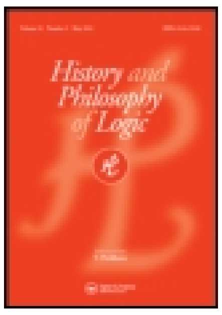


History and Philosophy of Logic
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/thpl20

The semantics of frege's grundgesetze

John N. Martin [a]
[a] Department of Philosophy, University of Cincinnati, 45221, Cincinnati, Ohio, U.S.A.

Version of record first published: 29 Mar 2007.

To cite this article: John N. Martin (1984): The semantics of frege's grundgesetze , History and Philosophy of Logic, 5:2, 143-176

To link to this article: http://dx.doi.org/10.1080/01445348408837068

HISTORY AND PHILOSOPHY OF LOGIC, 5 (1984), 143–176

The Semantics of Frege's *Grundgesetze*

JOHN N. MARTIN
Department of Philosophy, University of Cincinnati,
Cincinnati, Ohio 45221, U.S.A.

Received 22 September 1983

Quantifiers in Frege's *Grundgesetze* like $(\forall x)(Fx \,\&\, Gx)$ are not well-defined because the part $Fx \,\&\, Gx$ stands for a concept but the yoking conjunction is horizontalised and must stand for a truth-value. This standard interpretation is rejected in favor of a substitutional reading that, it is argued, both conforms better to the text and is well-defined. The theory of the horizontal is investigated in detail and the composite reading of Frege's connectives as made up of horizontals is rejected. The sense in which the *Grundgesetze* has a many-valued but classical logic is explained by proving that its propositional fragment (under the standard interpretation) and Bochvar's 3-valued logic are instances of the same metatheoretic methods. Historically, the paper argues for a more naive but well-defined reading of the text. Theoretically, it provides a formally adequate statement of that semantics, as well as developing the abstract metatheory which embraces Bochvar's language and the Fregean fragment.

1. The problem of undefined cases

Various interpreters of Frege's semantics have suggested that the reference relation as Frege envisaged it for his formal language of the *Grundgesetze der Arithmetik* is ill defined. The problem derives from the simultaneous application of two features of Frege's theory. First is the general principle that reference of part determines reference of whole, and second is what Frege has to say about the functions which combine to form wholes. If Frege's commitment to the first principle is read as a commitment to a categorial semantics in the modern sense, then functional expressions made up in turn of other functional expressions as parts, should have references that are defined for the references of the parts. But what Frege says is that the reference of the functional expression forming the whole is defined for only objects whereas the references of the parts in this case are functions, not objects.

My purpose in this paper is to explore the exact nature of this ill-definedness problem, and the evidence for the categorial interpretation that begets it. I shall argue that the categorial reading goes beyond the text in important ways and is really an attempt to fill out lacunae in Frege's exposition so as to produce a fully developed semantics in the modern sense. I shall also try to show that it is possible to fill out Frege's explicit semantics in a way that both conforms better to the text and avoids the undefinedness problem. The result is a claim that it is historically inaccurate to find in Frege a 'Fregean' semantics, as the term is understood in modern logic.

A large part of the discussion will concern Frege's horizontal operator. It will be shown how on the categorial interpretation, it can be understood to have essentially the same conceptual role as the truth-operator of Bochvar's three-valued logic, but that nevertheless Frege's own understanding of the horizontal was more modest.

2. Fregean semantics in modern logic

It has become the custom to describe a general approach to modern semantics as Fregean. Its main feature is a certain parallelism between syntax and semantics that is suggested in some of Frege's writings. In 'On sense and reference' in particular, he seems to assume a syntax in which parts determine wholes and a parallel semantic structure in which the references of syntactic parts determine the references of syntactic wholes. At one point he comes very close to using the modern notion of part and to explaining the influence of part on whole as one of determination (p.35/65: *for an explanation of the forms of citation of Frege, see the bibliography*). But most of the evidence for this interpretation derives indirectly from his argument that truth-values are the references of sentences (32/62*ff.*). He argues that just as substitution of co-referential parts preserves the reference of the whole when the parts are simple names, so does substitutivity of materially equivalent sentential parts. The reference of the whole in the first case is a truth-value, and so it should be in the second. If material equivalence is viewed as two sentences standing for the same truth-value, then a single principle will explain the substitutivity properties of both names and sentences: substitution of co-referring parts preserves the reference of the whole. Moreover, there seems to be no other candidate for the reference of sentences that would at the same time be both plausible and render the substitutivity principle true. Thus, Frege concludes, sentences refer to truth-values.

What is interesting in this argument is the presupposed link between substitutivity and parallelism. One seems to imply the other. A sentence can be compared to a brick wall. The wall is a whole composed of individual bricks as parts. The color of the bricks determines that of the whole and, as a result, substituting bricks of like color will not alter the color of the whole. Conversely, if color of parts alone did not determine the color of the whole, then substitutivity of parts of like color could not be relied upon to leave the whole unchanges. Frege seems to have some such metaphor in mind, but he does not say so in so many words and certainly comes nowhere near to laying out in a mathematically rigorous way the properties of structures needed to insure true parallelism and its related substitutivity. These properties have been quite generally defined in modern algebra, and one of the attractions of Frege's discussion is that he seems to make use of these ideas in a rudimentary way.

The modern notion of parallel structure is defined by reference to abstract structures consisting of a set followed by a series of functions on that set. If

$$S1 = \langle A, f_1, \ldots, f_n \rangle, \qquad S2 = \langle B, g_1, \ldots, g_n \rangle$$

are two structures of the same type in the sense that operations of the same rank take the same number of argument places, then a function from A into B is called a *homomorphism* from $S1$ to $S2$ iff for any i,

$$h(f_i(x_1, \ldots, x_m)) = g_i(h(x_1), \ldots, h(x_m)).$$

Such a function that is $1-1$ and onto is called an *isomorphism*. The notion of a function constructible in a structure $S = \langle A, f_1, \ldots f_n \rangle$ may be defined recursively:

$f_1, \ldots, f_n$ are constructible in S and if $g, h_1, \ldots, h_j$ are all functions constructible in S and

$$\{x_1, \ldots, x_m\} = \{x_q, \ldots, x_r, \ldots, x_s, \ldots, x_t, x_u, \ldots x_v\},$$

then k defined on A as follows is constructible in S:

$$k(x_1, \ldots, x_m) = g(h_1(x_q, \ldots, x_r), \ldots, h_j(x_s, \ldots, x_t), x_u, \ldots, x_v).$$

A relation $\equiv$ on A is said to be a *congruence relation* on S iff for any function k constructible in S,

$$k(x_1, \ldots, x_m) \equiv k(y_1, \ldots, y_m) \text{ iff } x_1 \equiv y_1 \,\&\ldots\&\, x_m \equiv y_m.$$

Parallelism determines equivalence under substitutivity in the sense that a homomorphism determined a straightforward congruence relation, and conversely a congruence relation determines an obvious homomorphism. If h is a homomorphism from $S1$ to $S2$ we define $\equiv$ as follows:

$$x \equiv y \text{ iff } h(x) = h(y)$$

Conversely, if B is the set of all $\equiv$-equivalence classes $[y] = \{y: y \equiv x\}$ for any element x of A and $g_i([x_1], \ldots, [x_m])$ is defined to be $[f_i(x_1, \ldots, x_m)]$, then the function h from A to B defined as $h(x) = [x]$ is a homomorphism form S to $<B, g_1, \ldots g_m>$.

In formal semantics these ideas are adapted to explicate the idea of a language with parallel syntactic and semantic structure. First a *syntax* is defined as a structure on a set of expressions with operations that map expressions into expressions in the manner of more traditional formation rules. A *semantics* is then defined as a structure of the same type on a non-empty set of semantic values or references. Finally, a *language* is defined as a pair consisting of a syntax and semantics of the same type. Intuitively, the operation corresponding to a formation function is a semantic rule that determines the reference of a whole formed by that formation operation given the references of its immediate parts as arguments. A *semantic interpretation* or *reference relation* is defined as any homomorphism from syntax to semantics. Usually the set of expressions of a syntax is defined as the closure under the syntactic operations of a set or series of sets of atomic expressions. In this case a reference relation is defined recursively by first stipulating the values of atomic expressions and assigning references to whole expressions by applying the corresponding semantic operation to the references of its immediate parts. If $S1$ and $S2$ as previously defined are respectively this kind of syntax and its parallel semantics, and R is a recursively defined reference relation between them, then the assignment of a reference to a molecular expression $f_i(x_1, \ldots, x_m)$ is determined from the references of the immediate parts and the corresponding semantic operation g_i according to the following rule:

$$R(f_i(x_1, \ldots x_m)) = g_i(R(x_1), \ldots, R(x_m)). \quad (1)$$

Thus, just as a syntactic operation is a formal version of a traditional formation rule, its corresponding semantic operation is a formal version of the traditional semantic rule used to interpret expressions formed by that rule in its clause in the recursive definition of the semantic value assignment. A language may also exhibit the property that every semantic operation is directly represented in the syntax by an expression that essentially refers to it. Such a language is known as a *categorial grammar* which for our purposes we may define as a syntax, a semantics of like type, and a set of recursively defined reference relations obeying (1) that in addition conform to the following constraint:

$$R(g_i(x_1, \ldots, x_m)) = R(x_m)(R(x_1), \ldots, R(x_{m-1})). \quad (2)$$

(Here the notation $f(x)(y) = z$ means $g(y) = z$ where $f(x) = g$.)

From (1) and (2) it follows that within a categorial grammar reference of a whole expression is determined completely by the references of its parts:

$$R(g_i(x_1, \ldots, x_m)) = R(x_m)(R(x_1), \ldots, R(x_{m-1})). \quad (3)$$

Such a grammar is sometimes further simplified syntactically. Since all complex expressions are assigned references in the same way, as explained in (3), there is no need to keep track of the syntactic differences among complex expressions. Indeed, formation rules may all take the same form. For example, they may all be simple concatenation rules taking the form:

$$f_i(x_1, \ldots, x_m) = x_m x_1 \ldots x_{m-1}. \quad (4)$$

We may then deduce a single principle of interpretation applicable to all expressions:

$$R(x_m x_1 \ldots x_{m-1}) = R(x_m)(R(x_1), \ldots, R(x_{m-1})). \quad (5)$$

Historically, categorial grammars were first studied by Ajdukiewicz, who was interested in their syntactic properties and did not provide semantic interpretations. The modern semantics with its structure parallel to syntax may fairly be said to find its inspiration in Frege to the extent that Frege too advocated a parallelism but of a less precise sort.[1] But to read Frege as intending a categorial semantics in the strict sense for the *Grundgesetze* is another matter.

1 Modern abstractions from Frege of homomorphic parallels between syntax and semantics can probably be traced to Rudolf Carnap, though clear algebraic expression of the idea seems to be somewhat later. It may be found in Wallace *1964*, 99–117, and throughout the work of Richard Montague. Abstractions from Frege to the additional constraints of a categorial grammar are found in Kaplan *1964*, Chapters I and II; Potts *1973* and *1977*; and Cresswell *1973*, esp. pp.5 and 75*ff*. The main root of categorial grammars, however, is independent of Frege and goes back to Ajdukiewicz. See Bar-Hillel *et alii 1964*.

3. The categorial interpretation

Commentators specifically interested in the precise statement of the semantics of the *Grundgesetze* either quite clear attribute to Frege a categoial semantics or ascribe to him general principles highly suggestive of a categorial framework. The explicit versions of the categorial interpretation all agree that the operators, both the connectives and the variable binding operators, are to be read as referring expressions standing for semantic operations on references. The reference of a whole made up of an operator is then computed in the manner of (3).

This interpretation rests largely on what Frege says about functional expressions. First of all is a claim about the syntax of the *Grundgesetze*. All complex proper names on this reading are understood to be generated by a formation rule that takes a functional expression with various sorts of place holders and an equal number of proper names as arguments and then produces the complex name by filling the positions of the place holders by the names occurring as arguments. Frege consistently remarks that the result of substituting names for arguments into argument-places of functional expressions yields a proper name: 'We obtain a name of the value of a function for an argument if we fill the argument-place in the name of the function with the name of the argument' (*Grundgesetze*, 1:6/34).

A typical example is his remark that ' $3^2 > 2$ / $3 > 2$ ' results from ' $\xi^2 > 2$ / $\xi > 2$ ' by '3' being substituted for 'ξ' (*Grundgesetze*, 13:23/54). These substitutional facts are then read as revealing the order of construction implicit in Frege's unstated but implicit formation rules. That Frege consistently speaks of obtaining complex proper names by substituting argument names for place holders in functional expressions is taken as evidence for a formation rule that constructs complex names in just this way. Hans Sluga puts the interpretation this way: 'Frege holds that we can divide the words of our language into those which determine the structure of a complex phrase in which they occur and those which do not. The latter are Frege's names, and the former his functional expressions'.[2]

This syntactic thesis is supplemented by a semantic one that is a result of viewing Frege's language as categorial. Together with principle (3) it follows that the reference of a proper name is computed by applying the reference of its component functional expression to the references of its component names. Edwin Martin describes the semantics as follows: 'The reference of a complex expression is the value of the function which is the reference of the expression's main function-name when it takes as arguments the references of the expressions which fill the main function-name's argument places' (*1974*, 441). Similarly, Dummett writes that a functional expression is '. . . a *part*—or more properly a (partial) *feature*—of some expression which as a whole stands for an object . . . '.[3] On this reading then Frege postulates a semantic operation that does the work of a semantic rule corresponding to the formation rule that takes names and functional expressions to yield names. The operation is referred to by the functional expression, and reference is determined as in (3), in a manner suggestive of a categorial grammar.

2 Sluga *1980*, 140. See also Dummett *1981*, 23, 24, 38, 39, and 176; and Currie *1982*, 23.

3 Dummett *1981*, 250. See also E. Martin *1971*, 49; Dummett *1981*, 159 and *1982*, 319; and Currie *1982*, 23.

A straightforward consequence of having adopted such a general categorial reading of the *Grundgesetze* is that it should apply case by case to the various sorts of sentences, all of which Frege includes with the category of a proper name. The analysis in each case proceeds by postulating a formation rule that breaks down the sentence type into a functional expression and component proper names, by determining the function that the functional expression refers to, and then by determining the reference of the sentence type by applying this function in the manner of (3). In the cases of the functional expressions for the primitive connectives of the horizontal, negation, and the conditional, Frege is fairly clear that he intends the connectives to stand for concepts and about what functions these are.[4] Frege's analysis may be stated rather simply in set-theoretic notation if we acknowledge his stipulation that functions are not to be identified with their extensions but are to be understood rather as undefined primitives. Frege calls a function's extension its course-of-values, and this corresponds essentially to the modern set-theoretic analysis of a function into sets of n-tuples. Let us represent the set of (non-extensional) functions from B into A by A^B, and the set of functions into A with first argument in B and second in C as $A^{(B\times C)}$. Let O be the set of all objects understood to include the truth-value T for the True and F for the False. Frege explains that the functional expressions for the connectives refer to functions from objects to objects that for an arbitrary reference relation R may be defined as follows:

$$R(\text{——}\xi) = (\iota f)(f\varepsilon O^O \,\&\, (\forall x\varepsilon O)(x = T \rightarrow f(x) = T .\&. \; x \neq T \rightarrow f(x) = F)): \tag{6}$$

$$R(\text{—}\!\!\top\!\!\text{—}\xi) = (\iota f)(f\varepsilon O^O \,\&\, (\forall x\varepsilon O)(x = T \rightarrow f(x) = F .\&. \; x \neq T \rightarrow f(x) = T)): \tag{7}$$

$$R\left(\begin{array}{l}\text{—}\!\!\top\!\!\text{—}\xi \\ \;\;\llcorner\text{—}\xi\end{array}\right) = (\iota f)(f\varepsilon O^{(O\times O)} \,\&\, (\forall x\varepsilon O)(x = T \,\&\, y \neq T . \rightarrow f(x,y) = F :\&: x \neq T \vee y = T . \rightarrow f(x,y) = T)). \tag{8}$$

Sentences with a quantifier ranging over objects can be obtained from second-level functional expression, those which refer to a function from concepts to objects, as follows:

$$R(\text{—}\!\smile^{\mathfrak{a}}\!\text{—}\,\phi(\mathfrak{a})) = (\iota f)(f\varepsilon O^{(O^O)} \,\& \; (\forall g\varepsilon O^O)(\forall x\varepsilon O)(g(x) = T) \rightarrow f(g) = T .\&. \; (Ex\varepsilon O)(g(x) \neq T) \rightarrow f(g) = F)). \tag{9}$$

4 As Heck and Lycan point out (*1979*, 488), there is some evidence for understanding the function referred to by the negation functional expression, analyzed here in (7), to be 'horizontal' in the sense of having the function g defined in (6) (or one like it restricted to just truth-values) applies to the relevant arguments. The issue is discussed in detail below. Briefly, I opt for the unhorizontalized (7) because it is equivalent to the horizontalized version and because it conforms to those functions corresponding to the conditional and the quantifier which in the *Grundgesetze* are not horizontalized.
For (6) see *Grundgesetze*, 5:9/38; for (7) see 6:10/39; for (8) see 12:20/51; and for (9) see 8:11/40*ff.* and 26:43/81.

The controversial part of the categorial interpretation now begins. An interpretation for the sentence types generated from these four functional expressions follows directly from the general form of a categorial interpretation (3), together with the functions just defined. Let Δ and Γ be sentences, let a functional expression $\Phi(\xi)$ with placeholder ξ contain no occurrences of $\mathfrak{a}$, and let $\Phi(\mathfrak{a})$ be like $\Phi(\xi)$ except for containing $\mathfrak{a}$ wherever $\Phi(\xi)$ contains ξ. Applying (3) to (6)–(9) then yields the standard account of sentential reference:

$$R(\text{——}\Delta) = R(\text{——}\xi)(R(\Delta)); \tag{10}$$

$$R(\text{—}\!\top\!\text{—}\Delta) = R(\text{—}\!\top\!\text{—}\xi)(R(\Delta)); \tag{11}$$

$$R\left(\text{—}\!\top\!\!\begin{array}{l}\text{—}\Delta\\ \text{—}\Gamma\end{array}\right) = R\left(\text{—}\!\top\!\!\begin{array}{l}\text{—}\xi\\ \text{—}\zeta\end{array}\right)(R(\Gamma),\ R(\Delta)); \tag{12}$$

$$R(\text{—}\smile^{\!\!\!\!\mathfrak{a}}\text{—}\Phi(\mathfrak{a})) = R(\text{—}\smile^{\!\!\!\!\mathfrak{a}}\text{—}\phi(\mathfrak{a}))R(\Phi(\zeta)). \tag{13}$$

Of these four formulas, I think (13) alone is an accurate rendering of what Frege intends (see *Grundgesetze*, 8:12/42 and 21:37/74). But in addition there is some evidence to think (10)–(12) do capture Frege's views. First of all is the general argument that Frege is committed to global categorial approach, a thesis which is in turn justified by what Frege says about functions as explained previously. What I shall try to show is that though Frege does intend the connectives to stand for the functions captured essentially in (6)–(8), he would not subscribe to the categorial analyses given in (10)–(12). I shall also explain how his views on functions admit of exceptions and how they may be read as inconsistent with universal application of a categorial rule like (3). A second sort of evidence for (10)–(12) is textual. Frege does sometimes write in a way suggestive of these formulas, and in due course I shall discuss the relevant texts.

But before discussing the reasons for the categorial reading, we should note its most striking consequence for Frege's semantic theory. On the categorial reading, the semantics is not well-defined; some well-formed expressions are not assigned references. This problem is acknowledged by defenders of the categorial reading. Edwin Martin describes one case as follows: '"&" is apparently a truth function when standing between sentences but is not when contributing to the references of larger quantificational contexts such as "$\forall x(Fx \;\&\; Gx)$"' (*1971*, 441). What Martin is suggesting is that conjunction as it appears in the expression '*Fx* & *Gx*' must be defined for the references of the parts '*Fx*' and '*Gx*', but the references of these are concepts, whereas the reference of conjunction is a function defined for objects, not concepts. Changing the example to the simpler case of negation, the same problem may be illustrated in terms of the categorial principles (1)–(13).

An instance of (13) is:

$$R(\text{—}\smile^{\!\!\!\!\mathfrak{a}}\!\top\!\text{—}\Phi(\mathfrak{a})) = R(\text{—}\smile^{\!\!\!\!\mathfrak{a}}\text{—}\phi(\mathfrak{a}))R(\text{—}\!\top\!\text{—}\Phi(\xi)). \tag{14}$$

Together (11) and (14) entail:

$$R(\,-\!\overset{\mathfrak{a}}{\smile}\!-\,\Phi(\mathfrak{a})) = R(\,-\!\overset{\mathfrak{a}}{\smile}\!-\,\phi(\mathfrak{a}))R(-\!\!\top\!\!-\zeta)(R(\Phi(\xi)))). \tag{15}$$

But the function on the right side of (15) is undefined for its argument because, by (9), the domain of the function $R(\,-\!\overset{\mathfrak{a}}{\smile}\!-\,\phi(\mathfrak{a}))$ is O^{O} but its argument $R(\!-\!\!-\!\zeta)(R(\Phi(\xi)))$ is in O by (7). A similar problem is pointed out by Heck and Lycan *1979* as following on their categorial interpretation of the horizontal. One instance of the law of the amalgamation of horizontals (*Grundgesetze*, 8:14/43) is:

$$R(\,-\!\overset{\mathfrak{a}}{\smile}\!-\,\Phi(\mathfrak{a})) = R(\,-\!\overset{\mathfrak{a}}{\smile}\!-\,(\!-\!\!-\!\Phi(\mathfrak{a})). \tag{16}$$

Together (16) and (13) imply:

$$R(\,-\!\overset{\mathfrak{a}}{\smile}\!-\,\Phi(\mathfrak{a})) = R(\,-\!\overset{\mathfrak{a}}{\smile}\!-\,\phi(\mathfrak{a}))(R(\!-\!\!-\!\Phi(\xi))). \tag{17}$$

Then from (17) and (10) follows:

$$R(\,-\!\overset{\mathfrak{a}}{\smile}\!-\,\Phi(\mathfrak{a})) = R(\,-\!\overset{\mathfrak{a}}{\smile}\!-\,\phi(\mathfrak{a}))(R(\!-\!\!-\!\xi))(R(\Phi(\xi)))). \tag{18}$$

But this identity cannot hold, because by (9) the entity named just to the right of '=' is a function from O^{O} to O while its argument term to its right names an object in O.

Strictly speaking, however, both these derivations are bogus; they do not follow on a strict reading of the categorial principles. In both derivations the rules governing sentential reference are taken as explaining the reference of functional expressions beginning with the same operator. In (15) the rule (11) for $R(-\!\!\top\!\!-\Delta)$ is applied to obtained $R(-\!\!\top\!\!-\Phi(\xi))$, and in (18) rule (10) for $(R\!-\!\!-\!\Delta)$ is applied to get $R(\!-\!\!-\!\Phi(\xi))$. But since neither $-\!\!\top\!\!-\Phi(\xi)$ nor $-\!\!-\Phi(\xi)$ is a sentence, these steps seem to be illegitimate.

There still is, however, a major technical flaw in Frege's semantics under the categorial interpretation, and which can be more accurately put as follows. On this reading no explanation is given of the reference of functional expressions other than the three particular functional expressions for the connectives given (10)–(13). But these three functional expressions will not suffice for determining the reference of an arbitrary functional expression. Moreover, on this reading the reference of an arbitrary functional expression is presupposed, because such an expression is used in the *definiens* of the quantifier in (13)—its reference conditions are defined in terms of an arbitrary functional expression $\Phi(\xi)$. The categorial account is ill-defined in the sense that, though it presupposes an analysis of the reference conditions for an arbitrary functional expression in the analysis of the reference of the quantifier, its analyses of the reference conditions for the connectives will not suffice for determining the reference of functional expressions of arbitrary grammatical complexity.

4. A substitutional interpretation

It is possible, I think, to construct a well-defined semantics for the *Grundgesetze* consistent with what Frege actually says. But before showing how, I should make clear how much we are departing from straight exegesis. It would be anachronistic to find in Frege an explicit set of syntactic formation rules, corresponding semantic rules, and a recursive definition of reference in the modern sense. The categorical interpretation which often presents itself as exegesis goes beyond the test in just this way by attributing to Frege a categorial grammar. Frege nowhere sets out grammar in the sense of generating all well formed expression from a limited lexicon and a restricted set of formation rules. He explains what expressions mean, but semantically not syntactically. The formulation of a grammar sufficient for the language of the *Grundgesetze* is a project entirely supplementary to what we find in the text. There are, however, some loose guidelines for any reconstruction of the implicit grammar. It must not be too narrow in ruling out any of the acceptable expressions actually used in the text, and it should also assign expressions to categories consistently with the examples which Frege gives. The grammar must not be too broad, either. Exactly how restricted it should be is hard to say. Obviously not everything should be admitted, but much more should be admitted than Frege actually uses. We shall encounter later a substantive issue of interpretation that turns on how broadly the grammar should be construed. The categorial interpretation reconstructs Frege's grammar to the extent that it postulates a definite construction of whole from part viewing all sentences as formed from functional expressions.[5]

Semantics, too, is a construction that goes beyond the text. Frege explains what expressions mean in semantic terms. But there is no reason to suppose that he does so in the precisely specified way of a recursive definition. His explanations of the meanings of expressions then should not be assumed to be clauses in a recursive definition. Rather, they should be viewed as semantic truths that serve as the measure of success of any attempt to organize Frege's semantics in a recursive fashion. Only that reconstruction is adequate in which Frege's actual explanations follow as metatheorems. We shall see shortly an example in which what Frege says about the semantics is motivated by quite clear interests that do not have anything to do with recursive definitions. Having made clear what it is we are about to do, let us now turn to the central issue, the interpretation of what he calls 'first-level functional expressions'.

Frege discusses these expressions at length in the *Grundgesetze* (1:5/33*ff.*). He is concerned to make the points that functions are not objects, that they are unsaturated, and that they are distinct from expressions. It is also clear from the *Grundgesetze* and other writings that first-level functional expressions are taken as referring to functions on objects. (See also 'Function and concept', as well as 'Coments on sense and meaning' and the letter from Frege to Husserl of 24 May 1891.) But what he never fully formulates is the quite specific view he seems to hold about the way in which the particular function picked out by a functional expression is determined. His idea seems to be that the expression stands for that function which takes as an argument any object that is denoted by a name that may fill the gap in the functional expression, and pairs with it as value the object referred to by the functional expression when its gap is so filled. Frege would thus be assuming that functions are defined for

5 For categorial interpretations with quite precise grammatical claims see E. Martin *1974*, Potts *1977*, and Heck and Lycan *1979*.

all objects, and that each object has a name. This interpretation for first-level functional expressions is properly called *substitutional*, in the commonly used sense of the term, because the interpretation of the expression in question is determined by substituting names in its gap. We may give the view a more formal statement by first postulating a set *BPN* of basic proper names mapped by the reference relation R onto the set O of objects. Let $\Phi(\xi)$, $\Phi(\xi, \zeta)$ be arbitrary first-level functional expressions of one and two arguments respectively, with ξ and ζ serving as placeholders in the gaps to be filled by names. Then

$$R(\Phi(\xi)) = (\iota f \varepsilon O^{O})(\forall x \varepsilon O)(En\varepsilon BPN)(R(n) = x \,\&\, f(x) = R(\Phi(n))); \tag{19}$$

$$R(\Phi(\xi, \zeta)) = (\iota f \varepsilon O^{(O\times O)})(\forall x, y\varepsilon O)(En, m\varepsilon BPN)(R(n) = x \,\& \\ R(m) = y \,\&\, f(x, y) = R(\Phi(n, m))). \tag{20}$$

Tracing the reference conditions of $-\!\!\overset{\mathfrak{a}}{\smile}\!\!-\ \Phi(\mathfrak{a})$ back through those of $\Phi(\xi)$, we thus see that the truth-value of the quantifier depends on the references of names and their substitution instances in the functional expression. Together (13) and (19) imply:[6]

$$R(-\!\!\overset{\mathfrak{a}}{\smile}\!\!-\ \Phi(\mathfrak{a})) = T \text{ if } (\forall n\varepsilon BPN)(R(\Phi(n)) = T); \tag{21}$$

$$R(-\!\!\overset{\mathfrak{a}}{\smile}\!\!-\ \Phi(\mathfrak{a})) = F \text{ if } (En\varepsilon BPN)(R(\Phi(n)) = F).$$

The textual evidence for this reading of functional expressions is indirect. In his discussion of function Frege remarks that 'We obtain a name of the value of a function for an argument if we fill the argument-place in the name of the function with the

6 See also 25:42/80.

In a passing remark in a passage not directly about the quantifier, Hans Sluga explains Frege's quantifier in what is essentially a substitutional way, though in passages directly about the quantifier he explains it in Frege's own loose way as being 'about' first-level functions. See Sluga *1980*, 87.

Leslie Stevenson *1973* discusses whether Frege uses a substitutional or referential interpretation of the quantifier and argues that he opts for a substitutional reading much like (21) in his early work but for a referential version, much like our (13) in the *Grundgesetze*. My point is that even (13) becomes referential once the proper analysis of its component $R(\Phi(\xi))$ is given.

The same point can be made with reference to Edwin Martin's suggestion in *1982* that even within the *Grundgesetze* one finds passages suggestive of both the referential and substitutional interpretation of the quantifier. It is Martin's thesis that a shifting between the two interpretations leads Frege into error in his quasi-inductive proof (30:46/85*ff*.) that every expression has a reference. In the discussion of this proof, commentators frequently interpret its result as clashing with the well known inconsistency of Frege's axiom system (see, for example, E. Martin *1982*, 151 and Thiel *1968*, 76). But technically speaking it is perfectly possible to have a well-defined reference relation (in the sense that every expression has a reference) and simultaneously have an 'inconsistent' theory formulated within that language. A first-order naive axiomatization of set theory with an unrestricted comprehension principle is an example; its semantics is well-defined, but the particular set of sentences singled out as a theory is unsatisfiable. As will become clear, I am suggesting that something of this sort holds in the *Grundgesetze*.

name of the argument' (*Grundgesetze*, 1:6/34). The principles (19) and (20) seem to be about the weakest that would be sufficient to imply the claim. In another relevant passage he explains not the denotation of functional expressions but the related issue of the conditions under which such a denotation exists. He does not characterize $R(\Phi(\xi))$ but formulates a condition C such that $(Ex)(R(\Phi(\xi)) = x)$ iff C: 'A name of a first-level function of one argument has a *denotation* . . . if the proper name that results from this function-name by its argument-places' being filled by a proper name always has a denotation if the name substituted denotes something' (*Grundgesetze* 29:46/84, cf. 8:11/41).

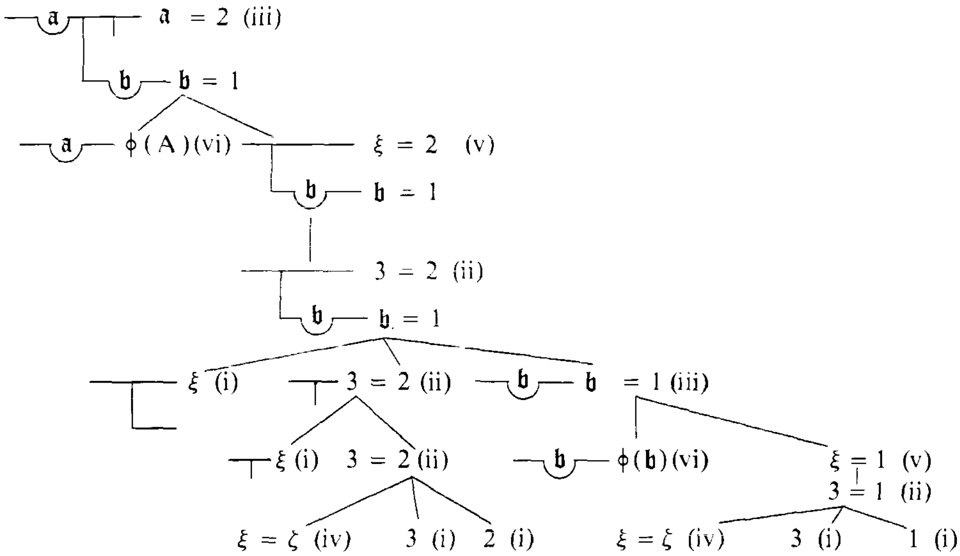
Montgomery Furth (*1966*, xxviii) points out that this passage seems to presuppose that all objects are named. His reasoning is presumably that if a purported function was undefined for an object that lacked a name, the condition C would not spot it and the test as a whole would be unreliable. Again, (19) and (20) find their justification as interpretations in that they insure the dependability of C without assuming much extra.

How this reading of functional expressions would mesh with the interpretations of the other expressions of the formal language, would depend in part on their grammar. Frege of course never defines a grammar in the modern way; the nearest he comes to doing so occurs in his quasi-inductive proof that every expression has a reference (30:46/85*ff.*). His reasoning in the argument roughly is that every expression has a reference because primitive expressions have references and the mechanisms for forming 'longer' expressions from 'shorter' preserve the property of having a reference. The argument thus assumes that there is a grammatical path from any expression through smaller intermediaries to basic expressions. The sort of grammatical analysis Frege assumes can be sketched by considering the case of functional expressions (functors) of one argument place:

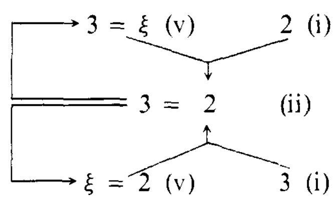
(i) every basic proper name is a proper name:
(ii) if $\Phi(\xi)$ is a first-level functor of one argument and n is a proper name, then $\Phi(n)$ is a proper name;
(iii) If $\mu_\beta\phi(\beta)$ is a second-level functor of one argument and $\Phi(\xi)$ is a first-level functor of one argument, then $\mu_\beta\Phi(\beta)$ is a proper name;
(iv) every basic first-level functor of one argument place, is a first-level functor of one argument place;
(v) if $\Phi(n)$ is a proper name and ξ is a place holder for proper names, then $\Phi(\xi)$ is a first-level functor of one argument place;
(vi) every basic second-level functor of one argument place is a second-level functor of one argument place;
(vii) if $\mu_\beta\Phi(\beta)$ is a proper name and ϕ is a place holder for first-level functors of one argument place, then $\mu_\beta\phi(\beta)$ is a second-level functor of one argument place.

Frege also allows for forming functors of one argument by filling the argument place of a functor of two arguments by an appropriate name, and a full grammar would also need provisions in the previous rules for generating to and from two argument functors in a manner analogous to those of one argument functors, as well as rules for third-level functors and the quantifier over functions. But this sketch is sufficient to illustrate two properties of the construction. First, the grammar does provide for any well-formed expression an graph connecting the expression through intermediaries to basic expressions. In the example below the graph is a tree, and each descent is justified by the application of a rule, annotated to the left of the expression generated by it, as is shown in Figure 1.

Figure 1



Thus, the grammar is sufficient to Frege's purpose in that it 'grounds' each expression in basic ones. If he could prove in addition that if every immediate 'part' refers then so does the outcome of a rule, he would have established his main point. (For a discussion of why he fails to do the latter see Thiel *1968*, 76–77.) An equally striking feature of the grammar, however, is that it does not yield a unique construction for each expression, and constructions may go in circles, as the following case shows:



It is considerations like these that lead Potts (*1973* and *1977*, especially p.9) to attribute to Frege the syntactic interderivability of sentences (names) and functors.

It is not clear what status to accord the grammar (i)–(vii). To view it as recording the grammar implicit in a recursive definition of reference seems too strong. But what is clear is that if it is adopted as the grammar of the *Grundgesetze* together with the substitutional analyses of (19) and (20) of functional expressions, it follows that the semantics provided by Frege is not categorial. Categorial grammars ordinarily call for unique grammatical composition, and more importantly, the reference of an expression is always determined by the reference of expressions used to generate it. But here the reference of a functor depends not on that of the sentential proper names used in an application of rule (v) but rather on that of a possible infinite set of substitution instances of that name. There is no standard sense in which these substitution instances could be construed as the 'parts' of the functor within a categorial grammar. We may speculate that on some suitable abstraction from the notion of a categorial grammar, one that allows for an infinite number of parts and non-unique parsing, all the substitution instances interpreting a functor would be counted as its 'parts', but what the merits of such an abstraction would be is hard to judge.

In the context of first-order logic the quantifier can be interpreted referentially as part of a categorial grammar in several ways that do conform to (3). In Tarski's early interpretation sentences are assigned truth-values as references relative to a particular assignment of values to the variables known as a satisfaction sequence. The result is that the syntax of sentences is homomorphic under an assignment of references to the structure of truth-values organized by truth-functions. More elegant perhaps is Tarski's later interpretation in which sentences, both open and closed, stand for sets of satisfaction sequences, for the set of all sequences relative to which they are true. An algebra of these sets is then straightforwardly homomorphic to syntax. It is even possible to obey the categorial restraints and follow Frege somewhat more closely by assigning truth-values to closed sentences and functions from objects to truth-values to open sentences. An algebra of operations definable for both truth-values and 'concepts' is then definable that is homomorphic to syntax.[7] But each of these three referential accounts of first-order logic departs from Frege in major ways. Since both of Tarski's accounts assign open and closed sentences the same sort of interpretation, neither draws Frege's semantic distinction between the two types of expression. Open sentences, the first-order equivalent to Frege's functional expressions, and closed do not stand for fundamentally different sorts of things. Tarski's later accounts also share a major non-Fregean feature with the third account. In both the semantic operations corresponding to the connectives are defined for both objects and 'concepts' rather than for just objects as Frege would have it. There is, I think, a lesson in the failure of first-order semantics to accurately match Frege's. It is difficult, perhaps impossible, both to assign to the various categories of expression the values Frege wishes and to simultaneously conform to the restraints of a categorial grammar. Frege's own semantics may be understood as coping with this dilemma by opting for what is essentially a non-categorial substitutional interpretation of the quantifier.

5. The formal semantics

The motivational remarks of the last section may now be drawn together by a formal statement of a fully defined semantic theory. The language of *Grundgesetze* will

7 For a statement of the two versions of Tarski's semantics and the Fregean alternative, as well as a discussion of their properties as categorial grammars, see my *1977*.

be simplified in several non-essential ways. Omitted are expressions for generality (§8), the operator for the definite article (§11), and functional expressions of more than one argument place, and the number of variables is limited to those Frege actually uses. As before no assumption is made about the set O of objects other than it contains T for the True and F for the False, and functions are to be understood as non-extensional primitives distinct from their courses-of-values which are identified with sets of n-tuples. For simplicity (AB) will do double duty as a type in the syntax and as the set A^B of (non-extensional) functions in the semantics, and likewise (ABC) is used for both a type and the set $A^{(B\times C)}$. The format of a categorial grammar is observed to the degree permitted by the interpretation, and predicates and operators are accordingly treated as genuinely referring expressions standing for functions that apply to the references of a sentence's immediate parts. The grammar is also defined so as to avoid syntactic ambiguity and loops. To each expression it is possible to associate a unique finite grammatical tree tracing its constructions through shorter expressions to basic expressions. In doing so I suppress one half of the syntactic interderivability mentioned in the grammar (i) – (vii) holding between sentential proper names and functors. It should be stressed that in doing so, I am not so much violating Frege's explicit grammar, as grammar is now understood, as I am regularizing in an anachronistic manner for the purposes of recursive semantics the syntactic remarks Frege makes for other purposes.

Since (i) – (vii) associate with each expression a connected graph grounded in basic expressions, the semantics below could be adapted to a grammar that preserves the interderivability, perhaps in the manner of Potts. The result would be well defined with the same set of assertable sentences, but not a categorial syntax.

Though I conform to the categorial tradition in syntax, doing so in the semantics would contradict Frege's explicit remarks on functional expressions. In violation of the categorial requirement (3), the reference of these expressions is not defined in terms of those of their immediate parts, but rather in the way spelt out in the substitution interpretation. In conformity with this substitutional interpretation, the assumption is made explicitly that all objects and all first-level functions are named. No assumption, however, is made about the cardinality of these sets. That is some intended interpretation these sets may be non-denumerable and beyond the naming capacity of a denumerable grammar is a limitation of any substitutional interpretation, and certainly not unique to Frege. (There is now a substantial literature on the limitations of the substitutional interpretation. See, for example Dunn and Belnap *1968*, and Kripke *1976*.) No attempt is made to reconcile the semantics with the proof theory of the *Grundgesetze*. The axiom system is inconsistent and would require major correction well beyond anything Frege himself contemplated, and well beyond the scope of this essentially semantic study.

For our purposes it will suffice to define a *complex symbol* in intuitive syntactic terms as any string of symbols, and we shall say that a complex symbol is a *part of* a longer string of symbols if it occurs within it. We let $A(B)$ stand for a complex symbol containing B as a part, and let $[A(B)]C/B$ refer to that symbol like $A(B)$ except for containing C wherever $A(B)$ contains B. The set of *types* is defined as the least set such that O is a type and if A and B are types then so are (OA) and (OAB). Parentheses surrounding type notation are dropped if convenient.

The grammar consists of a series of sets of *expressions* defined as follows. They are divided into those which are lexical and defined by stipulation and those which are complex and defined in clauses that describe their formation.

(1) Lexical expressions.

(A) The set of *basic proper names of type O* is $\{c_1, \ldots, c_n, \ldots\}$, and the set of *basic proper names of type OO* is $\{f_1, \ldots, f_n, \ldots\}$;

(B) the set of *predicates of type OO* is {—, ┬}, and the set of *predicates of type O(OO)* is { =, ┬└ };

(C) the set of *operators of type O(OO)* is { —◡₁— ,}, and the set of *operators of type O(O(OO))* is { —◡₂— };

(D) the set of *variables of type O* is { $\mathfrak{a}$, $\mathfrak{e}$, ε }, and the set of *variables of type OO* is { $\mathfrak{f}$ };

(E) the set of *place holders of type O* is $\{\xi, \zeta, \chi\}$, and the set of *place holders of type OO* is $\{\phi, \psi\}$.

(2) Complex expressions.

(A) The set of *proper names* (of type *O*) is the least set such that

(i) every basic proper name of type *O* is a proper name;

(ii) if *A* is a predicate or basic proper name of type *OO* and *B* is a proper name of type *O* then *AB* is a proper name, and if *A* is a predicate of type *O*(*OO*) and both *B* and *C* are proper names of type *O*, then *ABC* is a proper name;

(iii) if (a) $\Phi(n)$ is a proper name,
(b) *n* is the basic proper name of type *O*,
(c) *v* is a variable of type *O*, and
(d) $\Phi(v) = [\Phi(n)]\, v/n$,
then —$\overset{v}{◡}$₁— $\Phi(v)$ and $\grave{v}\Phi(v)$ are proper names (here *v* is said to be *bound* in each);

(iv) if (a) $\mu_\beta\Phi(\beta)$ is a proper name containing $\Phi(\beta)$ as a part,
(b) β is bound in $\mu_\beta\Phi(\beta)$ but not bound in $\Phi(\beta)$,
(c) for some basic proper name *n* of type *O*, $\Phi(\beta) = [\Phi(n)]\beta/n$,
(d) *v* is a variable of type *O*(*OO*), and
(e) $\mu_\beta v(\beta) = [\mu_\beta\Phi(\beta)]\, v(\beta)/\Phi(\beta)$,
then —$\overset{v}{◡}$₂— $\mu_\beta v(\beta)$ is a proper name.

(B) The set of *functors of level one* (of type *OO*) is the least set such that
if (i) $\Phi(n)$ is a proper name,
(ii) ξ is a place holder of type *O*, and
(iii) for some basic proper name *n* of type *O*, $\Phi(\xi) = [\Phi(n)]\xi/n$,
then $\Phi(\xi)$ is a functor of level one.

(C) The set of *functors of level two* (of type $O(OO)$) is the least set such that

if (i) $\mu_\beta \Phi(\beta)$ is as in (a)–(c) of (A–iv) above,
(ii) ϕ is a place holder of type OO, and
(iii) $\mu_\beta \phi(\beta) = [\mu_\beta \Phi(\beta)]\phi(\beta)/\Phi(\beta)$,
then $\mu_\beta \phi(\beta)$ is a functor of level two.

The partitioning of expressions determined by this grammar is used to define the key concept of the semantics, the notion of a reference relation. It will suffice to assign values to expressions according to their syntactic variety. Those for lexical expressions are defined directly and those for complex expressions by recursion. Let O be a non-empty set containing T and F. A *reference relation* is any function R on expressions that assign values as follows.

(1) Lexical expressions.

(A) For any basic proper name n of type O, $R(n)\varepsilon O$;

(B) R assigns functions in OO to basic proper names of type OO, and R assigns values to predicates as follows:

(i) $R(\text{——}) = (\iota f\varepsilon OO\,(\forall x\varepsilon O)(x = T \rightarrow f(x) = T .\&.\ x \neq T \rightarrow f(x) = F)$,
(ii) $R(\text{—}\!\!\top\!\!\text{—}) = (\iota f\varepsilon OO)(\forall x\varepsilon O)(x = T \rightarrow f(x) = F .\&.\ x \neq T \rightarrow f(x) = T)$,
(iii) $R(=) = (\iota f\varepsilon OOO)(\forall x, y\varepsilon O)(x = y \rightarrow f(x, y) = T .\&.\ x \neq y \rightarrow f(x, y) = F)$,
(iv) $R(\text{—}\!\!\sqsubset\!\!\text{—}) = (\iota f\varepsilon OOO)(\forall x, y\varepsilon O)(x = T \,\&\, y \neq T .\rightarrow f(x, y) = F :\&: x \neq T \vee y = T .\rightarrow f(x, y) = T)$;

(C) R assigns functions to operators as follows:

(i) $R(\text{—}\!\smile_1\!\text{—}) = (\iota f\varepsilon O(OO))(\forall g\varepsilon OO)([\forall x\varepsilon O, g(x) = T] \rightarrow f(g) = T .\&.\ [Ex\varepsilon O, g(x) \neq T] \rightarrow f(g) = F)$,
(ii) $R(\,'\,) = (\iota f\varepsilon O(OO))(\forall g\varepsilon OO)(f(g) = \{\langle x, y\rangle | f(x) = y\}$,
(iii) $R(\text{—}\!\smile_2\!\text{—}) = (\iota f\varepsilon O(OO)))(\forall g\varepsilon O(OO))([\forall h\varepsilon OO, g(h) = T] \rightarrow f(g) = T .\&.\ [Eh\varepsilon OO, g(h) \neq T] \rightarrow f(g) = F)$.

(2) Complex expressions (by category corresponding to clauses in the syntax).

(A) Proper names.

(i) For any basic proper name n, $R(n)$ is already defined;
(ii) $R(AB) = R(A)(R(B))$,
$R(ABC) = R(A)(R(B), R(C))$;
(iii) $R\ \text{—}\!\overset{v}{\smile}_1\!\text{—}\ \Phi(v)) = R(\text{—}\!\smile_1\!\text{—})(R(\Phi(\xi)))$, and
$R(\dot{v}\Phi(v)) = R(\,'\,)(R(\Phi(\xi)))$;
(iv) $R(\text{—}\!\overset{v}{\smile}_2\!\text{—}\ \mu_\beta v(\beta)) = R(\text{—}\!\smile_2\!\text{—})(R(\mu_\beta \phi(\beta))$.

(B) Functors of Level One.
$R(\Phi(\xi)) = (\iota f\varepsilon OO)(\forall x\varepsilon O)(En)$ (n is a basic proper name type O & $f(x) = R(\Phi(n))$.

(C) Functors of Level Two.
$R(\mu_\beta \phi(\beta)) = (\iota f\varepsilon OO)(\forall g\varepsilon OO)(Eh)$ (h is a basic proper name of type OO & $g = R(h(\xi))$ & $f(g) = R(\mu_\beta h(\beta))$
(recall by the syntax that for some basic proper name n of type O, $h(\xi) = [h(n)]\xi/n$ and $h(\beta) = [h(n)]\beta/n$).

A reference relation R is said to be *acceptable* iff it meets the following expressibility conditions:

(i) for every $x\varepsilon O$, there is a basic proper name n of type O such that $R(n) = x$,

and

(ii) for every $f\varepsilon OO$, there is a basic proper name h of type OO such that $R(h) = f$.

A proper name Δ may be said to be a *logical truth* or to be (semantically) *assertable* iff its reference is T under every acceptable reference relation, and we may abbreviate the statement that Δ is assertable by $||\!\!\vdash\!\!\!-\Delta$, using $||$ as a semantic correlate of Frege's proof theoretic assertion sign $\vdash\!\!\!-$.[8] The use of the horizontal here and its properties more generally are our next topic.

6. The horizontal

When studied in detail the horizontal is probably the most obscure feature of Frege's formal language. Some of these details are regularly used as backing for the categorial interpretation. I have postponed their consideration until now because of their complexity.

6.1. *Horizontal analysis.* There is a common view about the horizontal that is used to support the categorial interpretation. It requires that a version of the standard interpretation must hold for all the connectives and operators that in Frege's symbolism begin and end with a horizontal line segment. These include negation, the conditional, and the quantifiers. The view holds that any expression beginning with these symbols is to be analyzed in terms of immediate parts that themselves begin with horizontals, and for this reason we may call it the thesis of *horizontal analysis.* It is in part syntactic, in part semantic. Grammatically, any expression including both proper names and functional expressions that starts with one of these connectives or operators is viewed as constructed from an expression that starts with a horizontal, if the symbol is a one-place connective or operator, or from two horizontalized immediate parts in the case of the conditional. Each of these horizontalized parts is in turn constructed from what is to the right of the horizontal by the formation rule that adds the horizontal to it. Semantically, the analysis parallels the syntax in the categorial

8 In the language of Section 6, $||\!\!\vdash\!\!\!-$ is 'non-composite' and not intended to contain a horizontal as a part.

manner; and for this reason the thesis, if correct, would go a long way to justifying the categorial interpretation. Reference of a whole expression beginning with one of horizontalized connectives or operators is computed by applying the function corresponding to the connective or operator to its horizontalized part or parts. The references of these is in turn computed by applying the function corresponding to the horizontal to the reference of the expressions that occur to their right. What I shall argue is that this picture is much more precise than anything Frege actually says, and that the text does not require this interpretation of functional expressions beginning with horizontalized connectives or operators.

The thesis of horizontal analysis when applied together with the general assumptions of a categorial grammar entails that the horizontal operations occurring in the analysis be reflected in the grammar itself. This condition is satisfied by another common view of the horizontal, which we may refer to as the composite reading of the horizontalized connectives and operators. On this reading Frege is understood to have intended these symbols which all start and end with horizontal line segments to be literally a series of distinct expressions with a horizontal at beginning and end, and containing a more primitive sign for the operation or connective in question sandwiched between the outside horizontals. The logical relation of this reading to the thesis of horizontal analysis is rather subtle. Horizontal analysis does not alone entail that the language is categorial because it does not bear on all the expressions of the language; the thesis could be true yet one of its other expressions could still fail to conform to the categorial framework. But the thesis does lend support to the general categorial reading even if this justification is less than entailment. The categorial framework in turn requires that the semantic operations applied in the analysis of an expression be represented in the syntax by a part referring to that operation. Exactly how this representation should be carried out is not dictated either by the categorial framework or the thesis of horizontal analysis; there are many possible ways in which the syntax could meet this condition. The composite reading insures that one of these obtains. Thus it works together with the horizontal analysis to give, if correct, strong but not conclusive support for viewing Frege's language as a whole as categorial.[9]

There is yet a third view about the horizontal that though not entailed by horizontal analysis or the composite reading goes in conjunction with them to support a categorial reading. This is the thesis that Frege intended the semantic functions referred

9 Not only does the thesis of horizontal analysis not entail the composite reading, the converse fails also. It would be possible to read all Frege's horizontalized connectives and operators as composites of horizontals and more primitive connectives and operators, yet analyze these primitives semantically without any appeal to the horizontal's semantic operation. One way to do so would be to alter the primitive symbols of the syntax of Section 5 so that the horizontal line segments were deleted from both sides of the horizontalized connectives and operators. The syntax and semantics would then be defined for just unhorizontalized primitives and the horizontal. Composite expressions could then be generated by nesting the primitives inside horizontals, and the expression of the *Grundgesetze* could be read as composites of this sort. The composite reading would hold then in the context of a non-standard semantics. Though formally adequate, such an interpretation would leave unexplained why Frege never uses the primitives alone and systematically prefers their horizontalized uses. It is for this reason that in the non-standard interpretation I have given in Section 5, I have also rejected the composite reading.

to by negation and the conditional to be defined for just truth-values rather than objects in general. If defined for just truth-values, they could not in general be applied directly to proper names for then the resulting expression would lack a reference. Hence the role of the intervening horizontal. It transforms an object into a truth-value so that later applications of the connectives will be well-defined.

Before considering the evidence for and against these views, it should be pointed out that though if true they support the categorial reading, the categorial reading does not require them. It is perfectly possible to maintain a categorial reading of *Grundgesetze* yet view Frege's horizontalized connectives and operators as syntactic and semantic units, analyzed without reference to deeper horizontals, referring to semantic operations defined for all objects. Michael Dummett (*1981*, 315) consistently argues for just such a reading.

The main reason for doubting the thesis of horizontal analysis and its attendant views is that it leaves the reference relation undefined for a large variety of expressions. We have already seen in (16)–(18) how the view that the reference of $-\!\!-\,\Phi(\xi)$ determines that of $-\!\smile^{\mathfrak{a}}\!-\,\Phi(\mathfrak{a})$ breaks down because, on the view under consideration, the former stands for a truth-value whereas the second-level quantifier operation is defined not for truth-values but concepts. A similar problem arises for negation. Suppose, as Kneale and Kneale (*1962*, 504), and Heck and Lycan (*1979*, 488) suggest, that $R(\neg\,\xi)$ is defined by applying the negation function f, defined over truth-values only, to $R(-\!\!-\xi)$. But $-\!\!-\xi$ refers to a first-level function, and therefore f is undefined for its argument. A reconstruction of Frege's semantics that is consistent with the text but avoids these problems would be preferable.

A second rather global reason for questioning the interpretation is that the thesis that Frege intended the connectives to refer to functions defined only over truth-values is almost totally speculative. Indeed, in his introductory discussion of functions (2:7/35) he remarks that 'the domain of what is admitted as argument must also be extended to objects in general'. Though he is here discussing functions in general rather than the references of the connectives in particular, he does not in discussing the latter explicitly limit their domain to truth-values. Indeed it is possible to assert in the object language of the *Grundgesetze* the logical truth that all first-level functions are defined for all objects:

$$\vdash\!\smile^{\mathfrak{f}}\!-\;(\acute{\varepsilon}(\neg\!\smile^{\mathfrak{a}}\!\neg\;\mathfrak{f}(\varepsilon) = \mathfrak{a}) = \acute{\varepsilon}(\neg\!\smile^{\mathfrak{a}}\!\neg\,(\mathfrak{a} = \varepsilon))). \qquad (23)$$

Moreover, the restriction of connectives to just truth-values is unnecessary. As my main argument is designed to show, it is possible to construct a well defined semantics consistent with the text that does not make this assumption.[10] But despite these general reasons against the thesis of horizontal analysis there are arguments for it which I would like now to address. Three of these are matters of exegesis, and the fourth is a very interesting thesis about the conceptual role of the horizontal. I shall consider each in turn.

10 Thus in Section 5 only the single basic type O is used rather than two basic types O and $\{T, F\}$ common to some modern 'Fregean' languages. In this I follow Potts *1973* and *1977*.

6.2 *Frege's analysis of* $-\!\!\top\!\!-\xi$. As Heck and Lycan have observed (*1979*, 488) Frege sometimes does explain the functional expressions beginning with a connective and followed by place-holders by reference to the same place-holders beginning with horizontals. Both negation and the conditional are explained this way in 'Function and concept' (28/39*ff.*), and negation is explained this way in the *Grundgesetze* (6:10/39). He there explains the reference of $-\!\!\top\!\!-\xi$ in terms of $-\!\!-\xi$. If these explanations are construed as clauses in an attempt at a recursive definition of reference in the style of a categorial grammar, as the reading in question dictates, then $-\!\!-\xi$ must be viewed as a syntactic part of $-\!\!\top\!\!-\xi$, and the thesis of horizontal analysis is confirmed. But I have suggested that these explanations should be viewed rather as semantic truths to which Frege does indeed subscribe but not necessarily as clauses in a modern semantic definition. In trying to understand what the role of these explanations is, it is interesting to note that Frege does not always follow this pattern of explaining a horizontalized connective or operator in terms of a horizontal. In the *Grundgesetze* he explains the conditional $\begin{array}{l}-\!\!\top\!\!-\xi\\ \;\;\llcorner\!-\zeta\end{array}$ in terms of ζ and ξ, not $-\!\!-\zeta$ and $-\!\!-\xi$, and the quantifier $-\!\!\smile^{\mathfrak{a}}\!\!-\;\Phi(\mathfrak{a})$ is explained in terms of $\Phi(\xi)$ rather than the horizontalized $-\!\!-\Phi(\xi)$ (12:20/51, 8:12/42, 21:37/73). This switch of pattern is inexplicable in the account of the Kneales and that of Heck and Lycan, but if what Frege really has in mind is accurately implied by the semantics of Section 5, the variation is easily explained because it follows from that semantics that either style of explanation is correct. Indeed, they are equivalent.

Below in (24) two explanations of $R(-\!\!\top\!\!-\xi)$ are given. The first without a horizontal in the *analysans* describes $R(-\!\!\top\!\!-\xi)$ in the same terms as those used in the semantics of Section V; the second with the horizontal also accurately describes $R(-\!\!\top\!\!-\xi)$ but in terms whose accuracy is derivative from the more basic specifications of $R(-\!\!-\xi)$ and $R(-\!\!\top\!\!-\xi)$ in that definition. Likewise, (25) gives two descriptions of $R\left(\begin{array}{l}-\!\!\top\!\!-\xi\\ \;\;\llcorner\!-\zeta\end{array}\right)$.

$$(\iota f\varepsilon OO)(\forall x\varepsilon O)(x = T \rightarrow f(x) = F .\&.\ x \neq T \rightarrow f(x) = T); \tag{24a}$$

$$(\iota f\varepsilon OO)(\forall x\varepsilon O)(R(-\!\!\top\!\!-\xi)(x) = T \rightarrow f(x) = F .\&. \\ R(-\!\!\top\!\!-\xi)(x) \neq T \rightarrow f(x) = T); \tag{24b}$$

$$(\iota f\varepsilon OOO)(\forall x, y\varepsilon O)(x = T \,\&\, y \neq T .\!\rightarrow f(x, y) = F :\&: x \neq T \vee \\ y = T .\!\rightarrow f(x, y) = T); \tag{25a}$$

$$(\iota f\varepsilon OOO)(\forall x, y\varepsilon O)(R(-\!\!-\xi)(x) = T \,\&\, R(-\!\!-\xi)(y) \neq T .\!\rightarrow \\ f(x,y) = F :\&: R(-\!\!-\xi)(x) \neq T \vee R(-\!\!-\xi)(y) = T .\!\rightarrow f(x,y) = T). \tag{25b}$$

That Frege should shift between the two, supports the view that his explanations were not intended to be clauses in a recursive semantics, and if a recursive semantics is to be provided for what he does say, then these passages are rather more consistent with the non-categorial semantics of Section V than with the categorial account.[11]

6.3 *The argument ——x for the function —┬—x*. The second textual argument for horizontal analysis rests on Frege's metatheoretical law known as the *amalgamation of horizontals*. He gives various formulations but its essence may be summarized as follows: the addition or deletion of a horizontal operator immediately before or after a horizontalized operator governing either a sentence or a functional expression leaves the reference of the whole unaltered.[12]

For our purpose the law should be regarded as part of Frege's explicit semantics that must be captured by any attempt to fill in gaps so as to yield a fully adequate recursive account. The law does indeed follow from the semantics of Section 5. The reason it is suggestive of horizontal analysis rests not in the law itself but in his glosses on it. In 'Function and concept' (22/35), he says about —┬—x (using x for the place-holder ξ of the *Grundgesetze*):

> I conceive of this function with the argument ——x:
>
> $$(\text{—┬—}x) = (\text{—┬—}(\text{——}x)).$$

These remarks are suggestive of the horizontal analysis of negation in that it suggests that the function —┬—x is determined by the function ——x. But literally speaking ——x is not an argument for —┬—x, because the former is a function while the latter is defined for just objects. There are, I think, two possible readings of the remark that ——x is an argument of —┬—x that are consistent with Frege's choice of reference for these expressions. First, we can interpret him to be talking in the material mode, not about expressions, but about what expressions refer to. Then, rather than saying ——x is an argument for —┬—x, what he should have said is the non-controversial proposition that the *values* of ——x are arguments of —┬—x, and when speaking carefully this is what he does say. (See, for example, his remarks on ——ξ at *Grundgesetze*, 48:61/105.) Alternatively we may construe the remark as partly about syntax and partly about references. First, he makes the syntactic point that the expression '——x' is an argument for the expression '—┬—(——x)' in the sense that the latter may be obtained by filling the gap in '—┬—(   )' by '——x'. He then makes the semantic point that '—┬—x' and '—┬—(——x)' are co-referential. In the notation of Section 5 these points may be put:

$$\|\text{—┬—}(\text{——}\xi) = [\text{—┬—}\xi]\text{——}\xi/\xi; \tag{26}$$

$$R(\text{—┬—}\xi) = R(\text{—┬—}(\text{——}\xi)). \tag{27}$$

11 A small inelegance in the horizontalized formulation that may have inclined Frege in favor of taking the descriptions in Section 5 as definitional is that the horizontalized version uses an unnecessary circumlocution. In (24b) and (25b), but not in (24a) and (25a), the phrase '$\neq T$' may be simplified to '$= F$'. I have explained above why I think the functions range over all objects and not just T and F.

12 For the law applied to functional expressions, see 'Function and concept', 22/35, and *Grundgesetze*, 48:61/105*ff*; for its application to sentences, see *Grundgesetze*, 6:10/39, 8:14/43, 12:20/51.

This second reading is perhaps the better because unlike the first it makes sense of both parts of the quoted passage, and does not attribute to Frege a slip of the tongue. But under this interpretation the passage hardly supports the categorial interpretation as is seen from the fact that both (26) and (27) follow from the semantics of Section 5.

6.4 *The composite symbolism of the connectives.* Frege offers a second gloss on the amalgamation of the horizontals that is the strongest of the textual evidence advanced to support the thesis of horizontal analysis. The text is read as asserting not the thesis itself but the corroborating syntactic claim that the horizontalized connectives are composite. In 'Function and concept' Frege refers to the law which we would express as (28):

$$\Vdash((\text{—┬—}\xi) = (\text{—┬—}(\text{——})\xi))). \tag{28}$$

He then remarks that it provides an occasion in which the right hand horizontal line segment of '—┬—' is 'fused' with the horizontal. He also singles out the vertical line segment from its flanking horizontal lines and gives it the special name of the 'stroke of negation'. Explaining himself more fully he then says: 'I thus regard the bits of the stroke in "—┬—x" to the right and to the left of the stroke of negation as horizontals, in made to work; but, taken as a place-holder, as it is intended, the account breaks 105*ff.*). Commentators interpret these passages as direct statements of the composite nature for the negation sign '—┬—'. Syntactically, 'ξ' is joined with '——' to produce '——ξ'. Then '——ξ' is joined with '|' to yield '┌——ξ'. Finally, '┌——ξ' joins with '——' to make '—┬—ξ'. This syntactic analysis is then interpreted in the framework of a categorial grammar to yield a progressive semantic breakdown of $R(\text{—┬—}\xi)$ that runs essentially as follows:[13]

$$\begin{aligned} R(\text{—┬—}\xi) &= R(\text{——})(R\text{┌——}\xi); \\ R(\text{┌——}\xi) &= R(|)(R(\text{——}\xi)); \\ R(\text{——}\xi) &= R(\text{——})(R(\xi)). \end{aligned} \tag{29}$$

We have already seen that if ξ is understood as a sentence, the account may be made to work; but, taken as a place-holder, as it is intended, the account breaks down. The argument terms on the right side of (29) refer to entities for which the functions named to their left are undefined. The arguments are concepts but the

13 For examples of this reading, see Kneale and Kneale *1962*, 504, and Heck and Lycan *1979*, 488. These commentators and others (see Currie *1982*) take the connectives themselves as standing for functions rather than combinations of connectives and place-holders. In the explanation of functional expressions like those on the right of (29) Frege himself would use other functional expressions, i.e. combinations of connectives with place-holders, and I follow this practice in (10)–(12). Thus, instead of $R(\text{——})$ and $R(|)$ on the right of (29) we should have $R(\text{——}\xi)$ and $R(|\xi)$. In the formal semantics of *sentences* in Section 5, I do give references directly to the connectives (I–B–i, ii and iv) but these are intended as constructs consistent with but supplemental to Frege's text.

functions are defined for just objects. Fortunately it is not necessary to read the quoted passages in this way.

To read Frege's remarks about '—┬—' having bits consisting of '——' and '|' as a commitment to corresponding syntactic and semantic operations is to see him working in a modern framework which he nowhere clearly formulates. Moreover, if the details of categorial reading are to be believed, he does a rather bad job of employing the framework in that the resulting relation is poorly defined. A more likely reading of the remarks, it seems to me, is that they are intended to summarize in alternative words the metatheorem they gloss. Saying that '—┬—' consists of two horizontals flanking a negation stroke may mean nothing more than that adding or deleting extra horizontals to the right or left of '—┬—' leaves the reference of the resulting expression unchanged. That speaking of bits and distinct parts should lead later readers schooled in categorial semantics to understand these ideas in a later technical sense is not Frege's fault. The truth is that in his mouth these words need not have these technical meanings.

6.5 *The conceptual role of the horizontal.* The final argument to consider for the thesis of horizontal analysis is theoretical rather than textual. It contends that on the alternative analysis the horizontal is conceptually trivial and has no significant contribution to make to metatheory. Suppose, the argument goes, that horizontalized connectives are not composite and that their semantics are not defined in terms of horizontal operations. Then the horizontal line that forms a part of the connective symbol might as well not be there, and even the horizontal used on its own is replaceable in every occurrence by non-horizontalized equivalent expressions. The whole idea of the horizontal is therefore dispensible. But an interpretation that plausible attributes a more significant role to the idea that Frege apparently thinks to be important would be preferable. The analysis then goes on to spell out a non-trivial role for the horizontal on the composite categorial interpretation. (This argument is suggested by the remarks of Heck and Lycan *1979*, 487.)

The positive account of the horizontal advanced by the categorial reading is indeed interesting and will concern us in the next section. First I would like to show that on the alternative reading the horizontal is not trivial as alleged but a relatively interesting theoretical device. It is true that all occurrences of the horizontal can be replaced by other non-horizontalized expressions. Frege himself observes that ——ξ might be defined as $\xi = (\xi = \xi)$ (*Grundgesetze*, 16/46 and 49/89); and Dummett, who rejects the horizontal analysis and the composite view of connectives, remarks (*1981*, 315) that ——ξ may be defined by ┬┬ξ, which on the non-composite reading contains no horizontals. But this eliminability of free-standing horizontals does not show the expression is conceptually trivial any more than the definability of disjunction in terms of negation and conjunction in standard logic shows that that idea is trivial. All that is established is that Frege's primitive expressions are somewhat redundant from the viewpoint of expressive power. In axiomatics and other proof theoretic contexts the expressive minimum may lead to brevity and elegance, and questions of expressive power and functional completeness are sometimes significant semantically. But expressive redundancy certainly does not show that a potential eliminable expression is conceptually vacuous. Disjunction makes a real distinction in standard logic, and the horizontal picks out a distinct concept in Frege's semantics despite its eliminability.

It is also possible to attribute a conceptual role to the horizontal line segments of the connectives and operators even when these are taken as non-composite. In its broadest sense a non-trivial relation among sentences may be defined as any relation that holds among some but not all sentences. If a non-trivial relation of this sort happens to distinguish a subset of the logically valid arguments, it is not only meaningful but of some logical interest. The law of the amalgamation of the horizontals singles out just such a variety of valid inferences. The role of the horizontal line segments may be viewed similarly. When they occur before and after a connective or operator, they signal that the law of amalgamation applies to any expression it heads. They single out a non-trivial logical relation that an expression headed by a horizontal bears to other expressions just like it except for the addition or deletion of horizontals immediately before or after the horizontalized symbol. It is enough to think of Frege's horizontalized orthography for the connectives and operators as having for its theoretical role merely the display of this logical property.

That the relation in question is genuinely non-trivial is shown by the fact that the amalgamation of horizontals is invalid for some expressions. It fails both to the left (outside) and to the right (inside) of the abstraction operator, and within identities, as the following assertable sentences show. Let 2, 3 and < have their intended mathematical interpretations:

$$|\!\vdash\!\!\top\!\!-\acute{\varepsilon}(\varepsilon) = \acute{\varepsilon}(\text{——}\varepsilon); \tag{30}$$

$$|\!\vdash\!\!\top\!\!-\acute{\varepsilon}(\varepsilon < 2) = \text{——}\acute{\varepsilon}(\varepsilon < 2); \tag{31}$$

$$|\!\vdash\!\!\top\!\!-(3 = \text{——}3); \tag{32}$$

$$|\!\vdash\!\!\top\!\!-((2 = 3) = 3) = ((2 = 3) = \text{——}3). \tag{33}$$

The law does apply, however, to the outside of identity, and Frege could have indicated so in the symbolism by somehow adding a horizontal line segment, distinct from a free-standing horizontal, outside identity assertions, e.g. by writing $\text{——}\!\!\dashv\!(\xi = \zeta)$ instead of $\xi = \zeta$ or $\text{——}(\xi = \zeta)$.

7. Bochvar's many-valued logic

The case I have to make for the semantics of Section 5 is now complete. I have argued the interpretation is consistent with the text of the *Grundgesetze*, more consistent on the whole than the alternative categorial interpretation, and that it avoids serious problems of undefinedness. But there remains a portion of the categorial view to be expounded, and though interesting for its contributions to logical theory, it is not really attributable to Frege in a historically accurate sense. If his language is read as having composite connectives defined in terms of horizontals, then if many of the details of the language are suppressed and a propositional logic employing just negation, the conditional, and the horizontal is abstracted, then the resulting language can be explained to have a classical logic in an especially interesting manner. Truth-functional compounds that are in the form of classical validities yet have atomic components which do not stand for truth-values are rescued from meaninglessness by the application of horizontals. Horizontals seal off meaninglessness, converting it to a truth-value, so that a truth-functional compound is always supplied with arguments for which it is well defined. The manner in which this occurs is strikingly similar to the technique worked out in the twentieth century by D.A. Bochvar *1937*. Bochvar's representation of classical logic differs from that abstracted from Frege in that the representation theory is fully explicit and detailed. It is also formulated in a three-valued logic whereas Frege's functions are defined over a much wider set of objects. But in a suitable generalized statement of Bochvar's theory in which arbitrarily many values are allowed, it may be shown that Frege captures classical logic just in the way Bochvar explains. The three-valued language of Bochvar's original exposition and Frege's many-valued account prove to be special cases of a general method. It is this general theory and its application to Frege that shall occupy us here. But it should be clear at the start that what we are engaged in is the reconstruction of theory left inexplicit in Frege's work, if indeed he is motivated by the same ideas at all. The best that can be said is that if the composite readings of the connectives and its attendant categorial semantics were the right way to understand Frege, and I have argued that it is not, then a suitable abstraction from Frege's language would possess a classical logic in the manner of Bochvar.

The use of the horizontal to represent classical logic is relevant to questions of interpretation because it could be seen to illuminate the role of the horizontal on the categorial interpretation. After arguing that on the non-composite reading of the connectives they can find no significant function for the horizontal, Heck and Lycan *1979* try to find one on their categorial account. The only useful role they can find is that it groups together at an identifiable place in the syntactic construction of any sentence a rather arbitrary feature of the semantics. Proper names that occupy the syntactic role of a sentence in that they occur as arguments for connectives like negation and the conditional can in principle fail to refer to truth-values as their syntactic position suggests but rather stand for other sorts of objects. These pseudo-sentences are transformed into genuine names for truth-values by the addition of a horizontal, and since on the composite account a horizontal occurs to the inside of any sentential connective, it will be the horizontal attached to a compound's atomic parts that will regularize it. Heck and Lycan conclude, however, that they are unable to suggest any motivation for applications of horizontals additional to those attached to atomic expressions, because any expression containing horizontalized atomic parts is a genuine sentence and is in no need of regularizing (see their *1979*, 487–489, and its note 16). Their observation that, on the composite view, it is the regularizing function of the horizontal that is important, may be expanded upon in a way that shows the connective to be one shared by other many-valued logics and one with important implications for a language's ability to formulate classical logic.

Hans Herzberger and others have suggested that the horizontal and Bochvar's three-valued truth operator have much the same roles in the representation of classical logic. The suggestion here is not, I think, a historical claim that Frege is to be thought of as having Bochvar's representation consciously in mind, but rather that later thinkers can explain by appeal to Bochvar's theory how the propositional fragment of Frege's language under the composite interpretation can be shown to have a classical logic.[14] The link between Frege's semantics and more traditional many-valued logic may be sketched as follows. It ought to be possible on the composite view to construct grammatical sentences in which the primitive connectives occur without flanking horizontals. But unhorizontalized connectives are systematically avoided in Frege's text, and Kneale and Kneale remark that on the composite interpretation this absence needs an explanation. They suggest that we should attribute to Frege a grammatical prohibition against the occurrence of unhorizontalized connectives (*1962*, 504).

But it is not necessary to restrict the grammar in this way, and probably undesirable to do so inasmuch as Frege himself never formulates this rather precise syntactic principle. The semantics as it stands can already explain the absence. Recall that part of the composite analysis is the ancillary thesis that the functions interpreting the connectives are defined for just truth-values. These functions are analogous to predicates in three-valued significance theory. Such predicates are said to apply truly to some objects, apply falsely to others, and to be meaningless when applied to yet others. Predications of the last sort are category mistakes and in three-valued semantics are assigned the third value representing meaninglessness. (See, for example, Halldén *1949* and my *1975*.) Similarly, functions for the connectives yield *T* when applied to some objects, *F* when applied to others, and are undefined when applied to yet others. Moreover, if the semantic value of a part of an expression is undefined, the value for the whole is too because the well-defined set of arguments needed for its truth-functional computation is lacking.[14]

Likewise, Bochvar's three-valued primitive operators for the connectives have the property that the whole has the third value if any of the parts do. A striking feature of such semantics is that within it classical validity is not susceptible to its standard analysis. There are literally no tautologies in the sense of sentences that are always *T*, because in an interpretation in which any of the atomic parts of a sentence are meaningless the whole will be, too. Bochvar's solution and the one that may be applied to Frege's propositional fragment with the horizontal is to seal off the atomic sentences with a truth-operator before combining them with other connectives. Since the horizontalized atomic sentences are bivalent, compounds of them with the structure of classical truths will in fact be *T* under every assignment of values to the parts.

I shall now present details of the mathematical theory and do so in a rather more abstract form than Bochvar so as to show how its methods are applicable to a wide range of many-valued languages, including the Fregean fragment. Results are stated in terms of the basic properties of logical matrices on which they depend. Proofs consist of fairly straightforward applications of the definitions and are to be found in more detail but in a somewhat less abstract form in my *1977b*. That the composite categorial interpretation of the *Grundgesetze* should inspire these methods is to its

14 Stevenson *1973* seems to be presupposing some such interpretation of the horizontal in his discussion of how quantification over objects for which the quantifier is not defined (a sort of category mistake) is nevertheless bivalent for Frege.

credit, but as historical exegesis it is really quite implausibly complex to be seen as resting unspoken in Frege's mind. Seeing it in Frege is, I think, representative of a tendency found in several aspects of the categorial interpretation to attribute to Frege ideas of twentieth-century logic which he could not plausibly have held.

Let the notions of structure, morphism, constructible operation, language, syntax, semantics and reference relation be as defined in Section 2, with the exception that a semantics is to be enlarged to include as an additional element a subset of its domain of semantic references. This new element is known as the set of *designated values* for the semantics and the resulting structure is called a *logical matrix*. Let

$$S = \langle F, O_1, \ldots, O_n \rangle$$

be an arbitrary syntax. We assume F is the closure under the operations of some set of atomic formulas AF, and that accordingly F consists of the union of all the domains and ranges of the operations and AF consists of F minus the union of all the ranges. Let

$$M = \langle U, D, \phi_1, \ldots \phi_n \rangle$$

be an arbitrary matrix of the same type with set of designated values D. We let L be the langauge $\langle S, M \rangle$ and say that a subset X of F *entails* a formula p of F in L, briefly $X \Vdash_L p$, iff for any reference relation R of L if all q of X are such that $R(q) \varepsilon D$, then $R(p) \varepsilon D$. Langauges distinct from L and their corresponding features in syntax and semantics are distinguished by prime markers. The theory consists of formulating the conditions under which a many-valued language employing the horizontal has the same entailment relation as that of classical logic. It proceeds by first laying down a series of key concepts for the comparison of entailment between two many-valued languages and then stating metatheorems formulating the comparison in terms of that concept. The concepts are then applied to the special cases of the languages of Bochvar, Frege and classical logic.

A syntax S is *part* of a syntax S' iff $F \subseteq F'$ and for each i, $O_i \subseteq O_i'$. If f is an arbitrary function on a set A, then by $f(A)$ is meant $\{f(x) \mid x \varepsilon A\}$. Let L be $\langle S, M \rangle$ and L' be $\langle S', M' \rangle$.

If S is a part of S' and $M = M'$, then $X \Vdash_L p$ only if $X \Vdash_{L'} p$. (T1)

If S is a part of S', $p \varepsilon F$, and $M = M'$, then $X \Vdash_{L'} p$ only if $X \Vdash_L p$. (T2)

If $M = M'$, S is homomorphic to S' under some h, and $h(p) \varepsilon AF'$ whenever $p \varepsilon AF$, then $h(X) \Vdash_{L'} h(p)$ only if $X \Vdash_L p$. (T3)

If $M = M'$ and S is isomorphic to S' under some h, then $X \| \ p$ iff $h(X) \Vdash_{L'} h(p)$. (T4)

A matrix M is a *matrix extension* of a matrix M' of the same type iff $U \subseteq U'$, $D \subseteq D'$, and for each i, ϕ_i is non-empty and $\phi_i — \phi_i'$. The notion of morphism between semantic structures originally defined without provision for sets of designated values is extended to matrices by saying that h is a morphism from M to M' iff it is one between $<U, \phi_1, \ldots, \phi_n>$ and $<U', \phi_1', \ldots, \phi_n'>$. A morphism h from M to M' is said to *preserve designation* iff whenever $x \varepsilon D$, $h(x) \varepsilon D'$, and to preserve *non-designation* iff whenever $h(x) \varepsilon D'$, $x \varepsilon D$. Then

If $S = S'$ and M' is a matrix extension of M, then $X \Vdash_{L'} p$ only if $X \Vdash_{L} p$. (T5)

If $S = S'$, and M is homomorphic to M' under a relation that preserves designation and non-designation, then $X \Vdash_{L'} p$ only if $X \Vdash_{L} p$. (T6)

If $S = S'$ and M is homomorphic to M' under a relation that is onto U and that preserves designation and non-designation, then $X \Vdash_{L} p$ iff $X \Vdash_{L'} p$. (T7)

Let L be a *contraction* of L' iff S is a part of S', $U = U'$, $D = D'$, and for each i, either ϕ_i is empty or $\phi_i = \phi_i'$. (Usually AF will be a subset of AF'.) L' is a *conservative extension* of L iff S is part of S' and the entailment relation $\Vdash_{L}$ is identical to $\Vdash_{L'}$ restricted to F. Then

If L is a contraction of L', then L' is a conservative extension of L. (T8)

Let L' be a *definitional extension* of L iff L is a contraction of L', $AF = AF'$, and for each i, whenever both ϕ_i is empty and ϕ_i' is non-empty, ϕ_i' is constructable in M. Then

If L' is a definitional extension of L, then there is a translation function t mapping F' onto F such that $X \Vdash_{L'} p$ iff $t(X) \Vdash_{L} t(p)$. (T9)

For Bochvar's special application of these ideas we first define four languages. For simplicity, we shall employ the same symbol for both a syntactic operation and its semantic correlate. Let LC be the langauge $<SC, MC>$ for *classical logic* in which

$$SC = <FC, \&, \cup, \supset, \sim>$$

is a classical syntax constructed in the usual way from a set AFC of atomic formulas and

$$MC = <\{T, F\}, \{T\}, \&, \cup, \supset, \sim>$$

in which the operations are the classical truth-functions. The language LI of the *internal connectives* is $<SI, MI>$ such that

$$SI = \langle FI, \wedge, \vee, \rightarrow, \neg, — \rangle$$

is constructed from a set of atomic formulas AFI that is identical to AFC by means of the notational variants $\wedge$, $\vee$, $\rightarrow$, and $\neg$ for the usual connectives and by the truth-operator —, and

$$MI = \langle O, \{T\}, \wedge, \vee, \rightarrow, \neg, — \rangle$$

in which O is any set of values that include the classical truth-values T and F as elements, and the operations are generalizations of Bochvar's three-valued internal connectives defined as follows. Let an operation ϕ_n of a logical matrix be said to be *normal* iff for the corresponding operation ϕ_n' of the classical matrix MC, whenever $x_1, \ldots x_m$ are in $\{T, F\}$, then

$$\phi_n(x_1, \ldots, x_m) = \phi_n'(x_1, \ldots, x_n).$$

Also ϕ_n is said to be *sensitive* iff

$$\text{whenever } \{x_1, \ldots, x_m\} \subseteq \{T, F\}, \text{ then } \phi_n(x_1, \ldots x_m) \notin \{T, F\}.$$

The only conditions we impose on the generalized internal operations $\wedge$, $\vee$, $\rightarrow$, $\neg$ are that they be both normal and sensitive. The operation — is defined (like Frege's horizontal) as pairing T with T and everything else in O with F. It follows that if $O = \{T, F, N\}$, the internal connectives have the three-valued truth-tables stipulated by Bochvar:

	$\neg$	$\wedge$	T	F	N	$\vee$	T	F	N	$\rightarrow$	T	F	N	—
T	*F*		*T*	*F*	*N*		*T*	*T*	*N*		*T*	*F*	*N*	*T*
F	*N*		*F*	*F*	*N*		*T*	*F*	*N*		*T*	*T*	*N*	*F*
N	*T*		*N*	*N*	*N*		*N*	*N*	*N*		*N*	*N*	*N*	*F*

We now extend LI to a language $LI+ = \langle SI+, MI+ \rangle$ such that

$$SI+ = \langle FI+, \wedge, \vee, \rightarrow, \neg, —, \hat{\wedge}, \hat{\vee}, \rightarrow, \hat{\neg} \rangle$$

is constructed from the set $AFI+$ of atomic formulas identical to AFC by the connectives of SI and a second set $\hat{\wedge}, \hat{\vee}, \rightarrow, \hat{\neg}$ of variants for the usual connectives, and

$$MI+ = \langle O, \{T\}, \wedge, \vee, \rightarrow, \neg, —, \hat{\wedge}, \hat{\vee}, \Rightarrow, \hat{\neg} \rangle$$

has the operations of MI as previously defined and the new operations $\hat{\wedge}, \hat{\vee}, \Rightarrow, \hat{\neg}$ defined in terms of the internal operations:

$$x \,\hat{\wedge}\, y = —x \wedge —y; \; x \,\hat{\vee}\, y = —x \vee —y; \; —x \rightarrow —y;$$
$$\text{and } \hat{\neg} x = \neg —x.$$

To facilitate application of the previous ideas we shall regard an n place structure as identical to any $n + m$ place structure that is just like it except for having the empty set at positions $n + 1, \ldots, m$. It follows then that SI is a part of $SI+$, and that $LI+$ is a definitional extension of LI. The language LE of the *external connectives* is then defined as $<SE, ME>$ such that

$$SE = <FE, \hat{\hat{\wedge}}, \hat{\hat{\vee}}, \hat{\Rightarrow}, \hat{\Rrightarrow}>$$

and FE is the closure of the set of atomic formulas AFE identical to AFC by the syntactic operations $\hat{\wedge}$, $\hat{\vee}$, $\Rightarrow$, and $\Rrightarrow$ of $SI+$, and the operations of SE are the restrictions of corresponding externalized operations of $SI+$: $\hat{\hat{\wedge}}$ is the restriction of $\hat{\wedge}$ to FE, etc.; and ME is defined in terms of elements of $MI+$ to be

$$<O, \{T\}, \hat{\wedge}, \hat{\vee}, \Rightarrow, \Rrightarrow>.$$

It follows that LE is a contraction of $LI+$. Bochvar's result may now be demonstrated as an application of the previous theory to the languages just defined:

There is a translation function t from FC to FI such that $X \Vdash_{LC} p$ only if $T(X) \Vdash_{LI} t(p)$. (T10)

Proof. Since SE is isomorphic to SC under some h, it follows by T4 that

$$X \Vdash_{<SE, MC>} p \text{ iff } h(X) \Vdash_{LC} h(p).$$

Since ME is homomorphic onto MC preserving designation and non-designation, (T7) implies

$$X \Vdash_{<SE, MC>} p \text{ iff } X \Vdash_{LE} p.$$

Also since LE is a contraction of $LI+$, we know by T8 that $X \Vdash_{LE} p$ only if $X \Vdash_{LI+} p$.

Since $SI+$ is a definitional extension of SI, by T9 there is a h' such that

$$X \Vdash_{LI+} p \text{ iff } h'(X) \Vdash_{LI} h'(p).$$

The preceding together entail that

$$h(X) \Vdash_{LC} h(p) \text{ only if } h'(X) \Vdash_{LI} h'(p),$$

for any X and p of FE. Assume for arbitrary Y and q of FC that $Y \Vdash_{LC} q$. Then since h is an isomorphism, its inverse h^{-1} is also, and

$$h(h^{-1}(Y)) \Vdash_{LC} h(h^{-1}(q)).$$

Then by the foregoing

$$h'(h^{-1}(Y)) \Vdash_{LI} h'(h^{-1}(q)), \text{ and } t(p) = h'(h^{-1}(p))$$

for arbitrary p of FC is the desired translation. End of Proof.

It is possible to apply the foregoing theory in two ways to the language of the *Grundgesetze* so as to show that there is a form in which the intuitive validities of classical logic are respected. Which application is appropriate depends on which of the two general readings is given to Frege's connectives; the non-composite one (which I have argued is preferable) or the composite one (in which the horizontal operator genuinely occurs before and after each occurrence of the other connectives). It is only under the composite reading, however, that the representation of classical logic takes the form of Bochvar's as given in (T10). It must be stressed, however, that neither application of the theory is explicitly suggested by Frege's own words. Rather, they must be viewed as later theoretical extensions of his semantics.

On the non-composite reading, the representation of classical logic within the *Grundgesetze* is straightforward. First a propositional syntax is abstracted from Frege's original as fully defined, for example, in Section 5. This would consist of a set of formulas constructed in the usual way from a set of atomic formulas and the basic connectives. Frege's negation and conditional could be augmented by conjunction and disjunction (with their standard definitions) without altering the theory in any substantial way. The syntax would essentially be that of the external connectives *SE*. Semantically, the interpretations of these connectives would remain as stipulated in Section 5. There, negation and the conditional are interpreted by what are here the external operations of the matrix *ME*, and under the standard definitions in terms of negation and the conditional, both conjunction and disjunction would also stand for external operations. On this reading then the propositional fragment of the *Grundgesetze* is essentially the external language *LE*. But since the syntax of *LE* is isomorphic to that of the classical language *LC*, and its matrix *ME* is homomorphic-onto, preserving designation and non-designation, its entailment relation is identical to that of classical logic. Here the horizontal has played no role in establishing that Frege's propositional logic is classical.

On the non-composite interpretation of the connectives, however, the horizontal functions like Bochvar's truth operator in *LI*. This time we abstract from the *Grundgesetze* a syntax that includes the horizontal as well as negation and the conditional, and we may add by definition conjunction and disjunction as well. The resulting syntax is essentially the internal *SI*. Semantically, these operations are to be understood in the manner of the standard interpretation of the composite reading. They are classical if their arguments are well-defined, and they are undefined otherwise. Formally, all that is required is that the operations be normal and sensitive. The horizontal is interpreted by an operation mapping T to T and anything else to F, and the resulting matrix is essentially that for the internal connectives *MI*. If the state of being undefined is represented by the assignment of a third value, the operations are those of the internal three-valued matrix, and the resulting language abstracted from the *Grundgesetze* as its propositional fragment would be Bochvar's internal language. Its logic may be shown to be classical in the manner of (T10).

There is, however, one way in which Frege's text suggests that the application of the theory should be slightly different from Bochvar's. Bochvar does not allow for the externalized sentences of *LE* to contain any syntactic structure except that generated by the external connectives themselves. Frege, on the other hand, seems to view syntactic units even before the application of a horizontal as grammatically complex and allows for the possibility that complex expressions might be taken as atomic elements in an external language.[15]

Accordingly, we define a new language

$$LE+ \; = \; <SE+, ME>$$

in which the new syntax is constructed as follows. First, let us say here that in a syntax p is a *part* of q iff q is atomic and $p = q$, or q is some $\phi(r_1, \ldots, r_n)$ and p is a part of r_1 or . . . or p is a part of r_n. Let us say p is *horizontalized* (in *SI*) iff for any q, if q is a part of p and q is not some ——r, then q is part of some s which is also a part of p. Let $FE+$ be that set defined as the closure of the set $AFE+$ under the syntactic operations $\dot{\wedge}$, $\dot{\vee}$, $\dot{\rightarrow}$, $\dot{\neg}$. Here $AFE+$ is defined as the set of all ——p such that $p \varepsilon FI$ and p is not horizontalized, and the syntactic operations are just a notational variant on the usual connectives. We may now define an isomorphism from the classical syntax *SC* onto *SE*. Since *AFC* (the atomic formulas of *SC*) and

$$U = \{p \mid p \; \varepsilon EF \; \& \; p \text{ is not horizontalized}\}$$

are both denumerable, let them have a natural order. For $p_i \varepsilon AFC$, let $h(p_i)$ be ——q such that q is the i-th element of U. For $p = q \wedge r$, let $h(p)$ be $q \dot{\wedge} r$, and similarly for the other operations. Clearly h is an onto homomorphism, and a straightforward induction shows it is 1–1. Moreover $SE+$ is part of $SI+$, and the proof of T10 continues to hold if $SE+$ is substituted for *SE*, $LE+$ for $<SE, ME>$ and $LEC+ = <SE, MC>$ for *LEC*.

The importance of the horizontal in the methods of T10 can be illustrated in a slightly different way. We may use the horizontal to produce a matrix extension of the classical matrix with the result that the first two steps of the proof of (T10) can be shown to follow from a general result about horizontals. For two matrices M and M' we say that an *M horizontal relative to M′* is any function —— on U' such that (1) —— restricted to U is the identity function on U, and (2) for any x, $x \varepsilon D'$ iff ——$x \varepsilon D$. A function ϕ_i' of M' is M-external iff

$$\phi_i'(x_1, \ldots x_n) = \phi_i(\text{——}x_1, \ldots, \text{——}x_n),$$

where —— is an M horizontal relative to M'. Then

15 This difference in the external representations of Bochvar and Frege (on the composite interpretation) was first pointed out to me by Hans Herzberger.

(T11) If (1) S is homomorphic to S' under h and $h(p)\varepsilon AF'$ for any $p\varepsilon AF$, and
(2) there is some M'' such that
(a) M is homomorphic to M'' under some h' that preserves designation and non-designation,
(b) M' is a matrix extension of M'', and
(c) all ϕ_i' of M' are M-external,
then (3) $X \| \vdash_{L} p$ iff $h(X) \| \vdash_{L'} h(p)$.

Clearly, —— as previously defined in *MI* is an *MC* horizontal and the operations of *ME* are *MC* external. Moreover, under some h, *SE* is homomorphic to *SC* in a way that maps atomic formulas onto atomic formulas. The identity function also maps *MC* onto itself in the manner described by (2). Thus, we may apply (T11) to the languages of Bochvar and Frege:

For some homomorphism h from *SE* onto *SC*, (T12)

$$X \| \vdash_{LE} p \text{ iff } h(X) \| \vdash_{LC} h(p).$$

For some homomorphism h from *SE*+ onto *SC*, (T13)

$$X \| \vdash_{LE+} p \text{ iff } h(X) \| \vdash_{LC} h(p).$$

Acknowledgements

This research was supported in part by grants from the University of Cincinnati Research Council and Taft Fund. I am grateful to the referees for several useful references.

Bibliography

Works by Frege

In the paper, page numbers are cited in the form y/z, in which y is the page in the German original and z the page in the English translation. References in the *Grundgesetze* are in addition prefixed by the number x of the containing paragraph, in the form $x{:}y/z$.

1892–1895 '[Comment on sense and meaning]', in Hans Hermes *et alii* (eds.), *Gottlob Frege, posthumous writings* (Chicago: Univ. of Chicago, 1979); translation of *Nachgelassene Schriften und wissenschaftlicher Briefwechsel*, vol. I (Hamburg: Felix Meiner, 1969).

1891 'Function and concept', in Peter Geach and Max Black, *Translations from the philosophical writings of Gottlob Frege*, 2nd ed. (Oxford: Blackwell, 1966); from the original 'Function und Begriff' (Jena).

1964 The basic laws of arithmetic, trans. and ed. Montgomery Furth (Berkeley: Univ. of California); translation of parts of *Grundgesetze der Arithmetik*, vol. I (1893) and vol. II (1903) (Jena: Verlag Hermann Pohle).

Letter from Frege to Husserl, 24 May 1891 in Gottfried Gabriel *et alii* (eds.), Gottlob Frege, *Philosophical and mathematical correspondence* (Chicago, Univ. of Chicago Press, 1980); translation of *Nachgelassene Schriften und wissenschaftlicher Briefwechsel*, vol. II (Hamburg: Felix Meiner, 1976).

1892 'On sense and reference', in Geach and Black; translation of 'Über Sinn und Bedeutung', *Zeitschrift für Philosophie und philosophische Kritik*, *100* (1892), 25–50.

Other references

Bar-Hillel, Y., Gaifman, Y. and Shamir, E. *1964* 'On categorial and phrase-structure grammars', in Bar-Hillel, *Language and information*, New York (Addison-Wesley), 99–115

Merrie Bergmann, *1981* 'Translator's introduction' to Bochvar *1937*.

Bochvar, D.A. *1937* 'On a three-valued logical calculus and its application to the analysis of the paradoxes of the classical extended functional calculus', trans. in *History and philosophy of logic*, **2** (1981), 87–112.

Cresswell, M.J. *1973 Logics and language*, London (Methuen).

Currie, Gregory, *1981 Frege: an introduction to his philosophy*, Sussex (Harvester).

Dummett, Michael *1981 Frege: philosophy of language*, 2nd ed. London (Duckworth).

——— *1982 The interpretation of Frege's philosophy*, London (Duckworth).

Dunn, J.M. and Belnap, Nuel *1968* 'The substitution interpretation of the quantifiers', *Nous*, **2**, 177–185.

Furth, Montgomery *1966* 'Editor's introduction', to his translation of the *Grundgesetze*, v–lvii.

Halldén, Sören *1949 The logic of nonsense*, Uppsala (A.-B. Lundequistska Bokhandeln).

Heck, William C. and Lycan, William G. *1979* 'Frege's horizontal', *Canadian journal of philosophy*, **9**, 479–492.

Herzberger, Hans G. *1970* 'Truth and modality in semantically closed languages', in Martin, R.L. (ed.), *The paradox of the liar*, New Haven (Yale), 25–46.

Kaplan, David *1964* 'Foundations of intensional logic', Ph.D. dissertation (UCLA).

Kneale, William and Martha *1962 The development of logic*, Oxford (Clarendon).

Kripke, Saul *1976* 'Is there a problem about substitutional quantification?', in Evans, G. and McDowell, J. (eds.), *Truth and meaning*, Oxford (Clarendon), 325–419.

Martin, Edwin Jr., *1971* 'Frege's problems with "The concept horse"', *Critica*, **5**, 45–64.

——— *1974* 'A note on Frege's semantics', *Philosophical studies*, **25**, 441–443.

——— *1982* 'Referentiality in Frege's *Grundgesetze*', *History and philosophy of logic*, **3**, 151–164.

Martin, John N. *1975* 'A many-valued semantics for category mistakes', *Synthese*, **31**, 63–83.

——— *1977a* ' Fregean interpretation of the quantifiers', *Logique et analyse*, **20** 122–128.

——— *1977b* 'Many-valued logic, classical logic, and the horizontal', in *Proceedings of the Seventh International Symposium on Multiple-valued Logic*, London (Institute of Electrical and Electronic Engineers), 84–88.

Potts, Timothy C. *1973* 'Fregean categorial grammar', in Bogdan and Niiniluoto, I. (eds.), *Logic, probability and language*, Dordrecht (Reidel), 245–284.

——— *1977* 'Fregean grammar: a formal outline', *Studia logica*, **37**, 7–26.

Quine, William Van Orman *1959 Methods of logic*, rev. ed. New York (Holt).

Sluga, Hans D. *1980 Gottlob Frege*, London (Routledge and Kegan Paul).

Stevenson, Leslie *1973* 'Frege's two definitions of quantification', *Philosophical quarterly*, **23**, 207–223.

Thiel, Christian *1968 Sense and reference in Frege's logic*, Dordrecht (Reidel); translation of *Sinn und Bedeutung in der Logik Gottlob Freges* (Meisenheim am Glan: Verlag Anton Hain, 1965).

Wallace, John Roy *1964* 'Philosophical grammar', Ph.D. dissertation (Stanford).