher-level subsumptions, in which functions can themselves appear as logical subjects. In this way, Frege's philosophical logic yields a logically stratified notion of subjecthood that goes together with the logical stratification of the realm of *Bedeutungen* as such, and with a whole cluster of other logically stratified logico-philosophical phenomena.

Russell believed that Frege's function/object distinction is incoherent because it cannot be expressed in judgment. Frege's statements such as 'A function never takes the place of an object' are taken by Russell to be self-undermining, since 'a never takes the place of an object' has a taking the place of an object. In this way, Russell betrays that he is wedded to a conception according to which the proper business of philosophical logic is to present judgments that express logical truths. But Frege's conception of philosophical logic is such that there is a sharp distinction between the activity of judgment, on the one hand, and the activity of elucidation, on the other hand. Rather than constituting a subject matter for judgment, the function/object distinction is a subject matter for elucidation. Elucidating it, moreover, does not consist in failed attempts to state it, but rather proceeds through an elucidatory turn to language, so that an understanding of the function/object distinction must go through an understanding

of statements such as 'A function name never takes the place of a proper name', which concern the logical stratification of the *Begriffsschrift* rather than the realm of *Bedeutungen*.

Hilbert believed that his axioms are mere marks on paper that can be assigned *Bedeutung* through interpretation in a multitude of ways. This, he thinks, allows him to circumvent questions about the truth of his axioms and about the *Bedeutung* of his signs. In this way, Hilbert betrays that he considers the sign to be constituted logically independently of its use in the expression of judgment. But Frege's conception of the sign is such that its use is *internal* to the sign, so that nothing that does not by its very nature have the purpose of signifying is properly called a sign. Rather than presenting us with genuine axioms, Hilbert has presented us with empty husks. In thinking that he can arrive at the meaningful sign by starting from the mere figure, Hilbert neglects the logical priority of the full-blooded case of a sign that is used in the expression of judgment, which is such that the order of the mere figure has no role to play in a properly logical conception of the sign.

In the *Grundlagen*, Frege presents his famous three interconnected methodological principles: never to lose sight of the distinction between concept and object, never to ask for the *Bedeutung* of a word in isolation, and always to separate sharply the psychological from the logical (GL, x). To violate one is to violate them all. On the basis of Frege's dialectic with Kerry, Russell, and Hilbert, I wish to propose three additional – equally interconnected – such principles:[1]

(1) Never to lose sight of the phenomenon of logical stratification.
(2) Always to clearly distinguish between elucidation and judgment.
(3) Always to separate sharply the figure from the sign.

If (1) is not observed, the possibility of making judgments about functions has to be relinquished, leading to a blurring of judgment and elucidation through a blurring of sentences such as 'The concept *horse* is realized' and 'Concepts are fundamentally

[1] Kimhi has claimed that we must regard Frege's philosophical logic as, in the end, violating all three of his own original principles (Kimhi, 2018, 45ff.). This claim could be made of my own three new principles as well. We have seen that Frege has no workable conception of the *elucidatory use* of signs, because the nature of the sign is fundamentally wedded to its use in the expression of *judgment*. This, in effect, constitutes a violation of (2). We also saw that Frege seems to think of all signs as objects, which reveals that his notion of the sign is still invested with certain vestiges of the notion of the figure, yielding a violation of (3). The result is that the logical stratification of the *Begriffsschrift* is conceived as consisting in mere distinctions between species of objects, and thereby not as a properly *logical* stratification, which constitutes a violation of (1).

different from objects', thereby violating (2). Moreover, this prevents an adequate understanding of the function name as a *sign* that has the purpose of *bedeuten* a function in the expression of judgment, so that properly adhering to (3) becomes impossible.

If (2) is not observed, Frege's logically stratified elucidatory use of his logico-philosophical terms of art will be regarded as constituting a logically defective attempt at judgment, preventing a proper recognition of the logical stratification of phenomena such as aboutness and *Bedeutung*, thereby violating (1). The same goes for Frege's use of his notion of *fit*, which makes impossible a proper understanding of the logical stratification of the *Begriffsschrift*, so that (3) will also be violated.

If (3) is not observed, Frege's distinction between proper names and function names will appear as a distinction between mere figures, rendering completely mysterious the idea that function names are essentially unsaturated. One option is then to adopt a reading such as Geach's, according to which function names *are* functions, which falls into naïve ineffabilism, thereby violating (2). Or, alternatively, one will adopt a reading such as Wright's, according to which functions are no longer essentially unsaturated but constitute a subspecies of objects, so that (1) is abandoned.

In the Introduction, I presented Frege's predicament: at first sight, Frege's conception of the logical categories seems to be such that the elucidatory statements through which Frege purports to articulate that conception, turn out to be self-undermining. A facile way out of this predicament – which, as we have seen, is really only the illusion of a way out – is naïve ineffabilism. Frege, however, was not a naïve ineffabilist. Frege's elucidations are not to be conceived as failed attempts to state what cannot be expressed in judgment. To see how Frege sought to overcome his predicament, is to see how (1)-(3), if properly adhered to, yield the internal heterogeneity of compositionality and functionality *as* Frege's conception of the logical categories, in the way that I have sought to bring out in this dissertation. It is a conception on which, to be sure, the logical categories are ineffable, in the sense that this internal heterogeneity is not a subject matter for judgment. But it does not fall into *naïve* ineffabilism, since the elucidation of that internal heterogeneity is not modelled on judgment at all.

It may be said that my account is not satisfactory, since it ends up saying nothing more than that elucidations elucidate by elucidating what they elucidate, without giving us an account of what – if it is not what naïve ineffabilism takes it to be – such elucidation *is*. There is, however, no non-circular answer to be given to that question. What Frege's elucidations do, is to elucidate, and to appreciate the *sui generis* nature of elucidation is to

appreciate that this is the only answer that can be given. All that one can do, is to try to unfold that activity of elucidation, and thereby to bring into view its constitutive moments. That is what I have tried to do, and it has brought me to my characterization of the elucidation of the logical categories as the activity of bringing out – through a logically stratified elucidatory use of certain fundamental logico-philosophical terms of art and through an elucidatory turn to language – the internal heterogeneity of the compositionality of the *Begriffsschrift* and the functionality of the realm of *Bedeutungen*. That internal heterogeneity, I wish to say, is the fundamental logico-philosophical phenomenon from which all else flows in Frege's philosophical logic. To elucidate the logical categories, is to allow this phenomenon to show itself from itself, through the showing itself of functionality in compositionality. To understand this, is to understand Frege's way out of his predicament.

References

Bar-Elli, G. (2006). Identity in Frege's Begriffsschrift. *Canadian Journal of Philosophy, 36*(3), 355-370.

Beaney, M. (1996). *Frege: Making Sense*. London: Duckworth.

Bengtsson, G. (2018). Frege on Dichtung and Elucidation. In G. Bengtsson, S. Säätela, & A. Pichler (Eds.), *New Essays on Frege* (pp. 101-118). Cham, Switzerland: Springer.

Black, M. (1964). A Companion to Wittgenstein's Tractatus. Ithaca, N.Y.: Cornell University Press.

Black, M. (1968). Frege on Functions. In E. D. Klemke (Ed.), *Essays on Frege* (pp. 223-248). Chicago: University of Illinois Press.

Boghossian, P. (2008). Epistemic Rules. *The Journal of Philosophy, 105*(9), 472-500.

Bronzo, S. (2017). Wittgenstein, Theories of Meaning, and Linguistic Disjunctivism. *European Journal of Philosophy, 25*(4), 1340-1363.

Burge, T. (2005). *Truth, Thought, Reason*. Oxford: Oxford University Press.

Cavell, S. (2002). *Must We Mean What We Say?* (Updated ed.). Cambridge: Cambridge University Press.

Conant, J. (2002). The Method of the *Tractatus*. In E. H. Reck (Ed.), *From Frege to Wittgenstein* (pp. 374-462). Oxford: Oxford University Press.

Conant, J. (2020). Replies. In S. Miguens (Ed.), *The Logical Alien* (pp. 321-1028). Cambridge, MA: Harvard University Press.

Conant, J. (2020b). Wittgenstein's Critique of the Additive Conception of Language. *Nordic Wittgenstein Review, 9*, 7-36.

Davidson, D. (1979). Quotation. *Theory and Decision, 11*, 27-40.

Diamond, C. (1981). What Nonsense Might Be. *Philosophy, 56*(215), 5-22.

Diamond, C. (1984). What Does a Concept Script Do? *The Philosophical Quarterly, 34*(136), 343-368.

Diamond, C. (2010). Inheriting from Frege: the work of reception, as Wittgenstein did it. In M. Potter & T. Ricketts (Eds.), *The Cambridge Companion to Frege* (pp. 550-601). Cambridge: Cambridge University Press.

Dummett, M. (1981). *Frege: Philosophy of Language* (2 ed.). London: Duckworth.

Evans, G. (1982). *The Varieties of Reference* (J. McDowell Ed.). Oxford: Clarendon Press.

Frege, G. (1967). *Kleine Schriften* (I. Angelelli Ed.). Hildesheim: Georg Olm.

Frege, G. (1968 [1884]). *The Foundations of Arithmetic* (J. L. Austin, Trans.). Oxford: Basil Blackwell.

Frege, G. (1970 [1879]). Begriffsschrift, a formula language, modeled upon that of arithmetic, for pure thought. In J. van Heijenoort (Ed.), *Frege and Gödel: two fundamental texts in mathematical logic* (pp. 1-82). Cambridge (Mass.): Harvard University Press.

Frege, G. (1976). Wissenschaftlicher Briefwechsel (G. Gabriel, H. Hermes, F. Kambartel, C. Thiel, & A. Veraart Eds.). Hamburg: Felix Meiner Verlag.

Frege, G. (1979). *Posthumous Writings* (P. Long & R. White, Trans. H. Hermes, F. Kambartel, & F. Kaulbach Eds.). Oxford: Basil Blackwell.

Frege, G. (1980). *Philosophical and Mathematical Correspondence* (H. Kaal, Trans. G. Gabriel, H. Hermes, F. Kambartel, C. Thiel, A. Veraart, & B. McGuinness Eds.). Chicago: The University of Chicago Press.
Frege, G. (1984). *Collected Papers on Mathematics, Logic, and Philosophy* (M. Black, V. H. Dudman, P. Geach, H. Kaal, E.-H. W. Kluge, B. F. McGuinness, & R. H. Stoothoff, Trans. B. McGuinness Ed.). New York: Basil Blackwell.
Frege, G. (1983). Nachgelassene Schriften (H. Hermes, F. Kambartel, & F. Kaulbach Eds. 2 ed.). Hamburg: Felix Meiner Verlag.
Frege, G. (2004 [1910-1914]). *Frege's Lectures on Logic* (E. H. Reck & S. Awodey, Trans. E. H. Reck, S. Awodey, & G. Gabriel Eds.). Chicago, Illinois: Open Court.
Frege, G. (2013 [1893/1903]). *Basic Laws of Arithmetic* (P. A. Ebert, M. Rossberg, & C. Wright, Trans. P. A. Ebert, M. Rossberg, & C. Wright Eds.). Oxford: Oxford University Press.
Geach, P. (1961). Frege. In G. E. M. Anscombe & P. Geach (Eds.), *Three Philosophers* (pp. 127-162). Oxford: Basil Blackwell.
Geach, P. (1975). Names and Identity. In S. Guttenplan (Ed.), Mind and Language (pp. 139-158). Oxford: Clarendon Press.
Geach, P. (1976). Saying and Showing in Frege and Wittgenstein. In J. Hintikka (Ed.), *Essays in Honor of G. H. von Wright* (pp. 54-70).
Goldfarb, W. (2001). Frege's conception of logic. In J. Floyd & S. Shieh (Eds.), *Future Pasts* (pp. 25-42). Oxford: Oxford University Press.
Gustafsson, M. (2020). Wittgenstein on Using Language and Playing Chess. In S. Miguens (Ed.), *The Logical Alien* (pp. 202-221). Cambridge, Massachusetts: Harvard University Press.
Hacker, P. M. S. (1986). *Insight and Illusion* (Revised ed.). Oxford: Clarendon Press.
Hacker, P. M. S. (1999). Naming, Thinking and Meaning in the *Tractatus*. *Philosophical Investigations, 22*(2), 119-135.
Hale, B., & Wright, C. (2012). Horse Sense. *The Journal of Philosophy, 109*(1/2), 85-131.
Heck, Richard Kimberley, & May, R. (2013). The function is unsaturated. In M. Beaney (Ed.), *The Oxford Handbook of The History of Analytical Philosophy* (pp. 825-850). Oxford: Oxford University Press. (Originally published under the name 'Richard G. Heck, Jr')
Heck, Richard Kimberley. (2010). Frege and semantics. In M. Potter & T. Ricketts (Eds.), *The Cambridge Companion to Frege* (pp. 342-378). Cambridge: Cambridge University Press. (Originally published under the name 'Richard Heck')
Hugly, P. (1973). Ineffability in Frege's logic. *Philosophical Studies, 24*(4), 227-244.
Johnston, C. (2007). Symbols in Wittgenstein's *Tractatus*. *European Journal of Philosophy, 15*(3), 367-394.
Johnston, C. (2017). Frege on Syntax, Ontology, and Truth's Pride of Place. *European Journal of Philosophy, 26*(1), 569-588.
Jolley, K. D. (2007). *The concept 'horse' paradox and wittgensteinian conceptual investigations*. London and New York: Routledge.
Jolley, K. D. (2015). Once Moore Unto the Breach! Frege and the Concept 'Horse' Paradox. *Philosophical Topics, 43*(1-2), 113-124.
Jones, N. K. (2016). A Higher-Order Solution to the Problem of the Concept Horse. *Ergo, 3*(6), 132-166.
Kimhi, I. (2018). *Thinking and Being*. Cambridge, Massachusetts: Harvard University Press.
Klement, K. C. (2004). Putting Form Before Function: Logical Grammar in Frege, Russell, and Wittgenstein. Philosopher's Imprint, 4(2), 1-47.
Kneale, W., & Kneale, M. (1962). The Development of Logic. Oxford: Clarendon Press.
Kremer, M. (2006). Logicist Responses to Kant: (Early) Frege and (Early) Russell. Philosophical Topics, 34(1-2), 163-188.
Kremer, M. (2010). Sense and reference: the origins and development of the distinction. In M. Potter & T. Ricketts (Eds.), The Cambridge Companion to Frege (pp. 220-292). Cambridge: Cambridge University Press.

Kremer, M. (2019). Definitions in Begriffsschrift and Grundgesetze. In P. A. Ebert & M. Rossberg (Eds.), Essays on Frege's Basic Laws of Arithmetic (pp. 538-566). Oxford: Oxford University Press.
Long, P. (2001). Logic, Form and Grammar. London: Routledge.
Macbeth, D. (2005). Frege's Logic. Cambridge, Massachussets: Harvard University Press.
McDowell, J. (1998). Meaning, Knowledge, and Reality. Cambridge, Massachusetts: Harvard University Press.
Moore, A. W. (2012). *The Evolution of Modern Metaphysics*. Cambridge: Cambridge University Press.
Moore, A. W. (2019). *Language, World, and LImits*. Oxford: Oxford University Press.
Narboux, J.-P. (2014). Showing, the Medium Voice, and the Unity of the *Tractatus*. *Philosophical Topics, 42*(2), 201-262.
Narboux, J.-P., & Perrin, D. (2021). Introduction. In J.-P. Narboux & D. Perrin (Eds.), *New Essays on Frege's Logical Investigations* (pp. 7-42). São Paulo: Editora Nonada.
Noonan, H. (2006). The Concept Horse. In P. F. Strawson & A. Chakrabarti (Eds.), *Universals, Concepts and Qualities: New Essays on the Meaning of Predicates* (pp. 155-176). Burlington, VT: Ashgate.
Parsons, T. (1986). Why Frege Should Not Have Said "The Concept Horse is Not a Concept". *History of Philosophy Quarterly, 3*(4), 449-465.
Picardi, E. (1993). Kerry und frege über begriff und gegenstand. *History and Philosophy of Logic, 15*(1), 9-32.
Proops, I. (2013). What is Frege's 'concept horse problem'? In P. Sullivan & M. Potter (Eds.), *Wittgenstein's Tractatus* (pp. 76-96). Oxford: Oxford University Press.
Quine, W. V. (1934). Ontological remarks on the propositional calculus. *Mind, 43*(172), 472-476.
Ricketts, T. (2010). Concepts, objects and the Context Principle. In M. Potter & T. Ricketts (Eds.), *The Cambridge Companion to Frege* (pp. 149-219). Cambridge: Cambridge University Press.
Ricketts, T. G. (1985). Frege, The Tractatus, and the Logocentric Predicament. *Noûs, 19*(1), 3-15.
Ricketts, T. G. (1986). Objectivity and Objecthood: Frege's Metaphysics of Judgment. In L. Haaparanta & J. Hintikka (Eds.), *Frege Synthesized* (pp. 65-96). Dordrecht: D. Reidel Publishing Company.
Russell, B. (2010 [1903]). *Principles of Mathematics*. London and New York: Routledge.
Sheffer, H. M. (1926). Review of *Principia Mathematica*. *Isis, 8*(1), 226-231.
Stanley, J. (1996). Truth and metatheory in Frege. *Pacific Philosophical Quarterly, 77*, 45-70.
Sullivan, P. (2003). Ineffability and Nonsense. *Proceedings of the Aristotelian Society, Supplementary Volumes, 77*, 195-223.
Sullivan, P. (2006). Metaperspectives and internalism in Frege. In M. Beaney & E. H. Reck (Eds.), *Gottlob Frege: Critical Assessments of Leading Philosophers* (pp. 85-105). London: Routledge.
Textor, M. (2010). Frege's Concept Paradox and the Mirroring Principle. *The Philosophical Quarterly, 60*(238), 126-148.
Travis, C. (2020). Where Words Fail. In S. Miguens (Ed.), *The Logical Alien* (pp. 222-280). Cambridge, Massachussets: Harvard University Press.
Valberg, J. J. (1970). Improper Singular Terms. *Proceedings of the Aristotelian Society, 71*, 121-145.
Van de Vijver, G. (2013). Frege en Lacan: een transcendentale denklijn. *Psychoanalytische Perspectieven, 31*(4), 353-369.
van der Schaar, M. (2018). Frege on Judgment and the Judging Agent. *Mind, 127*(505), 225-250.
Vanrie, W. (2017). Logical syntax and nonsense in Wittgenstein's Tractatus. (BPhil in Philosophy). University of Oxford,
Vanrie, W. (2020). Why did Frege reject the theory of types? British Journal for the History of Philosophy. doi:10.1080/09608788.2020.1807908
Weiner, J. (2005). Semantic Descent. Mind, 114(454), 321-354.
Weiner, J. (2006). On Fregean Elucidation. In M. Beaney & E. H. Reck (Eds.), Gottlob Frege: Critical Assessments of Leading Philosophers (pp. 197-214). London: Routledge.

Weiner, J. (2010). Understanding Frege's Project. In M. Potter & T. Ricketts (Eds.), *The Cambridge Companion to Frege* (pp. 32-62). Cambridge: Cambridge University Press.
Wells, R. S. (1968). Is Frege's concept of a function valid? In E. D. Klemke (Ed.), *Essays on Frege* (pp. 391-406). Chicago: University of Illinois Press.
Witherspoon, E. (2002). Logic and the Inexpressible in Frege and Heidegger. *Journal of the History of Philosophy, 40*(1), 89-113.
Wittgenstein, L. (1981 [1922]). *Tractatus Logico-Philosophicus* (C. K. Ogden, Trans.). London: Routledge.
Wright, C. (1998). Why Frege did not deserve his *granum salis*. *Grazer Philosophische Studien, 55*, 239-263.

Summary

Frege's philosophical logic presents us with a distinction between *concepts* and *objects*. Objects – such as Elon Musk – are *complete*, whereas concepts are unsaturated and require completion by an object. An example is the concept *ξ is the father of X AE A-XII*. The proposition 'Elon Musk is the father of X AE A-XII' involves the completion of this concept by the object Elon Musk. Frege conceives of the distinction between concept and object as a sharp *logical* distinction. Objects are *essentially* complete, and concepts are *essentially* unsaturated. Moreover, they are entirely incommensurable, in the radical sense that any statement about an object becomes entirely meaningless if we try to apply it to a concept, and *vice versa*. In logical terms: a variable either has objects or concepts as values, but never both. All forms of logical generality are limited to either objects, or concepts.

Frege's sharp distinction between concept and object, however, seems to give rise to paradoxes. Take the statement 'A variable either has objects or concepts as values, but never both'. It seems that – in order to say this – we need to introduce exactly the form of logical generality that Frege takes to be impossible, one that encompasses *both* objects and concepts, as shown by the words 'never both'. The result is that Frege's concept/object distinction faces a peculiar philosophical predicament: in the attempt to articulate it, we seem to come out with statements that undermine it. The conclusion that Frege's conception is incoherent, seems inevitable.

My dissertation is a historical-philosophical investigation of the way in which Frege seeks to avert such a conclusion. It is an investigation of Frege's conception of the *logical categories*, of which *object* and *concept* constitute two examples, but not the only ones. The underlying question is: what *is* this radical sort of distinction between logical categories, and how can we give voice to it? My research shows that Frege's philosophical logic essentially involves the idea that logical insight never stands on its own, but always involves the development of a suitable logical notation, such as Frege's own *Begriffsschrift*.

The sort of insights that are at issue in logic, are inherently tied to logical notation *itself*. To understand this, we must understand both Frege's conception of the nature of logical notation, as well as this 'inherent tie'. To understand the first, it is crucial to realize that Frege's philosophical logic cannot be reconciled with the sort of distinction between syntax and semantics that has become commonplace in contemporary logic. To understand the second, we must understand how logical notation and the logical categories exhibit what I call an *internal heterogeneity*: they are intrinsically interconnected, yet not the same, in such a way that it is only *through* their heterogeneous interconnection that we can come to grasp either of them.

Samenvatting

In Frege's filosofische logica treffen we een onderscheid aan tussen *concepten* en *objecten*. Objecten – zoals bijvoorbeeld Elon Musk – zijn volgens Frege *volledig*, terwijl concepten *onverzadigd* zijn en vervollediging vragen door een object. Een voorbeeld is het concept '*ξ is de vader van X AE A-XII*. In de propositie 'Elon Musk is de vader van X AE A-XII' wordt dit concept vervolledigd door het object Elon Musk. Frege vat dit onderscheid tussen concept en object op als een *logisch* onderscheid. Objecten zijn *essentieel* volledig, en concepten zijn *essentieel* onverzadigd. Sterker nog: deze zijn volledig onvergelijkbaar, in de radicale zin dat eender welke uitspraak over een object volstrekt betekenisloos wordt indien we ze trachten toe te passen op een concept, en *vice versa*. In logische termen: een variabele heeft ofwel objecten als waarde, ofwel concepten, maar nooit beide. Elke vorm van logische algemeenheid is beperkt tot hetzij objecten, hetzij concepten.

Frege's radicale onderscheid tussen concept en object lijkt echter tot paradoxen te leiden. Neem de zin 'Een variabele heeft ofwel objecten als waarde, ofwel concepten, maar nooit beide'. Het lijkt erop dat we – om dit te zeggen – precies het soort logische algemeenheid moeten introduceren die volgens Frege onmogelijk is: een algemeenheid die *zowel* objecten als concepten omvat, zoals blijkt uit de woorden 'nooit beide'. Het resultaat is dat Frege's concept/object onderscheid met een eigenaardige filosofische moeilijkheid kampt: in de poging om dat onderscheid te formuleren, belanden we bij uitspraken die dat onderscheid zelf lijken te ondermijnen. De conclusie dat Frege's visie incoherent is, lijkt zo onafwendbaar.

Mijn doctoraat is een historisch-filosofisch onderzoek naar de manier waarop Frege deze conclusie tracht te voorkomen. Het is een onderzoek naar Frege's filosofie van de *logische categorieën*, waarvan *object* en *concept* twee voorbeelden zijn, maar niet de enige. De onderliggende vraag luidt: wat *is* dit soort radicaal onderscheid tussen logische categorieën, en hoe kan zo'n onderscheid kenbaar gemaakt worden? Mijn onderzoek

toont aan dat Frege's filosofische logica essentieel draait rond de gedachte dat de logica zich niet op zichzelf laat denken, maar enkel doorheen het ontwikkelen van een geschikte logische notatie, zoals Frege's *Begriffsschrift* er één tracht te zijn. Het soort inzichten die in de logica aan de orde zijn, zijn inzichten die steeds inherent gekoppeld zijn aan de logische schriftuur *zelf*. Om dit te begrijpen, moeten we zowel Frege's visie op de aard van die logische schriftuur begrijpen, als die 'inherente koppeling'. Om het eerste te begrijpen, is het cruciaal om in te zien dat Frege's filosofische logica incompatibel is met het onderscheid tussen syntax en semantiek dat quasi banaal is geworden voor de hedendaagse logica. Om het tweede te begrijpen, moeten we begrijpen hoe de logische schriftuur en de logische categorieën een *interne heterogeniteit* vertonen: ze zijn intrinsiek verweven met elkaar, maar toch verschillend, op een zodanige manier dat we ze enkel *doorheen* hun heterogene verweving filosofisch kunnen vatten.