Chapter 13
Reference and Denotation

Robert van Rooij

Abstract According to Frege, the meaning of an expression is the description that helps language users to determine what its reference is. Natural as the view might seem, it gives rise to the *conceptual problem* that it presupposes that we already know the meaning of the terms used in the description (Wittgenstein, Quine), and it is *empirically incorrect* because 'having a correct description in mind' is neither a sufficient nor a necessary condition for successful reference (Kripke, Kaplan). Perhaps reference for at least some times is non-descriptive, and depends on context. Anaphora have a referential use as well, picking up the speaker's referent of an earlier used indefinite description. The challenge of this view is to provide a satisfactory analysis of so-called donkey-sentences.

13.1 The Descriptive Theory of Meaning and Its Problems

The perhaps most 'natural' conception of 'meaning', at least in its point of departure, identifies 'meaning' with *naming*. The meaning of an expression is that what the expression *refers to*, or *is about*. What meaning does is to establish a *correspondence* between expressions in a *language* and things in the (model of the) *world*. For simple expressions, this view of meaning is natural and simple. The meaning of a proper name like 'John' or definite description like 'the number of major planets', for instance, is the object or number denoted by it, while the meaning of a simple declarative sentence like 'John came' could then be the *fact* that John came. Beyond this point of departure, things are perhaps less natural. What, for example, should be the things out in the world that common nouns and a negated sentence like 'John didn't come' are about? This referential, or *Millian*, theory of meaning gives rise to a serious empirical difficult as well: the *substitution problem*.

R. van Rooij (✉)
Institute for Logic, Language and Computation (ILLC), University of Amsterdam, Amsterdam , Netherlands
e-mail: R.A.M.vanRooij@uva.nl

S. O. Hansson, V. F. Hendricks (eds.), *Introduction to Formal Philosophy*, Springer Undergraduate Texts in Philosophy, https://doi.org/10.1007/978-3-319-77434-3_13

Assuming, by the principle of compositionality, that the meaning of a complex sentence depends only on the meanings of its part and the way these parts are put together, it follows that if two expressions have the same meaning, one can substitute the one expression for the other in a complex sentence without change of meaning. But because there are 8 major planets in our solar system, on the theory of meaning at hand the expressions '8' and 'the number of major planets' refer to the same thing, and thus have the same meaning. Still, we cannot substitute the expressions 'the number of major planets' for the number 8 in the sentence 'It is necessary that 8 is bigger than 7' without changing its truth value. Frege [5] concluded that the *meaning* of an expression should not be identified with the *reference*, or denotation, of that expression. Instead, the meaning of 'the number of major planets' is the description itself, and the meaning of a noun or name is given by a set of properties associated with the expression that give necessary and sufficient conditions for objects or stuff to be in its denotation. Obviously, the above substitution puzzle does not arise on such a view.

However appealing and natural this cluster theory of reference might be, it gives rise to at least two problems, one conceptual and one empirical in nature. The **first conceptual problem** concerns the predicates, or properties, used in the description that is supposed to identify the referent. What is the meaning of those predicates? The standard theory of meaning doesn't seem to do more than explaining the meaning of one part of the language in terms of other parts – the predicates in terms of which the descriptions are given. One proposal to solve this problem would be that we indeed have a set of 'basic predicates', e.g. the predicates that refer to *natural kinds*. But how does this reference come about, if not via description?

A possible way out of the above regress problem is to propose to study meaning in terms of outward and observable correlates of language behavior. Perhaps motivated by such concerns, [24], for instance, proposed that we should study the meaning of expressions in terms of their use, and the logical positivists proposed their verificationalist' analysis of meaning mainly because they considered the way to verify a sentence as a particularly good way to get clear how a certain sentence is used. The verificationalist' reductive analysis of meaning failed, for one thing because it is difficult – if not impossible – to interpret the terms of a language individually. To determine the meanings of expressions we have to look simultaneously at a whole group (or perhaps all) of expressions of the language *as a whole* [1, 9]. One way to cash out such a *holistic* theory of meaning idea would be to claim that the terms refer to whatever things, properties, and relations that do the best job of making the set of sentences true that speakers in fact consider to be true. Unfortunately, generalizing Quine's (1960) well-known argument for the indeterminacy of reference, Putnam [18] showed that this picture as such is not constrained enough to fix the meaning of the expressions of a language in the intuitively correct way. Even if one knows the truth value of a sentence in every possible circumstance, this doesn't necessarily mean that one knows the intuitively correct meanings of its constituents. For instance, it is possible to formulate highly counterintuitive meanings for expressions like *cat* and *mat*, so that in the actual world they refer to trees and cherries, respectively, without affecting

the meaning of *The cat is on the mat*. To determine the meaning of the terms of our language, knowing the truth value or meaning of a collection of sentences is not enough, because the terms of the language can be assigned weird and 'unintended' interpretations.

The **second problem** for the description theory of reference is **empirical** in nature. Donnellan [2, 3] and Kripke [12] have convincingly argued that this theory leads to counterintuitive results for proper names. They have shown that speakers can refer, and even can *intend* to refer, to particular individuals without being able to describe or identify those individuals. Ordinary people can, for instance, use the name *Feynman* to denote the physicist Feynman even though they have no uniquely identifying set of descriptions in mind. Kripke argued that uniquely fitting some set of descriptions that the speaker associates with a proper name is not a sufficient condition for its successful use either. Kripke [12] and Putnam [17] have similarly argued that the set of properties that speakers or agents associate with *natural kind terms* should also not be equated with the meaning of the noun. This is made very clear by the 'Twin Earth' stories given by Putnam [17] and others. In Putnam's story, the stuff that the inhabitants of the counterfactual situation call *water* is superficially the same as the stuff *we* call *water*, but its chemical structure is not H_2O, but XYZ. If, then, both the earthling and his twin assert 'Water is the best drink for quenching thirst', intuitively they have said something different. But how can this be if they associate exactly the same description with the word and if speaker's description determines reference?

13.2 The Causal Theory of Reference, and Context-Dependence

According to Kripke, Putnam and others, the meaning of at least proper names and natural kind terms is simply what they refer to. But this gives rise to the question of *why* these expressions have the references they in fact have. At this point, Kripke proposed his causal theory of reference. Kripke [12] argues that proper name 'N' can refer to a, only if, and because, a is the entity that is the *source* of the reference-preserving link from the initial baptism of the expression to the speaker's use of the name. This causal 'theory' of reference, or of meaning, is in accordance with a naturalistic philosophy and seems also the natural candidate to limit the possible interpretations of the expressions of 'our' language to solve Putnam's paradox (cf. [14]).

The causal account of meaning is not without problems. For instance, it is unclear how a causal theory could ever determine the meaning of functional words, or of prepositions like '*in*'. Moreover, it is not clear how to cash out the causal account in a completely naturalistic way and there are problems of how to account for our intuitions that we can have false beliefs. One way solve both of these problems involves making use of so-called 'normality conditions'. But in order for the resulting analysis to be wholly naturalistic, we need a naturalistic analysis of such conditions. A natural candidate that suggests itself to provide such an analysis

is Millikan's [15] bio-semantics, but it is controversial whether this theory can do the full job. The main problem of the causal theory, however, is the original substitution problem: if the meanings of 'Hesperus' and 'Phosphorus' are just their referents, 'Hesperus is Phosporus' is predicted to express the necessary true proposition. But, then, how can we account for the fact that agents can seriously doubt that such statements are true?

It is an obvious observation that what is expressed by a sentence is *context-dependent*: in different contexts the same sentence can express different things. What is expressed by 'I am living in Amsterdam' depends on who is the speaker in that context. Kaplan's (1989) theory of context dependence allows us to distinguish different reasons why a sentence is 'necessarily' true. First, what a sentence expresses in context c can be true everywhere. A sentence like 'Hesperus is Phosphorus' is necessary in this way, given that the meaning of a proper name is just its referent. But it might also be the case that a sentence is true in every context in which it is expressed. For instance, an English sentence like 'I am here now' is necessarily true for this reason. But now consider John's uttering of 'I am John'. Though this sentence is necessarily true, the sentence can, intuitively, still be informative. This is because the hearer might be ignorant of the identity of the speaker, or at least doesn't know that he is called 'John'. This intuition can be accounted for in the theory by claiming that the speaker doesn't know in which context he is. One might now propose to use this analysis to account for the other problems as well: people can doubt whether the identity statement 'Hesperus is Phosphorus' is really true, because the referent of a proper name is context dependent, just like the referent of 'I'. And indeed, not only the reference of expressions like 'I' and 'you' depend on contingent features of the context, but this is also true – at least according to the causal theory of reference – for proper names and natural kind terms. Notice, though, that there is a difference between the sense in which the reference of these expressions depends on context. The expression 'I' is context dependent, because *in English*, 'I' always refers to the speaker in direct speech, and the same expression *of English* might be uttered by different speakers. The reference of 'Phosphorus' and 'water', on the other hand, are context dependent only because in different worlds they have a different meaning, or causal origin. But, of course, in that sense the meaning of 'I' is context dependent as well, and depends on the language we speak. Assuming that we speak a particular language, it follows that we sometimes don't know the meanings of the expressions (names and nouns) we use. Though this conclusion is natural to some, to others it feels like a contradiction in terms.

13.3 Indefinites and Anaphora

According to Quine [20], our learning and use of pronouns marks our ability to refer. If we may believe [6], scholastic philosophers held that pronouns can refer back to indefinites because indefinites are referential expressions. The indefinite refers to

that object that the speaker intends to refer to by the use of the indefinite. Moreover, if a speaker uses a referential expression in his utterance, the proposition expressed by this utterance is object-dependent. Geach [6] has criticised this account. If John intends to refer to *d* by his use of the indefinite *an S*, and wants to say of *d* that he is *P*, even though *d* is not, John is not saying something false when he claims *An S is P*, according to Geach, if there actually is an *S* that is *P*. In order not to make such a prediction, according to Geach, it is better to represent an assertion like *An S is P* semantically simply by an existential formula, $\exists x[Sx \wedge Px]$. The specific/unspecific distinction belongs to pragmatics, which should be kept separate from semantics. To handle pronouns, we should follow Quine's insight and treat them as bound variables. A sequence of the form *Some S is P. It is Q* should, according to him, be translated as $\exists x[Sx \wedge Px \wedge Qx]$.

But there are well-known problems with this latter assumption. First, it leads to the unnatural consequence that we can interpret a sentence with an indefinite or other anaphoric initiator only at the end of the whole discourse: incrementality is given up. Second, if we want to interpret the pronouns in a donkey sentence like *If a farmer owns a donkey, he beats it* as bound variables, it seems we have to represent the indefinites in the antecedent as universal quantifiers to get the truth conditions right. But then it seems we have to give up compositionality. We cannot treat indefinites in all contexts in the same way. Finally, sometimes we cannot even get the truth conditions right by assuming that all pronouns should be treated as bound variables. This was shown by Evans [4] by sentences like *Tom owned some sheep and Harry vaccinated them*. According to a Geachian analysis of this sentence, we learn that Harry vaccinated *some* sheep that Tom owned if we accept what is expressed by the sentence; what we seem to learn, though, is that Harry vaccinated *all* of the sheep that Tom owned. Evans proposed that in a sequence of the form *Some S are P. They are Q*, the pronoun *they* goes proxy for the description *(all) the S such that P*.[1] Such pronouns he called *E-type pronouns*.

The above argument does not show that all pronouns are E-type pronouns. The pronouns occurring in sentences like 'Every man loves *his* cat', for instance, seems to function like the bound variables of quantification theory. Indeed, since Evans [4], proponents of the E-type approach normally make a distinction between *bound* and *unbound* pronouns, claiming that such a distinction can be made on purely syntactic grounds; and propose that only unbound pronouns should be treated as E-type pronouns.

However, if we use the term *unbound pronoun* in the above sense, it seems that not even all unbound pronouns go proxy for the definite or universal noun phrase recoverable from the antecedent clause and should be treated as E-type pronouns. Consider for instance *Yesterday, John met some girls. They invited him to their place*. We don't want to say that *they* needs to stand for all the girls John met yesterday. If we want to say that the pronoun is going proxy for a description

[1] Evans [4] claimed that the pronoun *rigidly refers to* (all) the *S* such that *P*. See Neale [16] for a motivation of the interpretation I have chosen.

recoverable from its antecedent, the relevant description should not be definite or universal, but *indefinite*: *some girls that John met yesterday*. To treat the pronoun as an abbreviation of an indefinite description also seems to be needed to get the right reading of a sentence like *Socrates owned a dog, and it bit him.* It seems that this sentence can be true if there was a dog that Socrates owned and it bit him, although at the same time there was also another dog that he owned that did not bite him.

A correct analysis for such discourses was given in Kamp's [10] 'Discourse Represention Theory', Heim's [9] 'File Change Semantics and Groenendijk and Stokhof's [8], 'Dynamic Predicate Logic'. These theories treat anaphoric pronouns simply as bound variables and indefinites as existential quantifiers. However, they interpret existential quantifiers dynamically in such a way that they introduce new objects, or discourse referents, that are available for reference. In this way, they solved Geach's incrementality problem. Moreover, they assure that with negation and conditionals, a universal quantification over assignment functions or sequences of individuals is involved, thereby accounting for donkey sentences and solving Geach's compositionality problem.

Note that, due to their use of *existential* closure, the anaphoric pronoun *he* in a sequence like *A man is walking in the park. He is whistling* is basically treated as an abbreviation of the *indefinite description* 'a man who is walking in the park'. But claiming that the pronoun is an abbreviation of an *indefinite* description would be very implausible. Pronouns are *definite* expressions. To quote Quine [19, p. 113], "'He', 'she', and 'it' are definite singular terms on a par with 'that lion' and 'the lion' " But if a singular pronoun cannot be treated as a definite description that (in extensional contexts) refers to (all) of the object(s) that verify the antecedent sentence, how then can a pronoun be treated as a definite expression? Some *empirical phenomena* also show/suggest that unbound anaphoric pronouns should in general have a more specific interpretation than the standard dynamic theories can offer.

One of those specific phenomena is the case of *pronominal contradiction*, originally due to Strawson [22]. When John asserts *A man is running through the park*, Mary may react by saying *He is not running, but just walking*. It is clear that in such examples the pronoun cannot be used as an abbreviation for the indefinite description *a man is running in the park*. The natural assumption to make here, however, is to say that in this case the pronoun is used referentially, referring back to the *speaker's referent* of the antecedent indefinite, the man the speaker had in mind for his use of the indefinite.[2]

Pronominal contradiction examples are well known to be problematic for the recently developed dynamic semantic theories. It is normally assumed that these

[2]Arguably, a type of pronominal contradiction is also involved in Geach's [7] notorious Hob-Nob sentences: *Hob believes that a witch blighted Bob's mare, and Nob believes that she killed Cob's cow.* Hob-Nob sentences are problematic for standard analyses because (i) witches do not exist, which rules out a *de re* analysis of the antecedent-anaphor relationship, and (ii) the discourse can be true in case Nob doesn't believe that the witch he talks about blighted Bob's mare, which rules out a descriptive E-type analysis.

problematic examples are rather special, though, and that for the 'standard' cases this notion of speaker's reference is irrelevant. The following example, however, suggests that anaphoric pronouns are usually used in this referential way. If John says *A man called me up yesterday*, it would be odd for John to reply to Mary's question *Did he have a gravel voice?* by uttering *That depends, if he called up in the morning he did, if he called up in the afternoon, he did not* if in fact two men called John up yesterday. It is not easy to see how this phenomenon can be explained if it is assumed that pronouns should simply be treated as variables bounded by dynamic existential quantifiers. A natural explanation can be given if it is assumed that for the use of the pronoun the speaker must have a specific object 'in mind'.

Note that according to this account we don't need to make use of existential closure to account for the non-exhaustive interpretation of pronouns, although the *definiteness* of pronouns can still be explained. A pronoun that takes the indefinite in a sentence of the form *Some S is/are P* as antecedent need not be interpreted exhaustively, i.e., need not refer to *all* the (relevant) *S*'s that have property *P*, because it only refers to the (all) speaker's referent(s) of the antecedent indefinite.

Of course, when we account for the anaphor-indefinite relation in terms of the notion of 'speaker's reference', we can no longer give the usual explanation for the asymmetry in acceptability between *John owns a donkey. Mary beats it* versus *John is a donkey owner. *Mary beats it*. Proponents of dynamic semantics, starting with Heim [9], explained the asymmetry solely in terms of the use of an explicit indefinite in the first, and the lack thereof in the second. Though there are problems with this view, a proponent of the alternative picture still has to explain the asymmetry.

According to Kripke [13] and Stalnaker [21], the speaker's reference is relevant to semantics, but only through pronominalisation. That is, it is irrelevant for what is expressed by the sentence (or clause) in which the indefinite occurs, but is truth-conditionally relevant for what is expressed by a later sentence with a pronoun that takes an indefinite as its syntactic antecedent.

How can we account for the *referential* treatment of pronouns on the one hand, and for the *existential* reading of indefinites on the other? One has to assume that possibilities should contain more information than is assumed in standard dynamic semantics, and that the dynamics is (relatively) independent of truth conditions (e.g. van Rooy, [23]). In particular, it should be clear in a possibility what the speaker's referent is of (occurrences of) indefinite expressions. Because the above sketched treatment of pronouns is exactly in line with [13] proposal, it is only to be expected that such an analysis, just like Kripke's, has problems to deal with *donkey sentences* like *If a farmer owns a donkey, he beats it*. The problem is that in such a sentence the indefinites don't seem to be used specifically, while the pronouns can arguably also not be treated as an abbreviation for *definite* descriptions, because it seems that the sentence can also be true in case one farmer owns more than one donkey. One way to solve the problem is to assume that a logical operator like *negation* is treated as an *intensional* operator, in that it allows part of the context, i.e. the choice function, to shift. One has to realize, however, that [11] would call such an 'intensional' treatment of negation a *monster*.

References

1. Davidson, D. (1967). Truth and meaning. *Synthese, 17*, 304–323.
2. Donnellan, K. (1966). Reference and definite descriptions. *Philosophical Review, 75*, 281–304.
3. Donnellan, K. (1978). Speaker reference, descriptions, and anaphora. In P. Cole (Ed.), *Syntax and semantics, volume 9: Pragmatics* (pp. 47–68). New York: Academic.
4. Evans, G. (1979). Pronouns, quantifiers and relative clauses (1). *The Canadian Journal of Philosophy, 7*, 467–536.
5. ***Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für philosophie und philosophische Kritik* (Vol. 50, pp. 25–50). [The classical paper making a distinction between meaning and reference].
6. Geach, P. (1962). *Reference and generality*. Ithaca: Cornell University Press.
7. Geach, P. (1967). Intentional identity. *Journal of Philosophy, 64*, 627–632.
8. Groenendijk, J., & Stokhof, M. (1991). Dynamic predicate logic. *Linguistics and Philosophy, 14*, 39–100.
9. Heim, I. (1982). *The semantics of definite and indefinite noun phrases*. Ph.D. dissertation, University of Massachusetts, Amherst.
10. ***Kamp, H. (1981). A theory of truth and semantic representation. In J. Groenendijk, et al. (Eds.), *Formal methods in the study of language* (pp. 277–322). Amsterdam: Mathematisch Centrum. [The first paper in the 'dynamic' tradition that treats indefinites uniformly as existential quantifiers, also in donkey sentences].
11. ***Kaplan, D. (1989). Demonstratives. In I. Almog, et al. (Eds.), *Themes from Kaplan*. Oxford: Oxford University Press. [The classic paper where the different roles of context of utterance and world of evaluation is forcefully argued for.]
12. ***Kripke, S. (1972/1980). Naming and necessity. In D. Davidson & G. Harman (Eds.), *Semantics of natural language* (pp. 253–355; 763–769). Dordrecht: Reidel. [The crucial paper which contains all the classical arguments against the descriptive theory of reference for proper names, and where the alternative causal theory is sketched.]
13. Kripke, S. (1977). Speakers reference and semantic reference. In: P. French et al. (Eds.), *Studies in the Philosophy of Language, Midwest Studies in Philosophy* (pp. 255–276). University of Minneapolis: Minnesota Press.
14. Lewis, D. (1984). Putnam's paradox. *The Australian Journal of Philosophy, 62*, 221–236.
15. Millican, R. (1984). *Language, thought and other biological categories*. Cambridge, MA: MIT Press.
16. Neale, S. (1990). *Descriptions*. Cambridge, MA: MIT Press.
17. ***Putnam, H. (1975). The meaning of meaning. In K. Gunderson (Ed.), *Language, mind and knowledge*. Minneapolis: University of Minnesota Press. [The classical paper where the twin-earth argument against the descriptive theory of reference of common nouns is introduced.].
18. Putnam (1981). *Reason, Truth, and History*. Cambridge: Cambridge University Press.
19. ***Quine, W. V. O. (1960). *Word and object*. New York/London: Technology Press/Wiley. [Contains the classical formulation of the indeterminacy of reference argument, and defends a holistic theory of meaning.].
20. Quine, W. V. O. (1974). *The roots of reference*. La Salle: Open Court.
21. Stalnaker, R. (1998). On the representation of context. *Journal of Logic, Language and Information, 7*, 3–19.
22. Strawson, P. (1952). *Introduction to logical theory*. London: Methuen.
23. van Rooy, R. (2001). Exhaustivity in dynamic semantics; referential and descriptive pronouns. *Linguistics and Philosophy, 24*, 621–657.
24. Wittgenstein, L. (1953). *Philosophical investigations*. Oxford: Blackwell.