Journal for the History of Analytical Philosophy
Volume 13, Number 1 (2024)

Frege's Conceptions of Elucidation

Wim Vanrie
wim.vanrie@ugent.be
Ghent University

ABSTRACT

I argue that discussions of Frege's conception of elucidation have suffered from a conflation of two distinct issues: elucidation of primitive scientific terms, and elucidation of the logical categories. The former seeks to bring us to grasp the *Bedeutung* of terms that stand at the beginning of the chain of definitions of a scientific system. The latter cannot be understood on the model of securing agreement in *Bedeutung* at all. I show how existing discussions of Fregean elucidation insufficiently take this difference into account. I adumbrate what I take to be a more accurate understanding of Fregean elucidation of the logical categories, starting from the observation that Frege, when engaged in such elucidation, consistently reverts to talking about signs. Frege, I argue, takes signs to possess *logical features*, and it is these logical features which his elucidations are meant to help us to grasp. For Frege, the nature of the logical categories lies *in* the signs. I argue that this reveals that Frege's approach to elucidating the logical categories is incompatible with a realist framework according to which there is a logical order of reality that is prior to the logical order of language.

1. Introduction

Frege's conception of the logical categories involves him in a well-known predicament. According to Frege, there is a sharp distinction between concepts and objects. There are objects, which are the *Bedeutung*[1] of proper names such as 'Seabiscuit', and there are concepts, which are the *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

[1] To prevent us from losing our bearing, I shall use Frege's original terms '*Sinn*' and '*Bedeutung*'. I shall also use the term '*Begriffsschrift*' to refer to Frege's fully developed logical notation, as it features in the *Grundgesetze*.

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 are different logical places; in some only objects can stand and not concepts, in others only concepts and not objects. (Frege 1984, 281–82)

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.[2]

Against Frege, one could take the concept/object 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.[3] 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

[2] In focusing on the concept/object distinction, I am following common practice. It should nevertheless be noted, first, that the more fundamental distinction is that between functions—of which concepts are a special case—and objects, and, second, that Frege's conception of function itself ramifies into a hierarchy of logical categories, including higher-level functions that take lower-level functions and/or objects as arguments. What is usually called a concept belongs to the logical category of first-level functions of one argument.

[3] This is how Russell thought about the matter in his *Principles of Mathematics* (Russell [1903] 2010, sec. 49).

ξ is a horse at all! As he puts it in what has come to be one of the most notorious statements in the history of analytical philosophy: "the concept *horse* is not a concept" (Frege 1984, 186).

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 seems that, for such a statement to express what it purports to express, it must do exactly what Frege claims cannot be done: generalize over both concepts and objects at once. 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 generality—the generality of being able to stand in some logical place—that encompasses both objects and concepts.

Here is how 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, 223–24)

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.

In this paper, I wish to align myself with the tradition in Frege scholarship that seeks to resist this conclusion.[4] Such readers believe that we find in Frege's *oeuvre* indications that he was aware of this predicament, and attempted to address it. To what extend he succeeded in doing so, remains an open question, but the minimal assumption is that Frege's conception of the logical categories is not *simply* incoherent in the straightforward way just sketched. Philosophically speaking, the hope is that reflecting on Frege's way of engaging this predicament may help us to understand the nature of the sort of reflections in which we engage, when we reflect on logic.[5]

What is needed, is an account of how statements such as 'Concepts are unsaturated' are supposed to give expression to 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 concept of 'concepts' under a first-level concept of 'unsaturatedness'—as he *would* be in saying, for instance, that oranges are sweet—since such a construal yields the sort of paradoxical results presented above.

It is precisely on this point that Frege himself was not silent. In fact, he systematically made it clear that his discourse about the logical categories cannot be understood on the same model as scientific discourse. 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

[4]Some prominent example are Diamond (1984), Conant (2002), and Weiner (2020). Readers of Frege who I take to accept the negative conclusion include Wright (1998), Moore (2012), and Proops (2013).

[5]My statement here suggests that it is a premise of this tradition that logic is distinct from the sciences—which may appear to be a decidedly un-Fregean commitment. I think it is a premise—although its exact upshot requires further articulation, of course—and I do not think it is un-Fregean. I hope that the remainder of my paper bears out this latter point. A flat-footed way of putting the matter would be to say that—while *doing* logic, i.e. constructing proofs within a scientific system, is a scientific activity—reflecting on what one is thereby doing is not.

> possible to hint at what is meant [*auf das Gemeinte hinzuführen*] and to make it clearer by relating it to what is known. Instead of a definition we must provide elucidations[6] [*Erläuterung*]; here of course we must count on a meeting of minds [*entgegenkommende Verständnis*]. (Frege 1979, 235)
>
> 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*]. (Frege 1984, 292)

What passages such as these reveal—and they appear throughout Frege's *oeuvre*—is that Frege did mark off his discourse about the logical categories as having a distinct status, which he characterizes by using such terms as *elucidation* (*Erläuterung*), *metaphorical* (*bildlich*), *hinting* (*hinzuweisen*), and *meeting of minds* (*entgegenkommende Verständnis*).[7] At the very least, this reveals that Frege was aware of the need for a distinction here.

As is customary, I will stick to the term *elucidation* to characterize Frege's statements through which he seeks to communicate his conception of the logical categories. If we want to take seriously the idea that there may be more to learn from Frege's conception of the logical categories, then we must take seriously Frege's use of this notion of elucidation to mark off his discourse about the logical categories. We must, simply put, seek to understand the nature of Fregean elucidation.

It is to this understanding that I seek to contribute. My main claim is this: attempts to understand Fregean elucidation have suffered from an overly coarse approach to the texts, resulting in a lumping together of distinct issues in a way that renders it impossible to clearly bring into view the phenomenon that is of interest here: the elucidation of the logical categories. I will show that such lumping together has indeed occurred, and indicate how it has yielded distorted accounts of Frege's conception of the logical categories. I end by presenting an outline of the sort of renewed understanding of Frege's elucidations of the logical

[6] The English translation has 'illustrations' here, but 'elucidation' is the more commonly used term.

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

categories that, I believe, becomes possible once these points are taken into account.

2. Elucidation of Primitive Scientific Terms

That the specific term 'elucidation' has become a technical term in the Frege literature, is perhaps in greater part due to the *Tractatus* than to Frege himself. Consider these two passages:

> The meanings of primitive signs can be explained by elucidations. Elucidations are propositions which contain the primitive signs. They can, therefore, only be understood when the meanings of these signs are already known. (Wittgenstein [1922] 1981, 3.263)
> The object of philosophy is the logical clarification of thoughts.
> Philosophy is not a theory but an activity.
> A philosophical work consists essentially of elucidations.
> The result of philosophy is not a number of 'philosophical propositions', but to make propositions clear. (Wittgenstein [1922] 1981, 4.112)

It is, I think, clear that Wittgenstein is not using the term 'elucidation' in the same way in these passages.[8] To put it bluntly: A philosophical work does not consist essentially of propositions which contain primitive signs whose meaning is to be explained.[9]

My first main claim in this paper is that a similar distinction is present in Frege. Frege also uses the notion of elucidation in at least two ways[10]—one having to do with the meaning of primitive signs, the other with philosophical clarification.[11] Whereas this is clear for the *Tractatus*,

[8] A third use appears in the notorious penultimate section 6.54. Since it has no direct analogue in Frege, I leave it aside here.

[9] I shall not argue for this point as a piece of *Tractatus* exegesis. I merely introduce it as a stepping stone to the similar point I want to make about Frege, for which I will argue in detail. Some of my arguments may be adapted to the *Tractatus*. I think it is fair to say that many scholars accept the point for the *Tractatus*, but fail to see that it equally applies to Frege.

[10] It is on these two ways that I will be focusing throughout. These can, I think, be clearly delineated and must—as I shall insist—be kept distinct. I am not, however, committing myself to the claim that there are no other strands of elucidation to be identified in Frege's *oeuvre*.

[11] To be clear, I am not claiming that these are simply the same in Frege and in the *Tractatus*. In this regard, I take my paper to lay the groundwork for a more accurate comparison of how Frege and Wittgenstein seek to put their respective notions of elucidation to philosophical work.

it is a point that has been largely overlooked with regards to Frege. There is a constant danger of confusing the different notions of elucidation that figure in Frege's thought, which is facilitated by the availability of general characterizations of elucidation that seem to capture all cases. Here is one such characterization: elucidation is what is required to secure mutual agreement in cases where definition is unavailable because a certain phenomenon is logically primitive. Because of the availability of such a general characterization—and because Frege is not as clear on these matters as one could hope[12]—interpreters have tended to underestimate or even completely miss the extent to which it obscures the distinct activities of elucidation that are at play in Frege's *oeuvre*. There is not one way in which the different topics for elucidation are such that definition is 'unavailable' for them, not one way in which the relevant phenomena are 'primitive', and not one way in which 'mutual agreement' is to be secured.

Let us, then, investigate Frege's use of the notion of elucidation. 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 scientific terms that stand at the beginning of the chain of definitions of a scientific system. I will call such terms *primitive scientific terms*. My first task is to bring 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.

Frege's conception of the elucidation of primitive scientific terms is articulated most explicitly at three places: a letter to Hilbert (Frege 1980, 36–37), the posthumous text *Logic in Mathematics* (Frege 1979, 207), and Part II of the published paper *Foundations of Geometry* (Frege 1984, 300–302). 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. 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

[12] 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.

Bedeutung.[13] It is because Frege takes mutual agreement in *Bedeutung* to be so 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 of giving definitions 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.

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 [*bezeichnen*] the same thing by the same sign (word). Once the investigators have come to an understanding about the primitive elements and their designations [*Bezeichnungen*], 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 [*Erläuterung*]. (Frege 1984, 300)

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 terms—such as, for instance, Frege's version of set membership '$\xi \cap \zeta$' (Frege [1893–1903] 2013, I, §34)—but are not themselves defined in Frege's system.[14]

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

[13] 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" (Frege 1984, 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*, which must hold in any case.

[14] An example of a non-logical primitive scientific term provided by Frege is that of a 'point' in geometry (Frege 1984, 301).

place by presenting statements in which the terms in question occur,[15] in the hope that they will help the reader latch onto their *Bedeutung*.

Frege continues:

> [Elucidation] 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. (Frege 1984, 300–301)

Once mutual agreement is secured, elucidation has achieved its purpose, and the actual business of science can get underway. The elucidation of primitive scientific terms is a *propadeutic* to their subsequent scientific use, which is logically on a par with the scientific use of any other scientific term, logically primitive or not. The only difference is that mutual agreement in *Bedeutung* is secured through elucidations instead of definitions.

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. (Frege 1984, 301)

It is precisely in this way, moreover—i.e. in the contingency of their success—that elucidations of primitive scientific terms differ from definitions:

[15] "Elucidations will generally be propositions that contain the expression in question, perhaps even several such expressions" (Frege 1984, 301). I say 'in the first place', since it cannot be excluded that other statements—in which the term to be elucidated does not itself appear—may also fulfil an elucidatory function.

> 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. (Frege 1984, 301–2)

Much remains to be said about Frege's conception of the elucidation of primitive scientific terms, which is a topic deserving of its own investigation. My cursory discussion suffices for our purposes, however, since my aim is to bring out how 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 elucidation of primitive scientific are, this core difference will remain intact.

3. 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* remain undefined. Securing mutual agreement, finally, means helping readers latch onto the right *Bedeutung*.

What I wish to argue now, however, is that *none* of the above characteristics of the elucidation of primitive scientific terms apply to 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 has shifted completely.

Let us start from the observation that Frege does indeed regard the logical categories as primitive:[16]

[16]For a similar remark about the notion of an object, see Frege (1984, 147).

> 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 1979, 235)

If logical category terms were primitive in the same way 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. The corresponding 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 '$O(\xi) = (\forall\varphi)\left(\varphi(\xi) \vee \neg\varphi(\xi)\right)$'. Thus, if '$O(\xi)$' had as its *Bedeutung* the logical category of objects, the notion of an object would not be primitive, as Frege claims it is.

More fundamentally, however, if the logical category of objects were 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 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 itself must be called incomplete, in need of supplementation, or 'unsaturated'. And in this respect functions differ fundamentally from [objects].[17] (Frege 1984, 140)

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

[17] Frege has 'numbers' instead of 'objects'—due to the specific context—but it is clear that he would endorse the same statement about objects.

unsaturated then becomes: $(\forall x)(O(x) \supset \neg U(x))$. By the same token, 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 '$\upsilon_\varepsilon\,(\varphi(\varepsilon))$'. We can then try to express the claim that all functions are unsaturated as '$(\forall F)\,(\upsilon_\varepsilon\,(F(\varepsilon)))$'. 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 not succeeded in articulating a difference between functions and objects, something true of objects while false of functions.[18]

It could be suggested that we should use an unequal-levelled relation to express directly that functions differ fundamentally from objects, along the lines of '$(\forall x)(\forall F)\,(x \neq_\varepsilon F(\varepsilon))$', in which the unequal-levelled relation sign '$\xi \neq_\varepsilon \varphi(\varepsilon)$' is meant to designate 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 '$(\forall x)(\forall F)\,(x \neq_\varepsilon F(\varepsilon))$' turns out to be true. But this does not mean that '$(\forall x)(\forall F)\,(x \neq_\varepsilon F(\varepsilon))$' 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.

However one tries to devise *Begriffsschrift* terms to capture Frege's statements about the logical categories, one will always run into the fact that such terms can never fully 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.[19] 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 search the Fregean

[18] Weiner (2020, chap. 3, §II.b) contains a similar argument pertaining to the notion of 'atemporality'.

[19] This elucidatory generality is explored from a different angle in (Vanrie, forthcoming).

realm of *Bedeutungen* for a *Bedeutung* that fits Frege's contrastive use of logical category terms, we see that there is no suitable *Bedeutung* for us to latch onto.[20] Frege himself never introduces a *Begriffsschrift* term that is supposed to stand for a logical category.[21] Given how central the logical categories are to his logic, this would be inexplicable if he actually believed that he could do so. Frege evidently thought 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 absolute 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 terms. The latter *have* a *Bedeutung*, only they 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 a chain of definitions at all. The very idea of definition is out of place with regards to the logical categories, since definitions serve precisely to fix the *Bedeutung* of a term. We may define a predicate '$O(\xi)$' as '$O(\xi) = (\xi = \xi)$', but this does not give us a definition of the logical category of objects. It simply gives us a definition of a first-level concept that subsumes all objects. 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 have on our hands an object, because only an object fits the argument place of the first-level function $\xi = \xi$.[22] 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 ask whether it is identical

[20] This has been observed many times. See 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, sec. 5), Conant (2020, 450), and Weiner (2020, 106).

[21] This has also been observed many times. See Conant (2002, 385), Ricketts (2010, 170, 50f), and Weiner (2010, 59–60). There is one notable exception in the correspondence with Russell. As the discussion in Vanrie (2021) reveals, it does not threaten my claims here.

[22] Compare Diamond (1984, 354ff).

to itself or not to settle this question, since the meaningfulness of the question already presupposes that the concept *horse* is an object.[23]

Some readers of Frege seem to disagree and think that one can provide definitions of logical categories. Sullivan, for instance, seems[24] to take '$(\forall x)(\varphi x = —\varphi x)$'[25] to provide a definition of the category of concepts (Sullivan 2006, 102).[26] 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*. From this vantage point, Sullivan's formula indeed gives us a definition of first-level concepts, now conceived as a species of first-level functions of one argument. But this is no longer a definition of a *logical category*, since first-level concepts no longer constitute a logical category in the logical system of *Grundgesetze*, only first-level functions of one argument do.[27]

My discussion so far gives us an initial impression of the difference between elucidation of primitive scientific terms and elucidation of the

[23]Can one then use the criterion of the *meaningfulness* of 'The concept *horse* is self-identical' instead of its truth? This question brings with it a host of complications that would take us too far afield here. For our current purposes, it is enough to observe that such a criterion does not move us closer to devising *Begriffsschrift* terms that have logical categories as their *Bedeutung*.

[24]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, he seems to be articulating the same sort of distinction between logical category terms and primitive scientific terms as I am. 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", 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 understand what Sullivan is after here. What is something that is at once 'absolutely indefinable' *and* 'straightforwardly definable in a particular formulation of logic'? It is unclear to me how Sullivan takes his definition—of a perfectly straightforward second-level concept that is true of all first-level concepts—to be related to the 'absolutely indefinable' logical category of concepts.

[25]The '—' is the content-stroke.

[26]Similarly, Dummett (1981, 216–17), Beaney (1996, 200–201), and Noonan (2006, 165) all seem to flirt with the idea that one can provide definitions of the logical categories.

[27]Unless this point is kept in mind, confusion is bound to arise when one engages with Frege's talk of concepts as constituting a logical category. The logical status of his earlier notion of concept as constituting a logical category should be understood as on a par with the logical status of his later notion of first-level functions of one argument.

logical categories. I will now show how a failure to appreciate this difference has led scholars into trouble in their attempts to understand Frege's conception of elucidation and the logical categories.

4. Naive 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*. They are not subject matters for judgment. 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 one could call an irresolute way of taking up this point that is prevalent in the literature, but which constitutes a misunderstanding of Frege's philosophical logic.[28]

While accepting that the logical categories do not constitute a subject matter for judgment, irresolute readers of Frege hold on to the idea that elucidations of the logical categories are modelled on the aim of securing agreement in *Bedeutung*. Only, because the logical categories are not *Bedeutungen*, what such elucidations seek to express is ineffable, so that it cannot, in fact, be expressed. As a result, these elucidations 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 try to but inevitably fail to express—thereby securing mutual agreement with regards to the logical categories. I will call such a view *naïve ineffabilism*. On a naïve ineffabilist reading, elucidations are construed as somehow succeeding in conveying ineffable truths about the logical categories—not by stating them—but by ineffably gesturing at them *in and through* their failure to state them.

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

[28] 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.

from there. In ascribing naïve ineffabilism to 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 elucidation of the logical categories as modelled on elucidation of primitive scientific terms, as attempting to do exactly what elucidations of primitive scientific terms do. In conceiving of statements such as 'Concepts differ fundamentally from objects' as conveying ineffable truths that 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* results in a conception of elucidations of the logical categories on which such elucidations end up being remarkably successful in emulating exactly the sort of statement which—on the record—they cannot really be.

All of this, however, is an illusion: this conception of elucidations as conveying ineffable truths, is a philosophical muddle. The idea of elucidations of the logical categories as failed attempts to do what elucidations of primitive scientific terms succeed in doing needs to be left behind. 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 ineffable truths. But this purported notion of success essentially depends on elucidations being such that they are first of all failed attempts—not to *elucidate* at all—but to directly state ineffable truths about the logical categories. It is only by aiming at the stars, that we end up shooting at the moon. 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 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[29] acknowledges that 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, all is well. 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 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*:[30]

> 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)

On the one hand, the logical categories cannot be referred to. On the other hand, we continue talking 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 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.

[29] 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 Conant (2020) and personal conversation.

[30] Similar naïve ineffabilist accounts are to be found in Geach (1976) and Witherspoon (2002).

As one would expect, Conant bases his reading of Frege's conception of elucidation of the logical categories on the passages—discussed above—in which Frege presents his conception of 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 elucidation of primitive scientific terms. Conant fails, however, to realize that this model cannot simply be transposed to the elucidation of the logical categories.

There are plenty of examples of readers who take the passages concerned with elucidation of primitive scientific terms and proceed to apply them to elucidation of the logical categories. Proops, for instance, infers from the fact that elucidation of primitive scientific terms consists 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 the elucidation 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. Proops' argument, however, rests on an unsupported generalization of what Frege says about elucidation of primitive scientific terms. Such examples could be easily multiplied: readers of Frege have systematically failed to distinguish clearly between elucidation of primitive scientific terms and elucidation of the logical categories, taking Frege's remarks about the former to set the framework for understanding the latter.[31]

A passage that is often taken to express a naïve ineffabilist view,[32] is the notorious grain of salt passage, in which Frege is replying to Kerry's objections to his distinction between concepts and objects:

[31] As we will see below, Weiner is also guilty in this in a more subtle way, by uncritically transposing the notion of a *propaedeutic* to the elucidation of the logical categories.

[32] See e.g., Conant (2002, 392), Witherspoon (2002, 108–9), Weiner (2006, 202), Ricketts (2010, 193), Hale & Wright (2012, 91–92), and Travis (2020, 222–23).

> 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. (Frege 1984, 193)

Two aspects of the passage are pertinent. The first is Frege's talk about his expressions missing his thought. 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. The second aspect is Frege's insistence that language is to blame for the difficulty. 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.

The reason why the grain of salt passage nevertheless does not express a naïve ineffabilist conception of the logical categories, is that it is not concerned with the elucidation of the logical categories at all. 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—an obstacle to talking about the logical categories as such—but an obstacle to making judgments about specific concepts, such as Kerry's own statement that the concept *horse* is a concept easily attained.[33] To avoid the complications arising from the notion of being easily attained, let us consider a simpler statement, such as 'The concept *horse* is realized'. In the *Begriffsschrift*, this is straightforwardly expressed by '$(\exists x)\,(x \text{ is a horse})$'. 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 a statement such as 'The concept *horse* is realized' is meant to express—namely, a second-level subsumption—but which it fails to properly express due to the logical deficiency of the grammar of ordinary language. Taken

[33]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.

in accordance with the grammar of ordinary language, the sentence 'The concept *horse* is realized' expresses a first-level subsumption. Thus, while intending to mention a concept, Frege ends up mentioning an object—exactly as he says. If we grant Frege his grain of salt, however, we will recognize that Frege aims to express a second-level subsumption, and we will understand him accordingly. This does not require any recourse to ineffable truths.[34]

Although most readers agree (as Conant does) that naïve ineffabilism is philosophically hopeless,[35] 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.

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. 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 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 conclude that he is always talking about roughly the same activity.[36] Nevertheless, an exegetical assumption is made here, which can be questioned. Once one does question it, one

[34] Thus, the issue of making judgments about concepts must also be clearly distinguished from the elucidation of the logical categories. For a more extensive discussion of that issue, and a more extensive defence of this reading of the grain of salt passage (and similar passages from Frege's *oeuvre*), see (Vanrie, forthcoming).

[35] See Diamond (1984, 367), Long (2001, 101), Conant (2002), Witherspoon (2002, 111ff), Johnston (2007, sec. 2.2), Hale & Wright (2012, 100ff), and Moore (2012, chap. 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).

[36] It is, of course, no coincidence that Frege uses the same term. It possesses the sort of general unity captured by the general characterization that was given above. But such a general characterization obscures the real differences between the multiple activities of elucidation at issue in Frege's *oeuvre*.

starts to appreciate that Frege's approach to the elucidation of the logical categories is quite different from his approach to the elucidation of primitive scientific terms. Before looking into that, however, I turn to another prominent interpretation of Fregean elucidation.

5. Weiner on Elucidation

One scholar who has discussed Frege's conception of elucidation extensively in her work, is Joan Weiner. Weiner constitutes an exception insofar as she clearly registers the difference between the two kinds of elucidation, pointing out that Frege's statements about the logical categories "are not designed to introduce primitive terms" (Weiner 2020, 106), and that "Frege never claims . . . that elucidation is confined exclusively to attempts to introduce primitive terms" (Weiner 2020, 107). Still, her discussions of elucidation fail to fully cash out this difference, as I will now try to explain.

Weiner characterizes Fregean elucidation in terms of the notion of a *propadeutic*:

> The character of the discussions and arguments in Frege's writings that play elucidatory roles varies dramatically . . . . 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)

This notion of propaedeutic is based on a distinction between what does and what does not belong to the logical theory:

> [An elucidation is] a statement or argument that, while it does not actually belong to the logical theory, it is useful for getting the reader to understand something that does belong to the logical theory. (Weiner 2020, 108)

As before, however, I wish to emphasize that such general characterizations leave the difference between elucidations of primitive scientific terms and elucidations of the logical categories unaccounted for. The senses in which these 'do not belong to the logical theory' and in which they are 'useful for getting the reader to understand something' are entirely different. If we insist on understanding Fregean elucidation in terms of the notion of a propaedeutic, then this notion must be understood differently depending on the sort of elucidation that is at hand.

It is a notion that, I think, is most fitting for elucidations of primitive scientific terms. Here, there is a clear sense in which elucidations are preparatory for the actual business of science: in helping us understand the *Bedeutungen* of primitive scientific terms, they prepare us for the actual scientific use of those terms. The elucidations themselves, however, do not belong to science because the terms in question stand at the beginning of the chain of definitions. In the case of logical category terms, however, there is no subsequent scientific use of those terms for which elucidations prepare us. The sense in which such elucidations 'do not belong to the logical theory' is stronger—it is because the logical categories do not belong to the subject matter of science at all, unlike the *Bedeutungen* of primitive scientific terms. I take it to be no coincidence, then, that Frege only employs the notion of a *propaedeutic* when he is discussing the elucidation of primitive scientific terms. Weiner fails to properly take this difference into account 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). If my arguments are correct, only elucidations of primitive scientific terms serve to communicate meanings.

While Weiner recognizes the diversity of the sorts of endeavours that Frege characterizes as 'elucidation', her discussion remains at a level of generality that leaves the specific nature of elucidation of the logical categories unaccounted for. In itself, this is not an objection to Weiner. One cannot, after all, investigate everything. The point I wish to make, is simply that we should not think that a characterization of elucidation in terms of the notion of a *propaedeutic* by itself suffices to equip us with an adequate understanding of the elucidation of the logical categories. There is a risk of assuming—since the notion of a *propaedeutic* is clear in the case of elucidation of primitive scientific terms—that we have thereby also achieved clarity about elucidation of the logical categories, a risk that Weiner's own discussion does not entirely keep at bay. With regards to the latter kind of elucidation, however, the work remains to be done: what kind of propaedeutic do we have here? What is it exactly, that elucidations of the logical categories are meant to help us understand? Weiner's general characterizations do not suffice to answer such questions.

6. Elucidation and logical features

The time has come to take some first steps towards the development of a novel understanding of Frege's elucidations of the logical categories, without seeking to model them on elucidations of primitive scientific terms. As announced, all I can do here is to give a general outline of the shape which, I think, such an understanding will have take, hopefully laying the groundwork for a more fully developed account.

Above, I acknowledged that Frege does not explicitly distinguish between elucidation of primitive scientific terms, and elucidation of the logical categories. I also suggested, however, that the difference is clearly registered in his *practice* of elucidation. We are now in a position to see how this is so. Weiner already notes the crucial feature of elucidations of the logical categories: "When Frege brings up the issue of what it is to be a function or concept, he invariably talks about linguistic expressions and how they work" (Weiner 2020, 85). Weiner herself does not, however, tie this observation in with her understanding of Fregean elucidation. Nevertheless, I believe that it is this feature that is the key to unlocking a proper understanding of Fregean elucidation of the logical categories.

Let me give an example of what Weiner has in mind. Talking about the proposition 'Two is a prime number', Frege writes:

> The two parts of the proposition are . . . 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. 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. 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. I am well aware that expressions like 'saturated' and 'unsaturated' are metaphorical and only serve to indicate what is meant—whereby one must always count on the co-operative understanding of the reader. (Frege 1984, 281)

This passage is not an outlier: Frege's elucidations of the logical categories consistently proceed in this manner.[37] In essence, my proposal is that we take this feature of Frege's elucidations seriously, which means that we cannot see it as an accident, or an aberration, that Frege consistently reverts to talking about signs in his attempts to elucidate the logical categories.

From a certain kind of philosophical perspective, however, it will appear as if this is a feature of Frege's elucidations that we *cannot* take seriously—at least, if we want to attribute to Frege a coherent philosophical view. Klement nicely gives voice to the sort of objection I have in mind:

> 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 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*. (Klement 2004, 4–5, emphasis added)

Klement is assuming a sharp distinction between 'the nature of things' and 'language'. Given such a distinction, no discussion of signs can pretend to equip us with a proper understanding of the nature of things. As Black puts it—targeting Frege—there is "something absurd in trying to base inferences about the logical structure of reality upon any physical characteristics of the signs we use" (Black 1968, 237).

Black is right. this is absurd. What Black (and Klement) do not realize, however, is that this need not exhaust the options. In talking about signs, is Frege seeking to base inferences about the logical structure of reality upon their physical characteristics? Such a characterization is based on at least two assumptions.

(A) There is a logical structure of reality that is prior to the logical structure of language.

[37] For further examples, see Frege (1984, 290–91), Frege (1979, 119, 177, 201, 271), Frege ([1893–1903] 2013 I, §1, §21), and Frege (1980, 141–42).

(B) The only relevant characteristics that signs possess, are physical characteristics.

This much is right: *if* these assumptions are correct, then Frege's practice of talking about signs cannot achieve its elucidatory purposes. Frege, however, is adamant that it is *by considering signs* that we can achieve an understanding of the logical categories. This is confirmed again in the *Grundgesetze.* Having presented the function sign '$\left(2 + 3 \bullet x^2\right) \bullet x$' as an example, Frege writes:

> The nature of the *function* lies therefore in that part of the expression [*liegt demnach in dem Theile des Ausdrucks*] 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 [1893–1903] 2013, I, §1)

In saying that the nature of the function lies in the part of an expression, Frege is doing precisely what Klement and Black finds problematic—characterizing the nature of the function by what can only be inessential linguistic features. From their vantage point, Frege's way of expressing the matter must be regarded as misguided. This is, however, not just some random piece of unpublished text that we have on our hands. It is the opening section of the *Grundgesetze*, in which Frege is seeking to clarify to his reader what is arguably the most fundamental notion of his logic, that of a function. If we can trust Frege to have properly articulated his views anywhere, it is here.

Rather than correcting Frege on his own behalf, we should take the fact that his way of expressing himself goes against (A) and (B) as an indication that (A) and (B) are philosophical assumptions that Frege did not make. Taken seriously, Frege's practice of talking about sign in elucidating the logical categories—and his insistence in the *Grundgesetze* that the nature of the function lies *in* the sign—reveal that Frege did not believe that there is the sort of prior logical order in reality that Klement and Black take for granted, and did not believe that there is nothing more to signs than mere physical characteristics.

I end by briefly suggesting how this may shape out. In discussing Frege's conception of the logical categories, Diamond talks about a logical order that is internal to the *Begriffsschrift*:

> 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)

It is this logical order—the inner logical order of language that is manifested in the *Begriffsschrift*—that Frege's elucidations of the logical categories are meant to get us to understand. It is precisely because it is the logical order *in the signs* that we need to understand, that talking about signs is an entirely natural way of seeking to convey such understanding. Signs, for Frege, do not merely have physical features—they also have *logical features*, features which they possess in virtue of their participation in the logical order that is inherent to language and thought as such. Unsaturatedness—the nature of the function that is at issue in the opening section of the *Grundgesetze*—is one such feature. The unsaturatedness of function signs is a logical feature of language that is at once revelatory of the unsaturatedness of functions. Rather than seeing the latter as prior to and independent from the former, we should see these as internally related. It is *through* an understanding of the unsaturatedness of function signs, that we achieve an understanding of the unsaturatedness of functions. That Frege believes that these cannot be separated, is what accounts for his approach to elucidating the logical categories.

This is, of course, non-accidentally reminiscent of the *Tractatus* conception of logical features as what shows itself in language. If my suggestions put us on the right track, then Frege's conception of elucidation of the logical categories is an important precursor to these ideas. Geach has equally argued that Frege must be seen as a precursor to the *Tractatus* conception of showing (Geach 1976). Geach, however, is working with the sort of naïve ineffabilist model that I have been arguing against, according to which elucidations constitute failed attempts to say what cannot be said—namely, that reality has certain features (Geach 1976, 54). The whole point of the turn to signs is precisely that elucidations are no longer understood as seeking to make metaphysical statements about reality—but rather as seeking to clarify, make clear, the logical order that is inherent in language, and in relation to which the

logical order of reality is not regarded as external and independent. So understood, what shows itself in language, is not external to language.[38]

This is also how we should understand the role of the *Begriffsschrift*. In his elucidations of the logical categories, Frege mostly talks about natural language signs, not *Begriffsschrift* signs.[39] This testifies to the fact that Frege believes that the logical category distinctions are present in natural language just as well as in the *Begriffsschrift*.[40] The logical order that Frege seeks to elucidate is inherent in natural language as well as in the *Begriffsschrift*. What the *Begriffsschrift* achieves, however, is precisely to make perspicuous that logical order in the construction of the signs themselves. We could say that, in natural language, the physical features of the signs obscure their logical features, whereas, in the *Begriffsschrift*, the logical features are made manifest in the physical features of the signs. Since the aim of elucidations of the logical categories is precisely to make us see those logical features, we must understand the *Begriffsschrift* as itself a device for elucidation, as itself part of Frege's activity of elucidating the logical categories.[41]

It could be objected that I am projecting too much of the Tractatus back into Frege. It is true that I am reading Frege in a way that brings him much closer to the Tractatus than most readers have allowed. My claim, however, is that this is the sort of reading that results from taking seriously how Frege approaches the elucidation of the logical categories, i.e., by engaging in discussions of signs. If I am correct, Frege's conception of elucidation is incompatible with the sort of realist

[38] Of course, there are many interpretations of the *Tractatus* conception of showing, not all of which will be congenial to the interpretation of Frege that I am proposing. For the sort of account that I think aligns well with my reading of Frege, see Narboux (2014). It remains an inescapable truth that one cannot interpret Frege without at least indirectly interpreting the *Tractatus*, and *vice versa*.

[39] Mostly, but not exclusively. Sections §§19-24 of the *Grundgesetze* constitute crucial elucidatory passages, and they proceed in terms of *Begriffsschrift* signs.

[40] At least, this is the case for the concept/object distinction. To what extent this applies to Frege's whole logical hierarchy, I regard as an open question. It is possible that Frege thought that natural language is logically incomplete insofar as it does not possess the means to express, say, a third-level relation between two second-level functions. More generally, the precise relation between the *Begriffsschrift* and natural language in Frege's elucidatory project, must be left open here as a matter for further investigation. I thank a reviewer for helping me see these matters more clearly.

[41] Of course, this is not the only purpose of the *Begriffsschrift*. It is also meant to be a device for constructing gapless proofs, for instance. These purposes are, in turn, related.

metaphysical framework that has often been ascribed to him. In this vein, Weiner has correctly pointed out that such ascriptions rely precisely on ignoring the fact that Frege's supposedly metaphysical pronunciations about objects, concepts, and the like, are the very statements which he systematically marks out as elucidations—which means that we cannot understand them as straightforward expressions of purported metaphysical fact.[42] There are no metaphysical facts in Frege, there are only logical facts, and these have a different status altogether.[43] That is why Frege consistently clarifies such purportedly metaphysical statements by reverting to a discussion of language. This only makes sense if Frege believes that there is a more intimate relation between world and language than Klement and Black allow. To grasp the logical features that are internal to language, is to grasp all logical order there is to grasp. There is no further, self-standing logical order 'in reality' that somehow transcends language. The logical order of language is all we get, and, according to Frege, all we need.

7. Conclusion

I have sought to lay the foundations for a novel approach to Frege's conception of elucidation of the logical categories. An essential prerequisite for such an approach is that we clearly distinguish such elucidation from elucidation of primitive scientific terms. While the latter revolves around getting us to grasp the *Bedeutung* of certain terms, no such thing is at issue in the former. Indeed, having made such a distinction, we start to see how Frege himself—if not in theory—distinguished between these two types of elucidation in practice. When elucidation of the logical categories is at issue, we consistently find Frege engaged in a certain kind of discussion of language, in which he highlights certain features of signs.

[42] See Weiner (2020, chap. 3) for relevant discussion of the sort of realist readings—advanced by philosophers such as Michael Dummett and Tyler Burge—that I have in mind.

[43] This is also how one should understand Frege's locutions about logical category distinctions being "founded deep in the nature of things" (Frege 1984, 156). Frege is referring to a *logical* nature, not a metaphysical one. See Diamond (1984, 365ff) for relevant discussion.

If we wish to understand Frege, we must take seriously this aspect of Frege's discussions. We must take seriously his idea that the nature of the logical categories lies *in* the notation. I have introduced a notion of *logical feature* that seeks to capture what Frege is after here. To understand the logical categories, is to understand the logical order that is internal to language, as it is manifested in logical features of signs—features that cannot be reduced to merely physical characteristics. Frege's notion of unsaturatedness, for instance, is meant to capture precisely such a logical feature.

As I have emphasized, this is but a beginning. Much more needs to be said, for instance, about this notion of a *logical feature*, and the sort of understanding in which the 'seeing' of such features is supposed to consist in. In what way, exactly, are logical features to be distinguished from merely physical features? How can this notion of 'seeing' logical features avoid falling back into an ineffabilist view? It is my hope that my discussion puts us in the right position to start thinking about these questions. What I do believe can already be concluded, is this: however the sort of account I am sketching here is to be fleshed out, it will be incompatible with a realist understanding of Frege as committed to a hard and fast distinction between 'the nature of things', on the one hand, and 'language', on the other. That is to say: for this notion of a logical feature to do the work that—on my account—it is meant to do, is for it to render incoherent the very idea that there is a clean separation between the logical order of the nature of things, on the one hand, and the logical order of language, on the other hand. Painting with a broad brush, we could say that my arguments, if successful, vindicate a Kantian reading of Frege.

Acknowledgements

The research for this paper was supported by an FWO Junior postdoctoral grant (project 3E008321). It has greatly benefited from numerous discussions with Jim Conant, as well as comments from Martin Gustafsson, Michael Kremer, Jean-Philippe Narboux, and Gertrudis Van de Vijver. I also wish to thank two anonymous referees for their insightful remarks.

References

Beaney, Michael. 1996. *Frege: Making Sense*. London: Duckworth.

Black, Max. 1968. "Frege on Functions." In *Essays on Frege*, edited by E. D. Klemke, 223–48. Chicago: University of Illinois Press.

Conant, James. 2002. "The Method of the Tractatus." In *From Frege to Wittgenstein*, edited by Erich H. Reck, 374–462. Oxford: Oxford University Press.

———. 2020. "Replies." In *The Logical Alien*, edited by Sofia Miguens, 321–1028. Cambridge, MA: Harvard University Press.

Diamond, Cora. 1984. "What Does a Concept Script Do?" *Philosophical Quarterly* 34: 343–68. https://doi.org/10.2307/2218765.

Dummett, Michael. 1981. *Frege: Philosophy of Language*. 2nd ed. London: Duckworth.

Frege, Gottlob. 1979. *Posthumous Writings*. Edited by Hans Hermes, Friedrich Kambartel, and Friedrich Kaulbach. Translated by Peter Long and Roger White. Oxford: Basil Blackwell.

———. 1980. *Philosophical and Mathematical Correspondence*. Edited by Gottfried Gabriel, Hans Hermes, Friedrich Kambartel, Christian Thiel, Albert Veraart, and Brian McGuinness. Translated by Hans Kaal. Chicago: The University of Chicago Press.

———. 1984. *Collected Papers on Mathematics, Logic, and Philosophy*. Edited by Brian McGuinness. Translated by Max Black, V. H. Dudman, Peter Geach, Hans Kaal, E.H.W. Kluge, B. F. McGuinness, and R. H. Stoothoff. New York: Basil Blackwell.

———. (1893–1903) 2013. *Basic Laws of Arithmetic*. Edited and translated by Philip A. Ebert, Marcus Rossberg, and Crispin Wright. Oxford: Oxford University Press.

Geach, Peter. 1976. "Saying and Showing in Frege and Wittgenstein." In *Essays in Honor of G. H. Von Wright*, edited by Jaakko Hintikka, 54–70. Amsterdam: North-Holland Publishing Company.

Hale, Bob, and Crispin Wright. 2012. "Horse Sense." *The Journal of Philosophy* 109: 85–131.

Heck, Richard. 2010. "Frege and Semantics." In *The Cambridge Companion to Frege*, edited by Michael Potter and Thomas Ricketts, 342–78. Cambridge: Cambridge University Press.

Hugly, Philip. 1973. "Ineffability in Frege's Logic." *Philosophical Studies* 24: 227–44. https://doi.org/10.1007/bf00356488.

Johnston, Colin. 2007. "Symbols in Wittgenstein's Tractatus" 15: 367–94. https://doi.org/10.1111/j.1468-0378.2007.00253.x.

Jolley, Kelly Dean. 2015. "Once Moore Unto the Breach! Frege and the Concept 'Horse' Paradox." *Philosophical Topics* 43: 113–24.

Jones, Nicholas K. 2016. "A Higher-Order Solution to the Problem of the Concept Horse." *Ergo: An Open Access Journal of Philosophy* 3: 132–66. https://doi.org/10.3998/ergo.12405314.0003.006.

Klement, Kevin C. 2004. "Putting Form Before Function: Logical Grammar in Frege, Russell, and Wittgenstein." *Philosophers' Imprint* 4: 1–47.

Long, Peter. 2001. *Logic, Form, and Grammar*. New York: Routledge.

Moore, A. W. 2012. *The Evolution of Modern Metaphysics: Making Sense of Things*. Cambridge: Cambridge University Press.

Narboux, Jean-Philippe. 2014. "Showing, the Medium Voice, and the Unity of the Tractatus." *Philosophical Topics* 42: 201–62. https://doi.org/10.5840/philtopics201442224.

Noonan, Harold. 2006. "The Concept Horse." In *Universals, Concepts and Qualities: New Essays on the Meaning of Predicates*, edited by P. F. Strawson and Arindarm Chakrabarti, 155–76. Burlington, VT: Ashgate.

Proops, Ian. 2013. "What Is Frege's 'Concept Horse Problem'?" In *Wittgenstein's Tractatus*, edited by Peter Sullivan and Michael Potter, 76–96. Oxford: Oxford University Press.

Ricketts, Thomas. 2010. "Concepts, Objects and the Context Principle." In *The Cambridge Companion to Frege*, edited by Michael Potter and Thomas Ricketts, 149–219. Cambridge: Cambridge University Press.

Russell, Bertrand. (1903) 2010. *Principles of Mathematics*. London and New York: Routledge.

Sullivan, Peter. 2003. "Ineffability and Nonsense." *Proceedings of the Aristotelian Society, Supplementary Volumes* 77: 195–223. https://doi.org/10.1111/1467-8349.00109.

———. 2006. "Metaperspectives and Internalism in Frege." In *Gottlob Frege: Critical Assessments of Leading Philosophers*, edited by Michael Beaney and E. H. Reck, 85–105. London: Routledge.

Travis, Charles. 2020. "Where Words Fail." In *The Logical Alien*, edited by Sofia Miguens, 222–80. Cambridge, MA: Harvard University Press.

Vanrie, Wim. 2021. "Why Did Frege Reject the Theory of Types?" *British Journal for the History of Philosophy* 29: 517–36. https://doi.org/10.1080/09608788.2020.1807908.

———. n.d. "Frege's Conception of the Absoluteness of the Logical Category Distinctions." In *Early Analytic Philosophy: Origins and Transformations*, edited by Gilad Nir and James Conant. London: Routledge.

Weiner, Joan. 2006. "On Fregean Elucidation." In *Gottlob Frege: Critical Assessments of Leading Philosophers*, edited by Michael Beaney and E. H. Reck, 197–214. London: Routledge.

———. 2010. "Understanding Frege's Project." In *The Cambridge Companion to Frege*, edited by Michael Potter and Thomas Ricketts, 32–62. Cambridge: Cambridge University Press.

———. 2020. *Taking Frege at His Word*. Oxford: Oxford University Press.

Wells, Rulon S. 1968. "Is Frege's Concept of a Function Valid?" In *Essays on Frege*, edited by E. D. Klemke, 391–406. Chicago: University of Illinois Press.

Witherspoon, Edward. 2002. "Logic and the Inexpressible in Frege and Heidegger." *Journal of the History of Philosophy* 40: 89–113. https://doi.org/10.1353/hph.2002.0021.

Wittgenstein, Ludwig. (1922) 1981. *Tractatus Logico-Philosophicus*. Translated by C. K. Ogden. London: Routledge.

Wright, Crispin. 1998. "Why Frege Did Not Deserve His Granum Salis." *Grazer Philosophische Studien* 55: 239–63. https://doi.org/10.5840/gps19985522.

Journal for the History of Analytical Philosophy

VOLUME 13, NUMBER 1 (2024)

ISSN 2159-0303

jhaponline.org