Abstract: According to Peter Aczel, the inconsistency of Frege’s system in *Grundgesetze* is due, not to the introduction of sets, as is usually thought, but to the introduction of the Horizontal. His argument is that the principles governing sets are intuitively correct and therefore consistent, while the scheme introducing the Horizontal amounts to an internal definition of truth conflicting with Tarski’s classic result on the undefinability of truth in the object language. The aim of this paper is to show that the Horizontal is innocent: Aczel’s diagnosis is based on a mistaken view of the structure underlying Frege’s ideal language.


1. OPENING

Frege’s Horizontal satisfies the scheme

\((H) \quad \gamma \text{ is the True if and only if } \gamma \text{ is the True,}\)

where ‘\(\gamma\)’ is a dummy standing for sentences and also for singular terms. Typical instances of \((H)\) are sentences like ‘\(\overline{\text{— 2+3=5 is the True if and only if 2+3=5 is the True.}}\)’.

In his seminal article “Frege structures and the notions of proposition, truth and set” (1980), Peter Aczel argued that the inconsistency of Frege’s system in *Grundgesetze* is due, not to the introduction of sets (or value-courses), as is usually thought, but to the introduction of the Horizontal, because the principles governing sets are intuitively correct and therefore consistent, whereas scheme \((H)\) gives an internal definition of truth conflicting with Tarski’s classic result on the undefinability of truth.

Frege’s critical mistake, we are told, was to construe the structure underlying his ideal language in such a way that it satisfies the condition that there are exactly two propositions, namely the True and the False, with the True being of course the only true proposition. For, this condition implies that in Frege’s system the collection of truths can be internally defined, as Frege actually does, in terms of the identity relation:

\[(D) \quad \Delta \equiv \Delta = (\Delta = \Delta).\]

This diagnosis comes as a surprise, considering that Frege himself subscribed to the undefinability of truth. Indeed, his conception of truth has often been regarded as a predecessor of Tarski’s precisely because of the doctrine that truth is a concept that is so fundamental that it cannot be reduced to something more fundamental.\(^2\)

In what follows, my aim is to make plausible the claim that the Horizontal is innocent. The main theses I wish to defend are: (i) the definition \((D)\) does not conflict with Frege’s own thesis of the undefinability of truth, because the Horizontal does not express the

\(^1\) Cf. Aczel, 1980, pp. 32, 40-1. I use ‘\(\equiv\)’ as an abbreviation for ‘means by definition that’.

notion of truth in Frege’s system, (ii) neither does \((D)\) conflict with
Tarski’s undefinability results, because \((D)\) does not define the collection
of truths in Frege’s system, (iii) nevertheless, there is a conflict between
Frege’s system and a more general theorem deriving from Cantor’s
diagonal argument according to which there can be no set of all truths,
(iv) this conflict derives from Frege’s ontology which implies that there
must be more truths than there can possibly be. The paper is structured
as follows. In section 1, the role of the Horizontal in Frege’s ideal
language is described. The task of section 2 is to demonstrate the
innocence of the Horizontal. Finally, in section 3, the conflict with
Cantor’s theorem is sketched.

2. THE ROLE OF THE HORIZONTAL IN FREGE’S SYSTEM

To explain the role of the Horizontal, we must briefly recapitulate
the general structure of the ideal language in *Grundgesetze*.\(^3\) The
characteristic feature of this language is that the sentences always have
the form ‘\(\vdash \Delta\)’, where the vertical stroke ‘\(\vdash\)’ is the so-called “judgement-
stroke”. It is commonly assumed that the judgement-stroke is an
illocutionary operator whose task is to make the attachment of assertoric
force explicit in order to overcome the “logical defect” of natural
language that the very same form of words – a declarative sentence – can
be used, now as having assertoric force, now as lacking it.\(^4\) By writing ‘\(\vdash
2+3=5\)’, for instance, the author makes explicit that s/he asserts or judges
the thought expressed by ‘\(2+3=5\)’ as true.

\(^3\) For an analysis of the role of ‘\(\vdash\)’ in Frege’s first system, see Simons, 1996,
p. 282 ff.

It can, however, be shown that, in the first place, the judgement-stroke is a truth-operator, not an illocutionary force marker; it expresses that something is true, not that the speaker asserts or judges something as true. This reading is forced upon us by the following pieces of evidence. With regard to the role of ‘ ‘ in his first logical system, which is formulated in his early Begriffsschrift (1879), Frege explicitly says that this stroke is used to “present” a content as “true” or as a “fact” (als wahr hinstellen). Thus, ‘ ⊢ 2+3=5’ expresses that the content that 2+3=5 is true, whereas ‘— 2+3=5’ expresses the same content without at the same time expressing that it is true. In his later writings, Frege explicitly states that in his revised system a sentence of the form ‘ ⊢ Δ’ expresses that the truth-value denoted by ‘— Δ’ “is the True”. In “Funktion und Begriff” (1891), e.g., he explains the difference between expressions like ‘ ⊢ 2+3=5’ and ‘— 2+3=5’ as follows: by writing ‘ ⊢ 2+3=5’, we are not “just writing down a truth-value”, as in ‘— 2+3=5’, but also at the same time “saying that it is the True”. These explanations suggest very strongly that the judgement-stroke is an operator whose primary linguistic function is to express the truth of a thought: it expresses that the truth-value of a thought (or judgeable content) is the True, not that the speaker judges or asserts the thought to be true.

Frege’s motive for introducing a truth-operator into his ideal language derives from his quest to distinguish syntactically between two linguistic acts whose difference is hidden by the surface grammar of natural language: the mere expression of a thought (“predication”), and

---

7 For a detailed defense of this interpretation, see Greimann, 2000a, pp. 215-25.
the presentation of a thought as a fact ("assertion"). When, for instance, a speaker poses the question, ‘Is 3 prime?’, or asserts the conditional sentence, ‘If 3 is prime, then the successor of 3 is not prime’, s/he expresses the thought that 3 is prime without at the same time presenting this thought as a fact. When, on the other hand, the speaker asserts, ‘3 is prime’, s/he does present the thought expressed as a fact. In Frege’s view, this feature of natural language is logically misleading, for the following reason:

In the formula ‘(2>3) ⊃ (7²=0)’ a sense of strangeness is at first felt, due to the unusual usage of the signs ‘>’ and ‘=’. For usually such a sign serves two distinct purposes: on the one hand it is meant to designate a relation (eine Beziehung bezeichnen), while on the other hand it is meant to assert the holding of this relation between certain objects (das Stattfinden der Beziehung zwischen gewissen Gegenständen behaupten). Accordingly it looks as though something false (2>3, 7²=0) is being asserted in that formula – which is not the case at all. That is to say, we must deprive the relational sign of the assertive force with which it has been unintentionally invested.

The important point is that, while in the conditional sentence ‘(2>3) ⊃ (7²=0)’ the relational sign ‘>’ is used merely to denote a relation, in ‘2>3’ it is used in addition to express the satisfaction (or “holding”) of this relation by the numbers 2 and 3, i.e. the truth of the predication that 2>3.

In natural language, truth is expressed, according to Frege, not by a special sign, but by the “form of the assertoric sentence”, which is considered by him to be the primary truth-operator of natural language, i.e. the primary and effective means of presenting a thought as a truth. In his fragment “Logik”, dating from 1897, he writes:

---


It is really by using the form of the assertoric sentence that we express truth [womit wir Wahrheit aussagen], and to do this we do not need the word 'true'. Indeed, we can say that even where we use the locution 'it is true that ...' the essential thing is really the form of the assertoric sentence.\footnote{Frege, 1983, p. 140, partly my translation. Compare also Frege, 1983, pp. 251-2.}

In Frege’s ideal language, the ascription of truth is divorced from the form of the assertoric sentence and wedded to a special sign, the judgement-stroke, which is considered by him as a \textit{sui generis} sign.\footnote{Cf. \textit{Grundgesetze}, vol. 1, §26.} The role of this stroke is identical with the role of the form of the assertoric sentence in natural language: its task is to express the satisfaction of concepts and relations by objects and, more generally, the truth of thoughts.

In §4 of \textit{Grundgesetze}, Frege reduces the satisfaction of concepts and dyadic relations by objects to the notion of truth by means of the following definitions: “We say that the object $\Gamma$ stands in the relation $\Psi(\xi, \zeta)$ to the object $\Delta$ if $\Psi(\Gamma, \Delta)$ is the True, just as we say that the object $\Delta$ falls under the concept $\Phi(\xi)$ if $\Phi(\Delta)$ is the True”. Note that this reduction of satisfaction to truth is the reversal of Tarski’s reduction of truth to satisfaction: for Frege, the notion of truth is conceptually prior to the notion of satisfaction, because to say of an object $\Delta$ that it satisfies the concept $\Phi$ is to say that the value of $\Phi$ for $\Delta$ as argument \textit{is the True.} Given this order of explanation, Tarski’s definition of truth for the language of the class calculus in terms of satisfaction must be regarded to be entirely circular.

The satisfaction of concepts and relations by objects is accordingly expressed in Frege’s ideal language in terms of truth, i.e. by means of ‘ ‘. Thus, according to §§4 and 5 of \textit{Grundgesetze} Vol. I, by

\begin{enumerate}
\item Cf. \textit{Grundgesetze}, vol. 1, §26.
\end{enumerate}

writing ‘\( \neg \Phi(\Delta) \)’, we are saying that the truth-value of ‘\( \neg \Phi(\Delta) \)’ is the True, and hence that \( \Delta \) actually falls under \( \Phi \). This notation achieves the envisaged separation of expressing a thought and presenting it as a fact: while ‘\( \neg \Phi(\Delta) \)’ merely expresses that \( \Delta \) is \( \Phi \) without expressing that this is also true, ‘\( \neg \Phi(\Delta) \)’ expresses both.

Since the formation rules of *Grundgesetze* acknowledge expressions like ‘\( \neg (2=2)=(3=3) \)’ as well-formed, the syntax of Frege’s ideal language seems to be characterized by a strange assimilation of sentences to singular terms which is based on the absurd assumption that sentences are a species of singular terms, namely, singular terms denoting a truth-value.\(^{12}\) There is, however, considerable evidence that the expressions of the form ‘\( \neg \Delta \)’ are to be understood exclusively as singular terms. For, both in *Grundgesetze* and in “Funktion und Begriff”, Frege explicitly says that from his peculiar use of the identity-sign it can already be seen that he is only “designating” a truth-value when he writes down an equation like ‘\( 2+3=5 \)’: just as ‘2^2’ designates the square of 2, so too ‘2+3=5’ designates “the truth-value of: that 2+3=5”.\(^{13}\) As a consequence, by writing down ‘\( 2+3=5 \)’ in the ideal language, we are not asserting that 2+3=5, but merely designating the truth-value of: that 2+3=5, “without saying which of the two it is”. Even if we wrote ‘\( (2+3=5)=(2=2) \)’ and presupposed that it was known that it is true that 2=2, Frege holds, we would not thereby “have asserted” that 2+3=5, but have “merely designated” the truth-value of: that ‘\( 2+3=5 \)’ refers to the same thing as ‘\( 2=2 \)’, without saying which of the two it is. Therefore, Frege explains, a

\(^{12}\) Cf. Dummett, 1973, pp. 7, 184, 196.

\(^{13}\) Cf. *Grundgesetze*, vol. I, §§2 and 5 and Frege, 1891, pp. 136-7 or the translation in Beaney, 1997, p. 142. The phrase ‘the truth-value of: that \( p’ \) is a direct translation of Frege’s German locution ‘der Wahrheitswert davon, dass’. It may be read as ‘the truth-value of the thought that \( p’ \).
A special sign is needed in order to be able to assert something as True, and this is the task of the judgement-stroke: by writing ‘$\vdash 2+3=5$’, we are saying that the truth-value of: that $2+3=5$ is the True.\textsuperscript{14}

These explanations imply that in the ideal language an expression like ‘$2+3=5$’ is not a sentence, but a singular term whose English counterpart is ‘The truth-value of: that $2+3=5$’.\textsuperscript{15} For clarity’s sake, let us write ‘$2+3=5$’ when we want to refer to ‘$2+3=5$’ considered as an expression of the ideal language, and ‘$2+3=5$’ when we want to refer to ‘$2+3=5$’ considered as an expression of natural language. Note that, according to Frege’s truth-conditional semantics, the sentence ‘$2+3=5$’ and the singular term ‘the truth-value of: that $2+3=5$’ do indeed have the same sense, because the conditions under which they denote the True are the same.\textsuperscript{16} Because of these features, the singular term can be used to express the thought that $2+3=5$ without at the same time expressing that its truth-value is the True.

Hence, the transformation of ordinary sentences into sentences of the form ‘$\vdash \Delta$’ actually achieves the intended separation of predication and assertion.\textsuperscript{17} In the ideal language, the predicates are “deprived” of the assertoric force simply by using them as functional signs that yield

\begin{itemize}
\item \textsuperscript{14} Cf. \textit{Grundgesetze}, vol. I, §§2, 5 and Frege, 1891, pp. 136-7 or the translation in Beaney, 1997, p. 142.
\item \textsuperscript{15} This agrees with Simons’ view that for Frege an equation like ‘$2^2=2+2$’, considered as an expression of the logical language, is “not a sentence, but a name of a truth-value”, and that Frege “takes care to express such things using a nominal expression like, ‘$2^2$’s being equal to $2+2$’ or ‘the truth-value thereof, that $2^2$’” (1996, p. 287).
\item \textsuperscript{16} The thought that $2+3=5$ may be considered as a “mode of representing” the True.
\item \textsuperscript{17} Cf. Frege, 1891, p. 32 and 1983, p. 214.
\end{itemize}

complex proper names of truth-values when applied to other singular terms.

Following Frege’s hints in *Grundgesetze* and “Funktion und Begriff”, we may translate his ideal language into English as follows:

1. ‘⊢Δ’ translates as ‘The truth-value — Δ is the True’.
2. ‘—Δ’ translates as ‘The truth-value of: that Δ is the True’.
3. ‘Φ(ξ)’ translates as ‘the truth-value of: that Φ(ξ)’.
4. ‘─┬─ξ’ translates as ‘the truth-value of: that ξ is not the True’.
5. ‘Φ(α)’ translates as ‘the truth-value of: that for every object x the truth-value of: that Φ(x) is the True’.

By these rules, the complex expression ‘(3=3)=(2=2)’ translates into English as ‘the truth-value of: that (the truth-value of: that 3=3)=(the truth-value of: that 2=2)’, and ‘⊢2’, as ‘The truth-value of: that (the truth-value of: that 2 is the True) is the True’.18

The semantics of Frege’s ideal language is characterized by his demand to “explain each expression with respect to its reference completely”.19 Frege rejects the recognition of vague concepts and of ambiguous and partial functions in his ideal language because this would violate tertium non datur.20 In his view, this law is just another form of the requirement that every concept should have a sharp

---

18 For a detailed defense of this reconstruction, see Greimann, 2000a, pp. 225-33.
19 *Grundgesetze*, vol. II, §57. See also *Grundgesetze*, vol. I, §§28 and 29, and *Grundlagen*, §§84, 85 and 74.
boundary, because according to the law any object either falls under a given concept or does not fall under it – *tertium non datur*.

In order to determine the *Bedeutung* of a one-place predicate completely, we must determine, with respect to any object \( x \), what the conditions are for the predicate to be true of \( x \). Analogously, in order to determine the *Bedeutung* of a functional sign ‘\( g(\ ) \)’ completely, its semantic interpretation has to be completed by suitable stipulations that jointly fix the truth-conditions of ‘\( x=g(y) \)’ for all given objects \( x \) and \( y \). Moreover, the stipulations must ensure that the value of \( g \) is unique for each object \( x \) as argument, i.e., they must rule out that there is a plurality of objects meeting the criterion of being the value of \( g(y) \), because otherwise the corresponding singular terms would be referentially ambiguous.\(^{21}\) To achieve this, the extension of the concept of being identical with the value of \( g(y) \) must be sharply delimited for any given object \( y \) as argument. Finally, to determine the *Bedeutung* of a singular term ‘\( a \)’ completely, we have to complete its semantic interpretation by stipulations that determine the criteria for being the *Bedeutung* of \( a \), i.e., we have to fix the truth-conditions of ‘\( a=y \)’ for any given object \( y \).

Since the usual truth-functions denoted by ‘and’ and ‘or’ are defined only for the truth-values as arguments, they must be extended by defining them for all objects. This is the task of the Horizontal, which denotes the function the truth-value of: that \( x \) is identical with the True assigning the True to the True and the False to every other object.\(^{22}\) This function is used in Frege’s system to convert the usual truth-functions into total functions, in order to fulfill the demand to determine the

---


The extended truth-functions of *Grundgesetze* accordingly result from “melting” the usual truth-functions with $\neg x$. Thus, the function $\neg x$, for instance, which is Frege’s extended negation, is the function the truth-value of: that (the truth-value of: that $x$ is identical with the True) is the False, which is coextensive with the function the truth-value of: that $x$ is not identical with the True, it assigns the False to the True, and the True to any other object. The prima facie nonsensical expression ‘$\neg\neg 2$’ can accordingly be translated into plain English as ‘The truth-value of: that 2 is not identical with the True’. Its *Sinn* is the thought that the True is the truth-value of: that 2 is not identical with the True, and its *Bedeutung* is the True, because 2 is not identical with the True. In Frege’s first logical system, the horizontal stroke is called the “content-stroke”. Its primary task is there to distinguish between contents that can be asserted (or judged) and those that cannot, where the condition for being assertible (or judgeable) is to express a thought. Thus, the content of ‘$2+3=5$’, but not that of ‘2’, is assertible. To account for this pragmatic difference, the formation rules of the system acknowledge an instance of the form ‘$\vdash \Delta$’ as a well-formed sentence only if ‘$\Delta$’ is a judgeable content.

Dummett holds that, in Frege’s mature system, the distinction between assertable and non-assertible contents is lost, because sentences are considered to be a species of names. Thus, according to

---

23 As Simons puts it, the Horizontal “has the role of eliminating third objects from consideration as values by assimilating them to the False (…)” (1996, p. 292).

the formation rules of this system, ‘⊢2’ is a well-formed sentence, although the sense of ‘2’ is not assertible. However, as Burge has shown, this picture is based on a “serious misunderstanding”: the system of Grundgesetze does justice to the difference between assertable and non-assertible contents, by stipulating that something can be asserted only if it is expressed by an expression of the form ‘―Δ’.

Since such an expression denotes a truth-value, its sense must be a thought, namely, the thought that the conditions under which the expression denotes the True are fulfilled. So it is built also into Frege’s mature system, if only discretely, that only thoughts can be asserted. Since this restriction is embodied in the formation rule that a sentence is well-formed only when it has the form ‘⊢Δ’, it follows that, in the system of Grundgesetze, the Horizontal plays also an important pragmatic role.

According to Burge, the Horizontal plays also the role of a means of expressing truth in Frege’s system. He writes:

> The horizontal expresses the notion of truth in Frege’s system. It means ‘is the True’ or ‘is the truth’ or ‘is truth’. It is present in the formulation of every assertion. It may accompany any declarative sentence without adding to its sense. The concept denoted by the Horizontal is the only one within Frege’s logic that meets the condition set by his redundancy conception of truth. Frege alludes to this condition without fanfare in Section 5 in Basic Laws, where he notes the equivalence:

\[ \Delta = \neg \Delta \]

where ‘Δ’ varies over truth values. The import of this condition comes [sic] clear if one sees sentences as substituting for ‘Δ’, reads ‘=’ (as in such cases

---


However, according to Frege’s own explanations, truth is expressed in his system by the judgement-stroke, not by the Horizontal. The expression ‘— 2+3=5’ cannot be used to express the identity of the truth-value of: that 2+3=5 with the True, but merely to designate the truth-value of: that the truth-value of: that 2+3=5 is identical with the True, without saying which of the two it is. To achieve the latter, we must use the judgement-stroke, which must therefore be regarded as the real device of expressing the identity of a truth-value with the True in Frege’s system. Thus, ‘⊢ 2+3=5’ expresses that the truth-value of: that 2+3=5 is the True.

Pace Burge, it is also highly questionable that “[t]he horizontal expresses the notion of truth in Frege’s system.” For, in his posthumous writings, Frege explicitly says that the intended notion of truth with which he is concerned is indicated, not by the word ‘true’, but by the assertoric force with which assertoric sentences are normally uttered:

[...] there is no doubt that the word ‘beautiful’ actually does indicate the essence of aesthetics, as does ‘good’ that of ethics, whereas ‘true’ only makes an abortive attempt to indicate the essence of logic, since what logic really is concerned with is not contained in the word ‘true’ at all but in the assertoric force with which a sentence is uttered.  

By the “assertoric force” Frege obviously means the truth-claiming aspect of assertions. Since, in his ideal language, the truth-claim is syntactically

---


represented by the judgement-stroke, not by the Horizontal, the former must be considered to express the notion of truth is Frege’s system. According to this notion, to be true is simply to be the case (or “a fact”), because this is what the judgement-stroke expresses.

Note also that the Horizontal does not correspond to the predicate ‘is true’, when the latter is understood along the lines of Frege’s redundancy theory of ‘true’. For, given the redundancy of ‘true’, sentences like ‘The thought that 2+3=5 is true’ and ‘2+3=5’ have the same sense and the same denotation, whereas singular terms like ‘— the thought that 2+3=5’ and ‘2+3=5’ have neither the same sense nor the same denotation: while ‘2+3=5’ denotes the True, ‘— the thought that 2+3=5’ denotes the False, because the True is not identical with the thought that 2+3=5.29

To understand the semantics of the Horizontal more closely, we must distinguish between the truth-predicate ‘x is true’ and the “truth-connective” ‘it is true that p’, as I would like to call it.30 The semantic difference is that the truth-predicate denotes a function from thoughts (or sentences) to truth-values, whereas the truth-connective denotes a function from truth-values to truth-values. The Horizontal is clearly derived from the truth-connective, not from the truth-predicate. The sole difference between the Horizontal and the truth-connective is that the function denoted by the truth-connective is defined only for truth-values as arguments, while the function denoted by the Horizontal is defined for all objects. Since the truth-connective is powerless to generate any inconsistencies, it is to be expected that the Horizontal does not generate any inconsistencies, either.

29 Cf. Frege, 1892, p. 150 or the translation in Beaney, 1997, p. 158.
30 See also Thiel, 1983, pp. 299-300.
3. THE INNOCENCE OF THE HORIZONTAL

Prima facie, Frege’s definition of the Horizontal conflicts with his own thesis of the undefinability of truth according to which the concept of truth is so fundamental that it cannot by explained in a non-circular way, i.e. without presupposing a prior grasp of it. This conflict, however, dissolves on closer examination, because the real truth-operator of Frege’s system is the vertical, not the horizontal stroke. His undefinability thesis accordingly refers to truth, considered as what is expressed by the judgement-stroke or by the form of the assertoric sentence, not to truth, considered as the concept expressed by the Horizontal. On this assumption, Frege’s argument for the undefinability of truth may be reconstructed as follows. When we want to explain truth by a definition of the form ‘x is true if and only if x is F’, we must presuppose that the hearer already understands what it means that F actually applies to x, i.e., that the truth-value of: that x is F is the True. To understand our definition, the hearer must already understand what it means for an object to fall under a concept, and hence what it means that something is true. This order of understanding immediately implies that truth cannot be explained in a non-circular way, i.e. without presupposing a prior grasp of it.

To show that the introduction of the Horizontal does not conflict with Tarski’s theorem of the undefinability of truth either, let me begin with a brief description of Aczel’s opposite view. Roughly speaking, by a “Frege structure” Aczel means a model of the lambda calculus, the

---

31 Cf. Frege, 1918, p. 344 and 1983, pp. 139-4 or the translations in Beaney, 1997, pp. 327 and 228.


elements of the model being called “objects”, together with a collection of objects called “propositions” and a subcollection of propositions called “truths”. It includes, in addition, a list of the usual logical constants, each satisfying a logical schema that expresses how a proposition is built up using the logical constants and further specifies truth conditions for the resulting proposition. With regard to the structure underlying the ideal language of *Grundgesetze*, Aczel’s central contention is this:

The underlying structure of the formal language in Frege’s *Grundgesetze* is essentially that of a Frege structure satisfying the following additional condition. There are exactly two propositions, “the true” and “the false”, with the former being of course the only truth. Russell’s paradox shows us that no Frege structure can satisfy this extra condition. I locate the flaw in Frege’s condition by formulating the notion of “internal” definability for collections of objects on a Frege structure. Frege’s condition implies that there is an internal definition of the collection of truths, and the argument of Russell’s paradox shows that there can be no such definition. This is clearly related to the classic result of Tarski on truth.\(^{33}\)

Aczel’s notion of “internal” definability is defined as follows. If \(c\) is a collection of objects on a Frege structure \(F\) and \(C\) is a propositional function in \(F\), then \(C\) internally defines \(c\) if for all objects \(a\): \(C(a)\) is true iff \(a\) is in \(c\), where \(F_1\) is a collection of one-place functions belonging to the structure \(F\). Given this notion, the horizontal stroke schema, which Aczel formulates as “For any object \(a\), --- \(a\) is a proposition such that --- \(a\) is true iff \(a\) is true”, asserts that the propositional function denoted by the Horizontal internally defines the collection of truths. Since the argument of Russell’s paradox shows that this is not possible, Aczel

---

\(^{33}\) Aczel, 1980, p. 32.

argues, we have recaptured Tarski’s result of the undefinability of truth for Frege structures.  

Now, I do not want to dispute Aczel’s claim that he has recaptured, by means of Russell’s paradox, Tarski’s undefinability result for his Frege structures. Rather, I wish to question his assumption that the underlying structure of Frege’s ideal language is that of a Frege structure satisfying the additional condition that there are exactly two propositions, the True and the False, with the True being the only truth.  

This assumption cannot be sustained, for the following reasons. 

[1.] On Frege’s view, the True is not a truth or a true “proposition”. It does not really make sense to say of the True that it is true, or of the False that it is false, because the truth-values are not truth-bearers at all. Although, for instance, — 2+3=5 is the True, it is not a truth or a true proposition. The True is the Bedeutung of ‘— 2+3=5’, and the truth expressed by ‘— 2+3=5’ is the Sinn of this expression, i.e. the thought that 2+3=5. So, the collection of truths in Frege’s system does not consist of the True, considered as the sole truth, but of true thoughts (or sentences).

[2.] In Frege’s system, the characteristic functions denoted by the predicates are not, as Aczel assumes, propositional functions, i.e. functions “all of whose values are propositions”. For, the values of these functions are truth-values, not propositions. Thus, the value of the function x is white for snow as argument is, not a proposition, but a truth-value. The opposite assumption rests on a confusion of the Sinn of sentences with their Bedeutung.

[3.] Aczel writes:

The Grundgesetze has a propositional function in $F_1$ called the horizontal stroke. When applied to an object $a$ it yields a proposition — $\bar{a}$ which is the true if $a$ is the true and is the false if $a$ is not the true. So it satisfies the schema: Horizontal stroke. For any object $a$, $\bar{a}$ is a proposition such that $\bar{a}$ is true iff $a$ is true.\textsuperscript{37}

This conclusion is faulty, because the predicates ‘is true’ (or ‘is a truth’) and ‘is the True’ are not coextensive. Although, for instance, the thought that snow is white is an object that is true, the value of the function $\bar{x}$ for this object as argument is the False, because the True is not identical to the thought that snow is white. Hence, Frege’s system does not satisfy the scheme “For any object $a$, $\bar{a}$ is a proposition such that $\bar{a}$ is true iff $a$ is true”.\textsuperscript{38}

[4.] Tarski’s theorem of the undefinability of truth refers to semantic truth-predicates satisfying the scheme

$$(T) \quad 'p' \text{ is } T \text{ if and only if } p.$$ 

But, in Frege’s system, the predicate satisfying $(T)$ is, not the Horizontal, but the predicate ‘denotes the True’ (bedeutet das Wahre), which may therefore be regarded as the proper semantic truth-predicate of the system. Let me explain. According to Frege’s analysis of ‘true’, the sense of this word is such that it does not make any essential contribution to the senses of the sentences in which it occurs, i.e., sentence-pairs like ‘Snow is white’ and ‘The sentence ‘Snow is white’ is true’ express exactly the same thought.\textsuperscript{39} This analysis implies that the predicate ‘is true’ cannot be used to set up his truth-conditional semantics, which is based on the idea that to know the sense of a sentence is to know what must be the

---

\textsuperscript{37} Aczel, 1980, p. 40.

\textsuperscript{38} A similar objection is made in Kutschera, 1989, p. 25, footnote.

case for the sentence to be true. For, his analysis of ‘true’ implies that an explanation of the form ‘The sentence ‘Snow is white’ is true if and only if snow is white’ does not contain any semantic information, but is synonymous with the tautology ‘Snow is white if and only if snow is white’.

From this some commentators have concluded that there is a conflict between Frege’s truth-conditional approach to semantics and his redundancy theory of ‘is true’. This would be a just complaint if Frege really used the predicate ‘is true’ to explain the truth-conditions of sentences. But, as a matter of fact, he does not use ‘is true’ to this end, but the predicate ‘denotes the True’ (bedeutet das Wahre), which is his substitute for a semantic truth-predicate. In contrast to ‘is true’, this predicate is not redundant, and, in contrast to the judgement-stroke, it is a genuine predicate expressing a genuine property, namely, the semantic property of denoting the True.

[5.] Tarski’s theorem of the undefinability of truth proves, roughly speaking, that the language $L_{\alpha}$ of arithmetic cannot contain its own

---

40 Ricketts, 1986 maintains that Frege’s system does not contain anything like modern truth-conditional semantics, as it has come to be understood in the light of Tarski’s work of truth. In Heck, 1999 it is shown that this view must be rejected.


42 Ironically, the conflict ascribed to Frege actually applies to Tarski’s system; see Etchemendy, 1988 and Soames, 1999, chap. 4.


44 In Frege’s system, one must distinguish between three different “truth-items”: (i) the semantic property denoted by the predicate ‘bedeutet the True’, (ii) the object denoted by the abstract singular term ‘the True’, and (iii) truth, considered as what is expressed by the form of the assertoric sentence, which is neither an object nor a property, but an element of a third category to which also “saturation” belongs.
truth-predicate, i.e., it cannot contain an open sentence applying to all and only the (Gödel-numbers of) true sentences of $L_{a}$. The core idea of Tarski’s proof is to show that the contrary assumption that $L_{a}$ does contain its own truth-predicate would entail a contradiction that is generated by the presence of Liar-sentences in the language.\textsuperscript{45} Since Frege’s ideal language includes the language of arithmetic, there can be no doubt that it is rich enough to provide a means for speaking about its own sentences (or about their Gödel-proxies). Nevertheless, there are good reasons to doubt that in Frege’s language the presence of Liar-sentences would really generate a contradiction. To see this, consider the sentence

$$(L) \vdash \text{the truth-value of: that (} L \text{) does not denote the True, which is the counterpart of the liar-sentence ‘This sentence is not true’ in Frege’s system. According to Frege’s principle of compositionality, (} L \text{) has a truth-value only when the singular term}

$$(S) \text{the truth-value of: that (} S \text{) does not denote the True}$$

has a denotation. If (} S \text{) has a denotation at all, then this must be the True or the False; \textit{tertium non datur}. If (} S \text{) denotes the False, then (} S \text{) must also denote the True, because in this case the truth-value of: that (} S \text{) does not denote the True is the True. If (} S \text{) denotes the True, then (} S \text{) must also denote the False, because in this case the truth-value of: that (} S \text{) does not denote the True is the False. So, if (} S \text{) has a denotation at all, it must denote both the True and the False. In this case, (} S \text{) is ambiguous, because it refers to two truth-values, just as the term ‘the square root of 9’ is ambiguous, because it refers to two numbers.}

Now, according to Frege’s criteria for having a \textit{Bedeutung}, a name has a \textit{Bedeutung} only when its denotation is determined completely. Since the concepts and functions expressed by predicates and functional signs

\textsuperscript{45} For a lucid explanation of Tarski’s proof, see Soames, 1999, chap. 5.
of natural language are vague and ambiguous, these signs have, strictly speaking, no *Bedeutung* at all!" Thus, in his “Ausführungen über Sinn und Bedeutung”, Frege writes:

> If it is a question of the truth of something – and truth is the goal of logic – we also have to inquire after *Bedeutungen*; we have to throw aside proper names that do not designate or name an object, though they may have a sense; we have to throw aside concept words that do not have a *Bedeutung*. These are not such as, say, contain a contradiction – for there is nothing at all wrong in a concept’s being empty – but such as have vague boundaries. It must be determinate for every object whether it falls under a concept or not; a concept word which does not meet this requirement on its *Bedeutung* is *bedeutungslos*.[i.e. has no *Bedeutung*].

Similarly, an ambiguous singular term like ‘the square root of 4’ has, according to Frege’s criteria, no *Bedeutung* either. Since, as we have seen, (\(\sqrt{4}\)) has either two denotations or none, this implies that (\(\sqrt{4}\)) does not fulfill the criteria for having a *Bedeutung*. By the principle of compositionality, it follows that the Liar-sentence (L) is neither true nor false but belongs to the realm of fiction, because it speaks about a truth-value that does not exist. The reason is that, in Frege’s system, the Liar-sentences are to be understood as statements about truth-values, whereas in standard systems these sentences are construed as statements about sentences. This is a consequence of Frege’s analysis of the logical form of sentences, according to which every sentence has the form of an equality expressing the identity of a truth-value with the True.

---

46 See also *Nachgelassene Schriften*, p. 195.
49 A curious implication of this analysis is that even the skeptic is committed to acknowledge the truth-values; see Frege, 1892, p. 149 or the translation in Beaney, 1997, p. 158.
From this we may finally conclude that the presence of a Liar-sentence would not generate an inconsistency in Frege’s system, but a truth-value gap. Since Tarski’s proof turns at some crucial point on the assumption that a Liar-sentence like $\langle L \rangle$ is either true or false, it cannot be used to prove the inconsistency of Frege’s system.50

[6.] As Frege explicitly notes, the use of the Horizontal can be eliminated in his system, because the Horizontal can be reduced to the identity sign ‘=’ in terms of the explicit definition

\[(D)\] For any object $a$, $a \equiv a = (a = a)$, which may be translated into English as ‘For any object $a$, $a \equiv$ the truth-value of: that $a = (the\ truth-value\ of: \ that\ a = a)$’.51 This definition implies that the introduction of the Horizontal, taken by itself, cannot be blamed for the inconsistency of the system. If there is any conflict between the system and Tarski’s result on the undefinability of truth, it must have a deeper source. Two possibilities deserve closer examination.

Aczel assumes that the inconsistency of Frege’s system derives from the fact that it has exactly two propositions: the True and the False. What he has in mind, is presumably this. According to Frege’s first system, the values of propositional functions are “circumstances” (“Umstände”), not truth-values.52 Thus, the value of the function $\text{White}(x)$ for snow as argument is the circumstance that snow is white, and the value for salt as argument is the circumstance that salt is white.

---

50 Note that the Strengthened Liar

\[\langle SL \rangle \vdash \text{the truth-value of: that $\langle SL \rangle$ denotes neither the True nor the False}\]

also belongs to fiction, because the definite description occurring in it does not apply to a unique truth-value.

51 Since the truth-value of: that $a = a$ is always the True, $a = a$ = the truth-value of: that $a = a$ if and only if $a$ is the True.

52 Cf. Begriffsschrift, §§2 and 9. Circumstances are “Russellian” or “singular” propositions.
Given this structure, the True cannot be defined in terms of the identity relation, because an equation like ‘White(snow) = White(snow)’ does not denote the True, but the circumstance that the circumstance that snow is white is identical to itself. Consequently, the equation ‘White(snow) = (White(snow) = White(snow))’ does not express the truth of the circumstance that snow is white, but the identity of the circumstance that snow is white with the more complex circumstance that the former circumstance is identical to itself. In his mature period, on the other hand, Frege construes the values of propositional functions as truth-values. As a consequence of this, his mature system has only two propositions, i.e. two values of propositional functions, namely, the True and the False. Since this revision allows us to define the collection of true propositions as the collection of values of the propositional function \( x = (x = x) \), it must be the root of the inconsistency of this system.

However, as we have seen, the collection of true propositions in Frege’s mature system must not be confused with the collection of “true truth-values”, i.e. with the unit-set of the True. Since the definability of the collection whose sole element is the True does not imply the definability of the collection of truths, Aczel’s diagnosis is not really persuasive.

Note, moreover, that the True itself cannot be defined in terms of function denoted by the Horizontal, because the latter is explained in terms of the True. Hence, if Frege’s identification of the True with the value-course of this function in §10 of *Grundgesetze* Vol. I is considered to be a definition of the True, this definition would be entirely circular, because it amounts to an identification of the True with its own unit-set.\(^{53}\)

Christian Thiel (1983) has made the attempt to verify Aczel’s suspicion that the Horizontal is responsible for the inconsistency of

\(^{53}\) This observation is also made in Simons, 1996, p. 288. For a reconstruction of Frege’s identification, see Greimann, 2000b, pp. 134-6.
Frege’s system. He argues that Frege’s critical mistake was to construe the Horizontal as a functional sign that is defined, not only for sentences, but also for all other names, for, what renders the scheme \((H)\) as an internal definition of truth is not the Horizontal as such, but its application to arbitrary objects including the truth-value names, i.e. sentences.\(^{54}\) Thiel’s conclusion is that, in the end, it is the ontology of Frege’s system, not the Horizontal, that falls prey of Tarski’s undefinability result.\(^{55}\) Although this diagnosis, as I shall try to show below, is perfectly correct, the location of the source of the inconsistency remains unsatisfactory. If I understand him correctly, Thiel holds that Frege’s critical mistake was to construe sentences as a species of singular terms, namely, as names of truth-values. As a consequence of this, Frege construed the truth-values as a species of objects, not as entities *sui generis* that are neither objects nor functions (including concepts and relations). If Frege had not assimilated the sentences to names, then he had not to define the function denoted by Horizontal for all objects, but only for the truth-values as arguments.

However, as we have seen, the sharp delimitation of the function — \(a\) does not render the Horizontal as a Tarskian truth-predicate, because the Horizontal does not apply to any truth-bearer. Hence, the source of the inconsistency cannot reside in the sharp delimitation of the function denoted by the Horizontal.

### 4. THE INCOMPATIBILITY WITH CANTOR’S POWER SET THEOREM

The main result of Tarski’s inquiry into the formal structure of truth is that there can be no coherent notion of all truths. In his


important book *The Incomplete Universe* (1991), Patrick Grim has tried to recapture this result by means of a simple Cantorian argument, which runs as follows.\(^{36}\) Suppose there is a collection of all truths \(T\), and consider all subsets of \(T\), i.e. all elements of the power set \(P(T)\). To each element \(x\) of \(P(T)\) there will correspond a truth – e.g., the truth that \(x\) will belong or not belong to \(P(T)\). The number of these truths must be larger than the number of elements of \(T\), because, by Cantor’s power set theorem, the power set of any set must be larger than the original set. Hence, there will be more truths than there are members of \(T\). From this it follows that \(T\) is not complete, and hence that \(T\) is not the collection of all truths.

Now, Frege is committed to assume that the universe of discourse of his logical system contains all objects, because he construes the laws of logic as the most general laws applying to all objects whatsoever, and because he demands the complete determination of the *Bedeutung* of every sign.\(^{37}\) Since Frege acknowledges thoughts as real entities, he is further committed to assume that the universe of discourse of his ideal language includes all truths, i.e. all true thoughts. Let \(U\) be Frege’s universe of discourse, and \(C\) be the subcollection of the truths. To each subset of \(C\) there will be a true thought, for instance, the thought that this subset belongs the universe of discourse of the system. By Cantor’s theorem, the number of these truths must be larger than the number of elements of \(C\). Hence, \(C\) cannot be the collection of all truths.


\(^{37}\) In §65 of *Grundgesetze*, vol. II, Frege explicitly rejects the procedure to extend the universe of discourse step by step. It might be objected that, in §10 of *Grundgesetze*, Vol. I, Frege presupposes that the domain of his system is restricted to value-courses and truth-values. In Greimann, 2000b, pp. 137-8, I have sketched an interpretation that dissolves this apparent conflict.
The conclusion to be drawn is that Frege’s system becomes inconsistent when the implicit claim that its universe of discourse includes all truths is made explicit: there can be no language whose variables range over all truths. Regarded from this point of view, the critical mistake in Frege’s theorizing about truth was the ontological positing of truths, i.e., the ontological acknowledgement of true thoughts as real objects. To remove the inconsistency, Frege would have to give up the ontological principle:

\[(OP) \text{ If it is true that } p, \text{ then there is an entity } x \text{ such that } x \text{ is the bearer of the truth that } p.\]

If there were no truth-bearers at all – neither thoughts nor sentences nor utterances nor any other sort of truth-bearers –, no thing would be true, but there would still be something that is true. In this case, it would, for instance, still be true that there are no truth-bearers. Hence, it is not necessary to assume that for every truth there is a truth-bearer: that something is true does not imply that some thing is true. The lesson of Cantor’s theorem is that it is furthermore not consistent to assume that for every truth there is a truth-bearer.

When this diagnosis is correct, then it is indeed the ontology of Frege’s system that is in conflict with modern results of the indefinability of truth. What Frege did not see was that there can be no collection including all true thoughts, because otherwise there would have to be more true thoughts than there can possibly be.

REFERENCES


HECK, R. “Frege and Semantics”, internet-manuscript,  


