

---

# How to Hintikkize a Frege

FABIEN SCHANG

---

## 1 Frege's theory of meaning and its critics

### 1.1 A one-sorted semantics

Frege's rationale is to uphold his own theory of meaning by means of an argument by contraposition  $(\alpha \supset \psi) \supset (\neg\psi \supset \neg\alpha)$ . It runs as follows: If

( $\alpha$ ) every sign (*Zeichen*) of an arbitrary sentence includes both a sense (*Sinn*) and a reference (*Bedeutung*),

then

( $\psi$ ) no substitution of equivalent components (with the same reference) alters the whole.

Now if

( $\neg\psi$ ) at least one counterexample of this substitution *salva veritate* can be found,

then

( $\neg\alpha$ ) Frege's theory of meaning collapses.

The kinds of argument under review are the following: Frege asserts both  $\alpha$  and  $\psi$ ; Kripke denies  $\alpha$  and asserts  $\psi$ ; Hintikka asserts  $\alpha$  and denies  $\psi$ .

#### FIRST OUTLINE OF FREGE'S THEORY OF MEANING

One-sorted semantics + Compositionality  
=  
extensional (standard) logic

The contrast between two modes of meaning (conveying information) through every linguistic item, that is the one between “sense” and “reference”,

primarily cancels some naïve theory of meaning, according to which every sign would purport to tag or label one object (recall Quine’s museum myth of meanings [13], or Saint Augustine’s assimilation between meaning and pointing something out [14]). How to account for the information gain from “ $a = a$ ” to “ $a = b$ ”, if meaning is reduced to reference?

For this purpose, Frege’s philosophy of language relies upon two main assumptions: compositionality, and extensionalism. According to the former, the reference of a complex sentence is determined by the reference of its components and sense (reference) cannot contribute to determinate a reference (sense). A crucial point for the following: the references of sentences are truth-values, among the True and the False. It is for this reason that Frege’s logic, because it has primarily occurred as a science of truth, is uniquely concerned with references and never deals with matters of sense concerning the informative value conveyed by a sign. The close connection between logic and truth has been famously claimed by Frege [4: 170] in these introductory words:

The word “true” indicates the aim of logic as does “beautiful” that of aesthetics or “good” that of ethics. All sciences have truth as their goal; but logic is also concerned with it in a quite different way from this. It has much the same relation to truth as physics has to weight or heat. To discover truths is the task of all sciences; it falls to logic to discern the laws of truth.

Because logic deals with truth, and truth is a reference, the very idea of a logic of sense cannot be but troublesome. Accordingly, a way out is to revisit the logical notions of truth and sense.

Following his one-sorted semantics, every sign expresses a sense and refers to an object: individual variables ( $x, y, z, \dots$ ) refer to individuals while expressing an individual concept; predicate variables ( $P, Q, R, \dots$ ) refer to properties while expressing concepts; sentence variables ( $p, q, r, \dots$ ) refer to truth-values while expressing “propositions”. Every sentence (*Satz*) includes both a thought as its sense (the *Sinn*, as the *Gegebenheitsweise* or way of giving references) and a truth-value as its reference (the True or the False, among its two possible objects). For according to Frege, the thought (*Gedanke*) is the sense of a sentence and corresponds to the other name of the proposition. In this sense, thought is an objective item of language and does not constitute a private representation (a “*Vorstellung*”). Now the process of internalization leads to an essential problem for analytical philosophy, concerning the boundary between objective and subjective domains. To put it in words: how far can we turn subjective contents of thought into objective terms of logic? We return to this point in the end of the paper.

Meanwhile, Frege’s theory of language entails that any two sentences with the same truth-value refer to the same thing; these merely differ by their sense. According to the so-called “rule of substitution” that prevails for every sign designating identical objects, it results in three variants of extensionalism – this is Frege’s second main assumption – for three distinct sorts of terms (individuals, predicates, and sentences):

Leibniz’s Law:  $\forall x\forall y\forall P[(x = y) \supset (Px \equiv Py)]$   
 Principle of Coextensivity:  $\forall P\forall Q\forall x[(P = Q) \supset (Px \equiv Qx)]$   
 Principle of Extensionality:  $\forall p\forall q\forall\delta[(p \equiv q) \supset (\delta p \supset \delta q)]$  (where  $\delta$  is a context variable)

Is this formal theory of language reliable in any case? As is known, there seems to be non-extensional, i.e. intensional contexts  $\delta$  in which the answer is negative. Let us consider the case of referential opacity in belief contexts. Although Tom believes that Cicero wrote *De Senectute* (symbols:  $B_{Tom}F(a)$ ), and Cicero denoted the same individual as Tullius (symbols:  $a = b$ ), Tom may not believe that Tullius wrote *De Senectute* (i.e.  $\neg B_{Tom}F(b)$ ). A counterexample seems to be found here to substitutivity of identicals, and so does accordingly for extensionalism. Formally speaking, the reasoning

$$\frac{B_{Tom}F(a) \quad a = b}{B_{Tom}F(b)}$$

does hold in Frege’s extensional logics but should not in an informal reasoning of natural language.

By analogy with quotation or indirect discourse contexts, Frege’s way out consists in making a distinction between direct (*gerade*) and indirect (*ungerade*) references: in a belief context, terms do not have their usual reference; sentences do not refer to truth-values any more but, rather, they express thoughts, so that the thought is that which is *expressed* in a direct context and *referred* to in an indirect context. Now following the principle of compositionality, no thought can contribute to the truth-value of a sentence, and the intensional counterexample is thus reduced to a “semiotic illusion” (i.e. a confusion between two tasks in a sentence) [3:116].

This arises from an imperfection of language, from which even the symbolic language of mathematical analysis is not altogether free.

The contemporary intensional or non-classical logics have been largely motivated by this problem of referential opacity; now the previous Fregean argument seems to challenge their very legitimacy: don’t their logical forms

result from a sort of category mistake, i.e. a symbolic confusion between the sense of a sign and its reference? At least three basic features of a “semantic shift” can be actually found in the current practice of these intensional logics: a treatment of incomplete thoughts as complete thoughts; a rejection of the one-sorted semantics; an internalization of sense, notably in epistemic logic where the truth-conditions of beliefs are specified.

Would such a process have something meaningless in it?

## 1.2 Objections to the one-sorted semantics

### Complete and incomplete thoughts

Non-classical logic is generally seen as a logic in which one of the cornerstones of Frege’s theory of meaning is discarded, whether for one-sorted semantics or compositionality. Throughout [3], the author means by incomplete thoughts (as parts of thoughts) the various sorts of sentences which currently characterize such modal logics as temporal, relevant, causal or counterfactual logics.

The contrast between Fregean and modal logics relies upon the logical form to be assigned to modalities: in the case of a modal sentence  $\delta p$ , with an operator  $\delta$  and a that-clause  $p$ ,  $p$  is incomplete for Frege and complete for a modal logician. In other words, the modal logician determines the truth-value of the modal sentence according to the truth-value of its component  $p$ ; Frege does not proceed in this way, because  $p$  does not have its usual reference in a modal context and, therefore, does not have any truth-value as its reference. How to account for such a difference in the logical analysis?

Let us turn again to the case of doxastic modal logic, i.e. the part of epistemic modal logic devoted to the concept of belief (“doxa”). In the sentence “Tom believes that Cicero wrote *De Senectute*”, the reference of the that-clause “Cicero wrote *De Senectute*” is the sense of the whole sentence and, thus, does not determine the extension of the corresponding belief: our knowing whether Cicero did write *De Senectute* or not does not decide whether Tom does believe it or not. Syntactically speaking, the Fregean distinction between the usual and unusual reference of a sentence turns the initial sentence into a name of that-clause: the logical form of these indirect contexts of discourse is not  $\delta F(a)$  but  $\Delta“F(a)”$ , where  $\Delta$  is no longer an operator  $\delta$  applied to a sentence  $F(a)$  but, rather, a predicate attached to an individual name “ $F(a)$ ”; such a transformation helps to extensionalize intensional contexts and has been used by other authors such as Carnap or Quine. In [1], for instance, Carnap handles the notion of necessity as a predicate for analyticity to which a name of sentence is attached. But the trouble with this extensional treatment concerns the quantified open sentences: the aforementioned paraphrase cannot be applied anymore, given

that the transformation of an individual variable  $x$  into an individual name “ $x$ ” leads to a nonsensical quantified sentence:  $(\exists x)\Delta(\text{“F}(x)\text{”})$ , where the relation between the bound variable  $x$  and the variable within quotes “ $F(x)$ ” is like the non-semantic relation between the word “cattle” and its composing phonem “cat”.

The limitation of Frege’s formal theory of language needs some change, consequently: either compositionality need be abandoned, or the trouble comes from the one-sorted semantics. The latter is questioned by Kripke’s theory of rigid designation. Can the problem of referential opacity be settled within such a non-Fregean semantics?

### A two-sorted semantics

To the problem of substitutivity in opaque contexts, Saul Kripke [11] and such other “New Theoricians of Reference” (thereafter: NTR) as Ruth Barcan-Marcus, Keith Donnellan, Nathan Salmon, Dagfinn Føllesdal replied by a two-sorted semantics: proper names have a reference but no sense (a reminiscence of John Stuart Mill’s theory of proper names), so that not every term includes both a reference and a sense. In order to emphasize this univocal relation between a proper name and its reference, Kripke and Barcan Marcus endorsed the thesis of necessity ( $N$ ) for identity sentences:

$$\forall x\forall y[(x = y) \supset N(x = y)]$$

Let us borrow an example from alethic modal logic, where the central modalities are those of necessity, possibility, or contingency. If 9 is necessarily greater than 7, and 9 is the number of planets, does it entail that the number of planets is necessarily greater than 7? Unless a contingent truth of astronomy turns out to be a necessary truth, we are led to conclude that the following inference does not hold and falsifies extensionalism:

$$\frac{N(9 > 7) \quad 9 = \text{the number of planets}}{N(\text{the number of planets} > 7)}$$

However, the theory of rigid designation sustains the substitutivity of identicals when designated by a *proper name*: it claims that a proper name has the same reference whatever the context of discourse may be (i.e. in every possible world) and keeps contributing to the reference of the composing sentence. The failure of the above inference is due to the fact that one of the substituted terms is not a proper name but, rather, a *definite description* (viz. “the number of planets”); the success condition of substitution depends upon the choice of the designating sign, and not only the designated object.

Assuming that this two-sorted semantics does work for necessity contexts, how to avoid the failure of substitutivity in intensional contexts like belief? In such a case, the rigidity of proper names and the view of a unique referring task for proper names don't seem to be sufficient: the substitution of proper names needn't preserve the truth-value (the reference) of the modified sentence. Note also that Hintikka was a conspicuous opponent of NTR, claiming that the reference of proper names is not "tagged on their forehead" but results from the very process of individuation performed by the epistemic agent [9]. At the same time, one can agree with Smith [15] that such a critic may seem to be unfair, given that Kripke's theory merely concerned the peculiar context of *alethic* modalities.

After Frege and Kripke, Hintikka's position appears to be more intuitive when applied to doxastic contexts: it maintains Frege's one-sorted semantics while rejecting its extensional principle of substitutivity. But then how to account for such a seemingly inconsistent stance, both asserting  $\alpha$  and denying  $\psi$  in the initial Fregean implication  $\alpha \supset \psi$ ? The following shows that the implicational form (i.e. the scheme  $\alpha \supset \psi$ ) in the reasoning pattern of [3] relies upon a Fregean theory of truth that is not shared by Hintikka.

## 2 Hintikka's internalization of Frege's sense

By contrast to the supporters of NTR, Hintikka does not argue for a two-sorted semantics and sticks to Frege's one-sorted version. Yet his view of logic gives rise to two major alterations, namely: (1) the use of a "possible-worlds semantics", as a co-univocal relation (one-many) between language and the world and according to which Frege's *Sinn* comes to be synonymous with a multiple reference [6]; (2) the claim of an affinity between Frege's judgement-stroke, assertion, and belief [5]. The status of Frege's judgement-stroke remains an open question, however. For example, while taking its psychological, performative or illocutionary reading into account, Greimann [5:215] notes that

According to Wittgenstein, the sign  $\vdash$  is logically quite meaningless, because "in the work of Frege (and Russell) it simply indicates that these authors hold the propositions marked with this sign to be true" (cf. Wittgenstein (1921, sentence 4.442)). On this interpretation, the judgement-stroke is a psychological operator whose linguistic function is to express certain propositional attitudes.

Rather, Greimann opposes an alternative reading of the judgement-stroke as a truth-operator marking the value of the judgeable content; the present paper follows [10] and maintains the psychological version which associates the propositional attitude of belief to the judgement-stroke.

## 2.1 Frege's sense as referring

Valuation in a possible-world (*à la* Hintikka) semantics consists in assigning a reference (a truth-value) to a sentence among a number of possible worlds, those worlds standing for sets of sentences compatible with an agent's beliefs. In other words, any agent believes that  $p$  if and only if  $p$  is logically compatible with each of these (com)possible worlds; whereas she does not believe that  $p$  whenever, in at least one of these worlds,  $p$  is not compatible with the other sentences (i.e. whenever  $p$  denotes the False).

Such a process turns Frege's theory of truth upside down: the Fregean connection between reference and sense comes to be translated into a connection between single and multiple reference, and Hintikka's epistemic logic affords a formal device to account for the notion of sense in terms of multiple reference or truth-values assignment. To assign a reference in a world (a model) is to give a sentence its sense. The outcome of this multiple valuation is a kind of individuation for linguistic items, and the function purported to accomplish such a task does apply to the Fregean *Sinn*. For let  $f$  be an individuating function applied to references in models; then the way of givenness (*Gegebenheitsweise*) characterizing sense results in the mapping  $f(w_k, p) \mapsto \{0,1\}$ , where  $k$  is the label distinguishing the different possible worlds  $w$  and  $\{0,1\}$  the bivalent pair of references for sentences (i.e. their truth-values).

Despite this technical formalization in terms of possible worlds, Hintikka *does* still agree with Frege's views about the two preceding theses. That is: one-sorted semantics and compositionality are still in order in his formal approach of beliefs, and the unusual reference of a term (its sense) cannot determine its usual one. A clear difference occurs in the terms employed to define the unusual reference, i.e. the sense of a sentence within a doxastic context of discourse: Hintikka associates truth-values to that which constitutes the sense of this sentence, while Frege would have taken this manoeuvre to be absurd since the True and the False cannot be senses from his nonmodal perspective. Where does the difference lie between Hintikka's and Frege's theories of truth? Not in their theory of truth properly speaking but, rather, in their theories of naming truth, i.e. in the category of terms which can be said to be true. According to Hintikka, the difference is to be expressed formally by rejecting the one-world assumption [8]: there is not only one but several domains of reference to be taken into account for a semantic assessment of thoughts (i.e. beliefs, in Hintikka's epistemic parlance).

## 2.2 Belief as holding true

The vernacular link between belief, assertion, and holding true does justice to what Hintikka realized with his formal system, namely: a formalization of sense, depicting the process of (making) sense – giving a reference – within a logical calculus. Indeed, a close connection can be seen between the notion of belief as truth-value assignment in possible worlds and Frege’s notion of assertion as the act of holding true. Do we have the same reading of “truth” in both cases, and what of truth in a possible world as a variable reference among several ones (other accessible worlds)? If everything is all right far, then sense can be determined by truth-values in epistemic logic; however, these do not stand for one and the same object referred by a sentence, as was the case with Frege’s semantics (the Fregean truth-value occurs in a unique available world, i.e. the real world). Now according to what Frege meant by sense as a public, non-psychological item, the individuating function or intension advocated by Hintikka to pick out individuals within possible worlds can be rendered in Fregean terms of sense *and references* (in worlds). In this respect, Hintikka agrees with Frege but is at odds with him as to the manyworld assumption; furthermore, belief results from an individuating function based on the very notion of *Sinn*, so that any belief embedded in a judgement *is* the “thought” denoted by sentences (isn’t belief a form of thought, after all?).

On the other hand, Hintikka accounts for referential opacity in terms of variable modes of reference: substitutivity fails whenever a given term fails to be given one and the same reference (truth-value) in each possible world compatible with the speaker’s belief, i.e. when it is not individuated.

Let us note finally that, following van Heijenoort [16], the discrepancy between the supporters of logic as a universal language (the early Wittgenstein, Quine, Russell, Frege) and logic as a calculus (Skolem, Löwenheim, Peirce, Hintikka, Kripke) helps to mark the contemporary history of logic as a deep disagreement between universalists and algebrists. However, such a partition in the history of modern logic has been questioned by Peckhaus [12]: according to the latter, van Heijenoort would have wrongly claimed that the champions of logic as a calculus ignored the use of *quantifiers* while these already occurred in the algebrist works of Schröder. At the same time, van Heijenoort’s historiographic dichotomy still makes sense according to us from another perspective, viz. the model-theoretical perspective of many-worlds assumption compatible with the algebrist’s view, whereas the universalists would have been unable to think about several models because of their one-world assumption. If such a distinction between one- and many-worlds assumption is taken into account to explain the impossibility of formalizing Frege’s sense, the result is the following revised version of his



theory of meaning:

**SECOND OUTLINE OF FREGE'S THEORY OF MEANING**

One-sorted semantics + Compositionality + One-world assumption  
 =  
 extensional (standard) logic

If belief is a function determined by a set of reference assignments, and each of these assignments results in a sense, then reference can be taken to be a basic notion of semantics. Therefore, the process of internalization has helped to go upstream and account for the rise of belief in functional terms; can it now go on computing the very givenness of sense, or does such an operation remain beyond the realm of analysis? This should go beyond the area of logicians, and a division of labour occurs here between epistemic logic and epistemology. Can the following problem, i.e. how to acquire a method of individuation for the objects of discourse, be treated by a logical calculus?

### 3 Conclusion: the limits of internalization

The answer to the question “Can a logic of sense make sense?” is a qualified “Yes, but”: a logic of Fregean sense does make sense, provided that the unique domain of reference of Frege’s logical theory be modified. Contemporary formal semantics clearly corroborated this increasing use of formalism within theories of language, as witnessed by the actual prominence of the so-called non-classical or intensional logics. How far can such an increase be accepted? Can we hope a purely formal theory of the rules of meaning for ordinary language?

Michael Dummett blamed the philosophy of language for not motivating the operations at work in formal semantics [2:92]:

A semantic explanation, entirely formulated with the help of the concept of [reference], adequately shows how a truth-value is determined by its composing words and the way in which these are composed. Yet the semantic explanation is lacking, for it does not go backwards enough: it posits a connection between its primitive symbol and an appropriate denotative, without telling us how such a connection is established. In logic, this is not mandatory; for a theory of meaning, it is essential.

Does epistemic logic bring this drawback that Dummett takes to be essential? Not as it was devised by the realist-minded logician Hintikka: it merely provides for a mathematical structure intended to describe the prefabricated function of sense, but without telling anything about its very constitution. Yet it would be worthwhile to have a closer look at the teachings of an antirealist semantics, in order to account for the modes of meaning before referring and attempt a formalization of the process of assertibility; that is: the conditions under which a subject is in position to assign a reference to a term should be taken into account. These details do not occur within a realist semantics, since the truth-conditions of a sentence transcend the cognitive skills of subjects. Among these skills is the capacity to recognize one and the same reference through a number of distinct expressions used by the same individual (i.e. the capacity to perceive the reference *transparently*).

The debate around internalization is nothing less than a debate about the place of logic within the philosophy of language. A debate that remains open to the moving history of the discipline.

## Bibliographie

- [1] Carnap, R. *Logische Syntax der Sprache*, Wien: Julius Springer, 1934.
- [2] Dummett, M. Frege: *Philosophy of Language*, Harvard University Press, 1993.
- [3] Frege, G. “Sens et dénotation”, in *Ecrits logiques et philosophiques*, Paris, Seuil, collection “L’ordre philosophique”, 1971, translated by Claude Imbert (original text: “Über Sinn und Bedeutung”, in *Zeitschrift für Philosophie und philosophische Kritik* **100**, 1892, 25-50).
- [4] Frege, G.: “La pensée”, in *Ecrits logiques et philosophiques*, Paris, Seuil, collection L’ordre Philosophique, 1971, translated by Claude Imbert (original text: “Der Gedanke. Eine logische Untersuchung”, in *Beiträge zur Philosophie des deutschen Idealismus* **2**, 1918-1919, 58-77).
- [5] Greimann, D. “Judgement-operator as truth-operator : a new interpretation of the logical form of sentences in Frege’s scientific language, *Erkenntnis* **52**, 2000, 213-38.
- [6] Hintikka, J. “Modality as referential multiplicity”, *Ajatus* **20**, 1957, 49-64.
- [7] Hintikka, J. *Knowledge and Belief (An Introduction into the Logic of the Two Notions)*, Ithaca Press, N-Y, 1962.
- [8] Hintikka, J.: “Problems of philosophy. Problem #3: one-world assumption and Frege’s sense-reference distinction”, *Synthese* **112**, 1997, 431-2.
- [9] Hintikka, J. and Sandu, G.: “The Fallacies on the New Theory of Reference”, *Synthese* **104**, 1995, 245-83.
- [10] Hocutt, M. “Is epistemic logic possible?”, *Notre Dame Journal of Formal Logic* **13**, 1972, 433-54.

- [11] Kripke, S. *Naming and Necessity*, Harvard University Press, 1980 (1st edition: 1972).
- [12] Peckhaus, V. “Calculus ratiocinator versus characteristica universalis? The two traditions in logic, revisited”, *History and Philosophy of Logic* **25**, 2004, 3-14.
- [13] W.V.O. Quine. *Ontological Relativity (and Other Essays)*, Columbia University Press, New York, London, 1969.
- [14] Saint Augustine. *De Magistro*, in *Against the Academicians and the Teacher* (trans. Peter King), Hackett Publishing Company, 1995.
- [15] Smith, D. W. “Kantifying”, *Synthese* **54**, 1983, 261-73.
- [16] van Heijenoort, J. “Logic as Calculus and Logic as Language”, *Synthese* **17**, 1967, 324-30.

Fabien SCHANG

National Research University Higher School of Economics

