Frege's Conception of Logic: From Kant to Grundgesetze

Øystein Linnebo

February 21, 2003

The last few decades have brought impressive new technical insights regarding Frege's logicism and his "reduction of arithmetic to logic."¹ This paper, however, deals with the complementary but far less investigated question how Frege understood the nature of logical truth and of logical knowledge. I shall examine Frege's conception of logic as it developed and matured, beginning with his early *Begriffsschrift* from 1879 and following it up through to *Grundgesetze* I from 1893.²

I shall make two main claims. My first main claim is that Frege started out with a view of logic that is closer to Kant's than is generally recognized, but that he gradually came to reject this Kantian view, or at least totally to transform it. My second main claim concerns Frege's reasons for distancing himself from the Kantian conception of logic. It is natural to speculate that this change in Frege's view of logic may have been spurred by a desire to establish the logicality of the axiom system he needed for his logicist reduction, including the infamous Basic Law V. I admit this may have been one of Frege's motives. But I shall argue that Frege also had a deeper and more interesting reason to reject his early Kantian view of logic, having to do with his increasingly vehement anti-psychologism.

1 Kant's Conception of Logic

I shall begin with a brief characterization of Kant's conception of logic. According to Kant, understanding and reason are powers of the human mind which operate according to certain rules. These rules can be either contingent or necessary. The contingent rules are ones that depend on and vary with the class of objects that is considered. There are, for instance,

¹See e.g. the papers collected in Demopoulos (1995), especially those by George Boolos and Richard Heck.

²I shall use the word 'Begriffsschrift' as a general term to characterize formal languages of the sort Frege was trying to develop. I shall use the italicized '*Begriffsschrift*', or simply '*BS*', to refer to Frege's 1879 book. For the other abbreviations, see the references at the end.

different uses of the understanding in mathematics, metaphysics, and morality. Kant defines general logic as the part of logic that is concerned exclusively with "absolutely necessary rules of thinking, without which no use of the understanding takes place" (A52/B76).³ This distinguishes general logic from various special logics, which are concerned with special employments of the understanding. An important example of a special logic is Kant's own transcendental logic. Next, Kant says that logic is *pure* when "we abstract from all the empirical conditions under which our understanding is exercised" (A53/B77). For instance, it is not the task of pure logic to advise people how better to remember and keep track of the premises of their arguments. What corresponds to our notion of logic is thus Kant's notion of pure general logic.

In this paper, I shall only be concerned with pure general logic. So I shall adopt the definition of logic that Kant gives in his *Logic*, where he shares this concern. Here Kant defines [pure general] logic as the "science of the necessary laws of the understanding and reason in general" (*Logic*, p. 15).

Kant explains why the necessary rules of the understanding hold for all use of the understanding, irrespective of the class of objects considered, by appealing to the *formality* of these rules. According to Kant, the necessary rules of the understanding concern only the *form* of thought, not its *matter*. These rules therefore abstract completely from all differences between objects. To further elucidate this difference between the formal and the material aspects of thought, Kant compares logic with grammar. In being the science of the form of thought, logic is like "a general grammar which contains nothing beyond the mere *form* of a language in general, without words, which belong to the *matter* of the language" (*ibid.*, my emphasis). In this sense, logic may be said to be the grammar of thought. Just like an ordinary grammar abstracts from the meanings of individual words, logic abstracts from all particular characteristics of objects.

Kant's definition of logic as "the science of the necessary laws of the understanding and reason in general" may give the impression that the task of logic is to describe the laws that govern actual human thought. This would be a psychologistic view of logic, which would make the subject into a branch of empirical psychology. But Kant explicitly warns against this sort of psychologism. Bringing psychological principles into logic, he says, would be "as absurd as taking morality from life" (*Logic*, p. 16)—and as we know, Kant took this to be *very* absurd. Kant's writings offer at least three reasons why a psychologistic understanding of logic is

 $^{^3\}mathrm{References}$ of the form '(Ax/By)' are always to Kant's Critique of Pure Reason.

unacceptable. First, an investigation of logic can proceed entirely *a priori*, without any need for empirical study of actual human thought. In Kant's words, logic is "a self-cognition of the understanding and of reason" (*ibid.*). Secondly, empirical psychology can only "lead to the cognition of merely contingent laws" (*ibid.*), whereas the laws of logic are *necessary*. Thirdly, the laws of pure logic have a *normative force* which no findings of empirical psychology could have. For instance, Kant writes

In logic we do not want to know how the understanding *is and thinks*, but how it *ought to* proceed in thinking. Logic shall teach us the right use of the understanding, i.e. the one that agrees with itself. (*ibid.*, my emphasis)

Kant here calls attention to the distinction between *descriptive* and *normative* laws. These two kinds of law have different kinds of validity. A descriptive law cannot be violated by any object of which it is valid. For instance, if "the universal law of gravitation" really is a universal law, there can be no exceptions to it. In contrast, a normative law can be valid for a person although this person fails to obey it.

Kant's emphasis on the normative character of the laws of logic gives rise to a problem. The laws of logic appear to be *both* descriptive laws of certain necessary rules of the understanding *and* normative laws of how we ought to think. But if the laws of logic really are descriptive laws of the understanding, doesn't it follow that the understanding is bound to observe these laws? But if so, logical mistakes would be impossible! Kant is aware of this problem. He writes that a logical mistake would be "a form of thinking contrary to the understanding" (*Logic*, p. 59), and since the understanding is the faculty of thought, this would be thinking that conflicts with its own faculty. Illogical thought therefore appears as impossible as a "force ... deviat[ing] from its own essential laws" (*ibid*.).

But Kant offers an ingenious solution to the paradox that appears to make logical mistakes impossible. All that follows from our premises, he writes, is that, "if we had no other power of cognition beside the understanding, we would never err" (*ibid.*). Logical mistakes are therefore possible provided they do not originate in the understanding itself. Not surprisingly, Kant locates the origin of logical mistakes in "the unnoticed influence of sensibility upon the understanding, or, more exactly, upon judgment."⁴ So when we make a logical mistake, the fault of the understanding is one of omission—of ignorance and inattention—rather than one of commission—of deviating from its own essential laws.

 $^{^{4}}$ *Ibid.*, but see also A294/B350.

This solution provides the resources to reconcile the normative and the descriptive aspects of the laws of logic. When a person violates a law of logic, the problem isn't that he is thinking an illogical thought but that he isn't thinking any thought at all—although it seems to him he does. As Kant is fond of saying, without the laws of logic there could be no thought at all. In this sense the laws of logic are *constitutive* of thought: They tell us how we are to use the understanding if we are to succeed in thinking at all—namely in accordance with its own essential laws, uncorrupted by illegitimate sensible influences.

Summing up and fixing our terminology for the rest of the paper, I shall call the position based on the following two theses the Kantian conception of logic.

- *The Formality Thesis.* Logic is purely formal. Its relation to thought is like the relation of a "general grammar" to a language.
- *The Constitutivity Thesis.* Logic is constitutive of thought. More precisely, there are non-empirical notions of thought and understanding, and the laws of logic are constitutive of thought and understanding thus understood.

2 Frege and the Formality Thesis

My first main claim is that Frege started out with a view of logic that is closer to Kant's than is generally recognized, but that he gradually distanced himself from this view. More precisely, I shall defend the following view. In BS and other writings from the same period, Frege attempted to understand his logic in accordance with the Formality Thesis. But this thesis disappears from his philosophy already in GLA, a few years later. The Constitutivity Thesis too was endorsed around the time of BS. But this thesis remained for a longer time, and by the time of GLA, it played a central role Frege's conception of logic.

Let's begin with the Formality Thesis. In a series of articles from the early 1880s Frege set out to explain his Begriffsschrift and to defend its superiority over competing logical systems, such as those devised by Boole and his followers. A recurring theme in these articles is that the Boolean logicians merely have constructed a *calculus* for pure logic, whereas the Begriffsschrift is a *language* capable of expressing content. The following two quotes are characteristic.

I did not wish to present an abstract logic in formulas, but to express a content through written symbols in a more precise and perspicuous way than is possible with words. In fact, I wished to produce, not a mere *calculus ratiocinator*, but a *lingua characteristica* in the Leibnizian sense. (CN, pp. 90-1 (1-2))

Right from the start I had in mind the expression of a content. What I am striving after is a *lingua characterica* in the first instance for mathematics, not a *calculus* restricted to pure logic. (*PHW*, p. 12)

Some interpreters have taken Frege's distinction between a calculus and a language as evidence that Frege regarded logical truths as contentful.⁵ But a careful reading of the relevant passages shows that what Frege had in mind wasn't this strong claim that the purely logical part of the Begriffsschrift has content all by itself, but the weaker claim that this language allows content to be expressed when contentful signs from the special sciences are added to it. Consider, for instance, the following explanation, which occurs only two pages after the first passage quoted above.

But [the rendering of a content] is exactly my intention. I wish to blend together the few symbols which I introduce and the symbols already available in mathematics to form a single formula language. In it, the existing symbols [of mathematics] correspond to the word-stems of [ordinary] language; while the symbols I add to them are comparable to the suffixes and [deductive] formwords that logically interrelate the contents embedded in the stems.⁶

Here Frege says that his Begriffsschrift is to be supplemented with existing symbols of mathematics and *in this way express content*. The task of the Begriffsschrift is to contribute the "suffixes and formwords" needed to "logically interrelate" the contentful signs of arithmetic.

Frege is even more explicit in his essay "Boole's Calculating Logic and the Begriffsschrift," from which the second of the passages quoted above is taken. In the very paragraph from which the passage quoted is taken, Frege explains that "every highly developed language must consist" of "two components": a "formal part which in verbal language comprises endings, prefixes, suffixes and formwords" and "a material part proper" (*PHW*, p. 13)). So far, the language of arithmetic has only had the material part and completely lacked "the logical cement that will bind these building stones together" (*ibid.*). For this reason, mathematicians have had to rely on natural language to express logical relations. Boole's symbolic logic, on the other hand, represents only the formal part of language. Frege's goal with his Begriffsschrift was to develop a language in which *both* the formal *and* the material parts of language would come to their right. In order to construct such a Begriffsschrift, in the first place for the

 $^{{}^{5}}$ See van Heijenoort (1967), and Goldfarb (1979) and (2001).

 $^{^6\,}C\!N\!,$ p. 93 (4); all explanations in brackets except the first are by the translator.

limited field of mathematics, Frege says that we must "supplement the signs of mathematics with a formal element" (*ibid.*). This supplementation is Frege's Begriffsschrift. It follows that Frege took the logical part of his Begriffsschrift to be purely formal, yet of a logical structure rich enough to enable a perspicuous representation of any arithmetical proposition, once this language has been supplemented with the appropriate contentful arithmetical signs.

Frege's comparison of the logical and the extra-logical signs of the Begriffsschrift with the formal and the material signs of natural language parallels a similar comparison we found in Kant. For Kant, we recall, explains the formality of logic by comparing the science of logic with "that of a general grammar which contains nothing beyond the mere form of a language in general, without words, which belong to the matter of language" (*Logic*, p. 15). The close similarities between Kant's and Frege's comparisons of logic with the formal grammar of a language makes it plausible that Frege was aware of this view of Kant's and that he agreed with it.⁷

It will be objected that other passages from the same period indicate that Frege *did* take his basic logical laws to have content. For instance, at the beginning of part III of *BS* Frege claims that "pure thought, irrespective of any content given by the senses or even by an intuition *a priori*, can, solely from the content that results from its own constitution, bring forth judgments that at first sight appear to be possible only on the basis of some intuition" (*BS*, p. 55 (55)). This indicates that he took at least some logical truths to have content. Moreover, in *BS* §2 Frege explains that "Whatever follows the content stroke must have a content that can become a judgment." Since any well-formed sentential expression of the Begriffsschrift can follow the content stroke, it appears that logical truths too must have content.

I believe Frege's explication in BS §3 of the notion of content he is working with opens for a reconciliation of these passages with the ones I have focussed on. Here Frege says that his notion of content, which he calls "conceptual content," is determined by considering "only that which influences ... possible inferences [Folgerungen]" (BS, p. 12 (3)). This means that contents are at least partially individuated by their logical relations to one another. Since all logical truths are equipollent, it can therefore be argued they share one and the same conceptual content, namely the *null-content* that entails no content other than itself.

Although I find this proposal plausible, I won't here pursue it any further. For my present argument it suffices that Frege went out of his way to construe his logic in accordance with

⁷We know that Frege was familiar with Kant's *Logic* because he quotes it in GLA, p. 19.

the Formality Thesis. And for this claim, the above discussion provides ample evidence.

The Formality Thesis was a rather short-lived part of Frege's view of logic. It conflicts with Frege's view in GLA that the numbers are logical objects. If logical objects exist, then logic will be a science with its own realm of objects, in which case logic can no longer be said to "disregard the particular characteristics of objects" (BS, p. 5 (iv)). It therefore comes as no surprise that in GLA, the Formality Thesis disappears from Frege's conception of logic. The Kantian grammar analogy disappears as well.

3 Frege and the Constitutivity Thesis

There is some evidence that Frege held the Constitutivity Thesis already in BS. For instance, Frege there refers to the laws of logic as "laws of thought that transcend all particularities" (BS, p. 5 (iv); my emphasis). Moreover, in a passage from BS already quoted Frege claims that pure thought can bring forth judgments "solely from the content that results from its own constitution" (BS, p. 55 (55); my emphasis).

However, since Frege's commitment to the Constitutivity Thesis is much more articulated in GLA, I shall focus on this work instead. One of the clearest statements of the thesis is found in GLA §14, where Frege writes that

we have only to try denying any one of [the truths of logic], and complete confusion ensues. Even to think at all seems no longer possible. (GLA, p. 21)

Frege goes on to explain that logic holds in the widest domain of all, namely for *everything* thinkable. His idea is that, if we are to succeed in thinking at all, we must observe the laws of logic. In GLA §105 we find another indication of how close Frege took the connection to be between logic and the faculty of thought. In this section Frege characterizes arithmetic and analysis as part of "the reason's proper study of itself." So he took the source of logic and arithmetic to lie within reason itself. And in so doing, Frege is echoing Kant's Constitutivity Thesis. Yet another example is Frege's denial in GLA §3 that his characterization of analyticity in terms of his new logic assigns a new sense to Kant's old label. All he does, Frege says, is "state accurately what earlier writers, Kant in particular, have meant by them" (GLA, p. 3). I believe this footnote was sincere and that Frege really took himself to be clarifying Kant's original definitions. But if so, he must have been sympathetic with the conception of logic that underlies Kant's original definitions.

The Constitutivity Thesis plays a central role in Frege's discussion of other important

topics of GLA as well. For instance, it plays an essential role in his characterization of objectivity. In philosophy we say that something is objective if it is, in an appropriate sense, independent of the cognizers and their activities and capacities. But in GLA Frege insists that, when characterizing the notion of objectivity, we should not go as far as to demand independence from our own faculty of reason. Just like Kant, Frege was convinced that no question would make sense were we to abstract from our reason. If we make the notion of objectivity so strong that a truth qualifies as objective only if it holds independently of our reason, the search for objectivity has become a search for something impossible and at bottom meaningless. Seeking objectivity of this kind is, in Frege's words, like wanting to "judge without judging, or to wash the fur without wetting it" (GLA, p. 36).

Accordingly, the notion of objectivity must be *based on reason* rather than attempt to *abstract away from it.*⁸ Frege's particular way of doing this is to construe objectivity as independence from individual psychology.

I understand objective to mean what is independent of our sensation, intuition and imagination, and of all construction of mental pictures out of memories or earlier sensations, but not what is independent of reason. (GLA, p. 36).

Thus, the subjective is what is peculiar to the minds of individual people, and the objective, what can be shared by several minds.

These examples show that Frege's commitment to the Constitutivity Thesis was deeper than that to the Formality Thesis.

4 Frege's Rejection of the Constitutivity Thesis

Nevertheless, I shall now argue that in GGA, the Constitutivity Thesis—at least in anything resembling the original Kantian form—suffered a similar fate. I shall offer three arguments.

The Constitutivity Thesis, we recall, attempts to base the normativity of the laws of logic on the constitutive role of these laws in our thought. The thesis is expressed in GLA §14, where Frege claims that, if a law of logic is denied, "complete confusion ensues" and "even to think at all seems no longer possible" (GLA, p. 21). This argument is reconsidered in the introduction to GGA.

If we step away from logic, we may say: we are compelled to make judgments by our own nature and by external circumstances; and if we do so, we cannot

⁸See GLA, p 38.

reject this law—of Identity, for example; we must acknowledge it unless we wish to reduce our thought to confusion and finally renounce all judgment whatever. (GGA, p. 15 (xvii))

But Frege's assessment of this argument has changed:

I shall neither dispute nor support this view; I shall merely remark that what we have here is not a logical consequence. What is given is not a reason for something's being true, but for our taking it to be true. (*ibid.*)

Frege now sees the argument from GLA §14 as bearing only on our takings-to-be-true [Fürwahrhalten]. But this is an empirical notion, which belongs to psychology and has nothing to do with objective truth.⁹ The argument from GLA §14 is thus relegated to empirical psychology: That our nature compels us to judge in certain ways is relevant only to the study of how we *take* various claims to be true and has nothing to do with the question whether these claims actually *are* true.

Second, in the introduction to GGA, Frege explicitly rejects Kant's comparison of logic with grammar, and along with this, the idea that the validity of the laws of logic may somehow be based on the structure of human thought. The laws of logic, Frege says,

are boundary stones set in an eternal foundation, which our thought can overflow, but never displace. [...] They do not bear the relation to thought that the laws of grammar bear to language; they do not make explicit the nature of our human thinking and change as it changes. (GGA, p. 13 (xvi))

Frege now regards all accounts of "the nature of our human thinking" as pieces of empirical psychology and thus as knowledge of something contingent and uncertain. This means he has rejected Kant's belief that a non-empirical study of the human mind is possible. For if such a study were possible, it would enable us to establish eternally valid principles about the nature of human thought. But this is exactly what Frege now denies.¹⁰

Third, Frege no longer states logicism as the thesis that arithmetic is *analytic*. Starting with GGA, the word 'analytic' disappears from Frege's writings. The logicist thesis is now expressed simply as the claim that arithmetic is *reducible to logic*. So it appears that Frege no longer finds Kant's notion of analyticity an apt characterization of the truths of logic.

 $^{^{9}}$ Evidence for this interpretation of the concept of "Fürwahrhalten" is provided by *PHW*, pp. 146 and 149.

¹⁰However, Frege's argument does not call into question his technical notion of a thought. For a Fregean thought is the objective sense of a sentence, which exists regardless of the existence and nature of human thought.

Why, then, did Frege reject the Constitutivity Thesis? One motive might have been to guarantee the logicality of the axiom system he needed for his logicist reduction. In the last section we saw how the *GLA* view that the numbers are logical objects led Frege to reject the Formality Thesis. Frege's adoption of Basic Law V in *GGA* may have played a similar role in leading him to reject the Constitutivity Thesis. For it is hard to see how Basic Law V can be said to be constitutive of thought. Unlike the other logical axioms, Basic Law V doesn't seem to be the kind of truth which, if denied, would make all thought collapse into confusion.

I don't want to deny that this may have been one of Frege's motives for rejecting the Constitutivity Thesis. But I do want to suggest that Frege's philosophy provides another, at least as powerful and much more interesting, route to the rejection of the Constitutivity Thesis. This route goes via Frege's increasingly vehement anti-psychologism.

Although the philosophies of Kant and Frege both are motivated by questions about the epistemic status of various kinds of knowledge, they provide very different frameworks for assessing epistemic status. Kant characterizes the difference between a priori and a posteriori in terms of the source of knowledge: The question is whether our knowledge of a truth "arises from experience" or is something that "our own cognitive faculty ... provides out of itself" (B1). Obviously, this question presupposes Kant's theory of the mind. Moreover, as we saw in Section 1, Kant's conception of logic too presupposes his theory of the mind. In Frege's framework for assessing epistemic status, there are no explicit references to the mind or to mental notions. Starting in BS, Frege insisted that the epistemic status of a truth is a matter of the best and deepest justification that can be given of this truth; it doesn't matter how we happen to discover this truth. In GLA this idea is incorporated into the definitions Frege gives of the notions of analytic and synthetic, a priori and a posteriori. The task of assessing the epistemic status of a truth thus becomes "that of finding the proof of the proposition, and of following it up right back to the primitive truths" (GLA, p. 4).

How deep this apparent difference between Frege and Kant is depends on whether Frege's framework for classifying the epistemic status of truths somewhere makes *implicit* reference to the mind or to mental notions. One possible implicit reference of this kind is Frege's early acceptance of the Kantian conception of logic.

As Frege's anti-psychologism hardened, I believe Frege became more and more uncomfortable with this appeal to the human mind lingering at the heart of his philosophy of logic. Although the theory of mind to which Kant's conception of logic appealed is a non-empirical theory, Frege began to feel that even this would threaten the absolute validity of logical laws. More generally, Frege seems to have thought that the absolute validity of logic would be compromised if our philosophical account of logic were to appeal to *anything* extra-logical. He thus writes, in the introduction to GGA, that it is "a sure sign of mistake if logic has need of metaphysics and psychology—sciences that require their own logical first principles" (GGA, p. 18 (xix)). Because of this very strict requirement on an account of logic, Frege sees no alternative but to take the validity of the logical laws as basic. On this view, the validity of the logical laws does not flow from the structure of the human mind but is a matter of "boundary stones set in an eternal foundation."

Admittedly, since this eternal foundation is nothing but the logical structure of Frege's "third realm" of mind-independent thoughts, there remains an attenuated sense in which logic still is constitutive of thought. But this constitutivity claim is far removed from the original Constitutivity Thesis, which says that logic is constitutive of actual human thought.

If I am right that Frege's anti-psychologism was his primary reason for rejecting the Constitutivity Thesis, one may wonder whether this anti-psychologism led him to go too far. When the Constitutivity Thesis says that the laws of logic have to be observed if thinking is to be at all possible, perhaps this reference to human thought should be understood, not as a *condition on the validity of the laws of logic*—which is what Frege objects to—but as part of an *account of what makes our thoughts the thoughts they are*; that is, as part of an account of the identity of the items that stand in logical relations. Only an account of this sort, it seems to me, can provide an informative answer to the question what distinguishes a logical truth from a non-logical one.¹¹

¹¹This is a revised version of a talk I gave in August 1999 at *The 11th International Congress of Logic, Methodology, and Philosophy of Science*, in Kraków, Poland. This talk was based on a longer manuscript, which in 1998 served as my second-year paper at Harvard University. I am grateful to my advisor for this second-year paper, Warren Goldfarb, for valuable discussion and criticism. Thanks also to Tony Corsentino, John MacFarlane, Richard Heck, Charles Parsons, Marco Ruffino, and participants at the Kraków conference for useful discussions of the ideas presented in this paper.

References

Works by Frege:¹²

- BS. Begriffsschrift. Transl. and repr. in J. van Heijenoort (ed.), From Frege to Gödel: A Source Book in Mathematical Logic, Cambridge: Harvard University Press, 1967.
- CN. Conceptual Notation and Related Articles. T.W. Bynum (ed.), Oxford: Clarendon Press, 1972.
- CP. Collected Papers on Mathematics, Logic, and Philosophy. B. McGuinness (ed.), Oxford: Blackwell, 1984.
- GGA. Grundgesetze der Arithmetik. Georg Olms Verlagsbuchhandlung, 1962. Partial transl. by M. Furth as Basic Laws of Arithmetic, Berkeley: University of California Press, 1964.
- GLA. Grundlagen der Arithmetik. Transl. by J.L. Austin as Foundations of Arithmetic, Oxford: Blackwell, 1953.
- PHW. Posthumous Writings. H. Hermes et al (eds.), Oxford: Blackwell, 1979.

Works by Kant:

- Kritik der reinen Vernunft. Transl. as Critique of Pure Reason by P. Guyer and A. Wood, Cambridge: Cambridge University Press, 1997.
- Logik. Transl. as Logic by R.S. Hartman and W. Schwarz, New York: Dover Publications, 1974.

Works by other authors:

- Demopoulos, William (ed.). Frege's Philosophy of Mathematics. Cambridge, MA: Harvard University Press, 1995.
- Goldfarb, Warren. "Logic in the Twenties: The Nature of the Quantifier." Journal of Symbolic Logic, 44, 351-68, 1979.

van Heijenoort, Jean. "Logic as Calculus and Logic as Language." Synthese, 17, 324-30, 1967.

¹²Where reprints or translations are listed, my page references are to these. Where the pagination of the reprints differs from that of the originals, I list the latter in parentheses.