1 Reference: Problems and Promises

What is reference?

As advanced in the Preface, this book is about how the conventional reference of some linguistic expressions or expression uses is fixed or determined (when it is). But what is reference, in the sense in which we want to talk about it here?

As in so many other cases where a question as to what something is is asked, in this case we can list examples that we want, at least at first sight, to count as examples of conventional linguistic reference, and we can give examples that we want, at least at first sight, to count as examples where reference to some particular thing, or reference *tout court*, does not take place. However, giving a general, sufficiently informative but exceptionless characterization seems elusive, and may well be impossible—in fact, this will be one of the ideas made plausible by the discussions in this book. Still, it is possible to state some plausible general ideas about reference in the sense that interests us here, ideas that we will take as guiding in the remainder of the book.

Like so many other theorists in the tradition of the theory of reference, we want to include as cases of reference some cases where a use of a linguistic expression intuitively stands for a certain thing as an effect at least in part of linguistic conventions concerning the expression, and in those same cases we want to say that there is no reference to things different from that thing; and also, we want to say that there is no reference at all in some cases where intuitively, as an effect at least in part of linguistic conventions, it turns out that there is no thing that a relevant use of an expression stands

for. We want to say, for example, that 'Aristotle', as normally used in an ancient philosophy class, refers, and that it refers to Aristotle, the great ancient Greek philosopher. We want to say that the madman who believes of himself that he is Aristotle does not use 'Aristotle' so that it refers to the madman himself, much as the madman believes that he is Aristotle—and much as *he* (as opposed to the expression 'Aristotle') refers, in the sense of 'refers' in which a person refers, to himself with 'Aristotle'. We want to say that if I utter *That is a beautiful mountain*, while intentionally pointing and referring (in the personal sense of 'referring') to a mountain in plain view, with no other mountain in sight, my use of 'that' refers to the mountain in question. We want to say that the agonizing man who in his delirium hallucinates a demon, and utters *He is coming to get me*, does not use 'he' so that it refers to anyone. In all these cases, the reference, when there is such, is the object that, in an intuitive sense, the relevant expression use stands for, as a matter at least in part of what seem to be the linguistic conventions applying to that expression use. And when there is intuitively no reference, or reference is not to some particular thing, it would seem that this is because somehow those same conventions imply (together with the surrounding circumstances) these facts about the relation of standing for failing to hold between uses of linguistic expressions and objects.

But beyond examples of this sort, involving uses of names and demonstratives where what we want to say about their referents or lack thereof is more or less intuitively clear, there may not be a general, informative but exceptionless characterization that we want to accept, one determining for an arbitrary expression use whether it has a linguistic reference or not, and what that reference is to be if there is to be one. We will see in chapters 2 and 3 that in a fair number of actual and potential cases of uses of demonstratives and names, it is unclear whether there is some thing that

a relevant expression use conventionally stands for, and thus unclear whether reference occurs or not. And when we move beyond names and demonstratives altogether, and consider other expressions, such as adjectives, common nouns, Arabic numerals, etc., it seems as if we can't even rely on the intuitive idea of the thing that an expression use stands for as a guide to the question of reference. For example, is there any thing that a use of a verb, a use of 'snored', say, intuitively stands for? Well, even if there need not be a general, informative and exceptionless characterization of reference that definitely solves questions such as this, there are probably a few things that can be said in this area, which suggest that the answer may lie in a certain direction.

Frege, to whom the introduction of the modern theoretical notion of reference at stake here can be attributed in all essential respects, postulated that a variety of meaningful expressions of diverse grammatical types are susceptible of having referents. And he postulated, more specifically, that the reference of an expression use (when it has one) is the thing which constitutes the expression use's contribution to the truth condition of the sentence in which it appears, the thing specifically contributed by the expression use on which the truth or falsehood of the sentence depends.[1] This,

---

[1] According to a relatively uncontroversial way of reading Frege, which I am adopting here, it is his view that each well-formed complex expression arises from the application of a symbol referring to a function to symbols referring to arguments of that function. Then, in the case of a complete assertoric sentence, its reference, which is a truth value, is a function of the referents of its parts, obtained by successive applications of functions to the referents of simpler constituents. This is a principle of compositionality for referents. In this sense the truth or falsehood of a sentence depends on the referents of its constituent expressions, which thus contribute their referents to the determination of the sentence's truth value. For example, suppose that

which we will call the Basic Fregean Idea, is an appealing idea providing a general

criterion for reference, one consistent with the desired connections between the

theoretical notion of reference and the intuitive notion of standing for an object in the

case of names and demonstratives, and an idea that helps in the search for referents for

expressions which are not demonstratives or names. If one says *Aristotle snored* in the

ancient philosophy class, presumably the thing specifically contributed by 'Aristotle',

on which the truth or falsehood of *Aristotle snored* depends, is Aristotle. And Frege

supposed, not implausibly at first sight, that the contribution of 'snored' to the truth

condition of *Aristotle snored* was the function s assigning the truth-value truth to a thing

that snored and the truth-value falsehood to a thing that didn't snore. One can

reasonably think that the thing contributed by 'snored' on which the truth or falsehood

of *Aristotle snored* depends is something that encodes the information whether Aristotle

snored or not, and the function s surely does that—if Aristotle is assigned the truth-

value truth by s, that means he snored; if he is assigned the truth-value falsehood, that

means he didn't. But furthermore, it is reasonable to think that the thing contributed by

---

Aristotle is the referent of 'Aristotle' and that the referent of 'snored' is the function s

assigning the truth value truth to a thing that snored and the truth value falsehood to a

thing that didn't snore; and suppose that the sentence *Aristotle snored* arises from the

application of 'snored' to its argument 'Aristotle'. Then the truth value truth is

obtained as the reference of the sentence *Aristotle snored* by application of s to

Aristotle. Like many others, I take the attribution of a principle of compositionality

for referents to Frege as well-grounded, e.g. given his acceptance of the principle that

substitution of co-referential expressions doesn't alter a sentence's referent (in Frege

(1892)); see e.g. Pelletier (2001). But there are of course dissenting voices (see e.g.

Janssen (2001)).

'snored' in the case of *Plato snored* ought to be the same thing, and ought to encode the information whether Plato snored or not (and *mutatis mutandis* for similar predications of 'snored' of other things). And again the function *s* surely does that. Frege thus postulated that the reference of 'snored' in *Aristotle snored* is the function *s*, or equivalently the set of things that snored, also known as the *extension* of 'snored'.

But the idea and its associated criterion do not solve all the questions about the reference of 'snored' that one could think of. It would seem that the Basic Fregean Idea implies that 'snored' does not refer to the set of things that snored when used in an utterance of *Aristotle might not have snored*, say. For the set of things that (actually) snored would not seem to encode enough information to determine, together with whatever reasonable referents we might assign to the other expressions in the sentence, whether Aristotle might not have snored, i.e. whether he would not have snored in other possible circumstances. It would seem that if 'snored' is to have a reference here, it should be a richer thing than the set of things that actually snored; perhaps it should be the set of all pairs formed by a possible circumstance and the set of things that snored in that circumstance (a thing also known as the *intension* of 'snored'). And this thing does also encode the information whether Aristotle (or any other actually existing thing, such as Plato) actually snored or not. So was this more complex thing the reference of 'snored' in *Aristotle snored*, after all? And then we may also ask, was the intension of 'Aristotle' (the set of all pairs formed by a possible circumstance and the thing that is Aristotle in that circumstance) the reference of 'Aristotle' in *Aristotle snored*, after all?

Frege himself pointed out that it seems as if the reference of 'Hesperus' cannot be the same as the reference of 'Phosphorus' in *The Greeks in Homer's times believed that Hesperus was not Phosphorus*, despite the fact that they would appear to stand for the same thing, the planet Venus, which would seem to be the thing they contribute on

5

which the truth or falsehood of simple sentences containing them depends. For *The Greeks in Homer's times believed that Hesperus was not Hesperus* seems clearly false, but if 'Hesperus' and 'Phosphorus' contribute the same thing to the determination of truth or falsehood in *The Greeks in Homer's times believed that Hesperus was not Phosphorus*, then this sentence should have the same truth value as *The Greeks in Homer's times believed that Hesperus was not Hesperus*, where the two uses of 'Hesperus' certainly must contribute the same thing. And yet *The Greeks in Homer's times believed that Hesperus was not Phosphorus* seems intuitively true. So, what is the reference of 'Hesperus', or 'Phosphorus'? Frege, as we know, postulated that their reference in simple sentences is the planet Venus, but also that their referents when embedded in verbs of psychological attitude are other things, appropriate *senses* containing modes of presentation of the planet Venus, each including enough information as to determine whether the subject of the attitude holds the attitude toward Venus under that mode of presentation. The same goes for 'Aristotle' when embedded under verbs of psychological attitude, and a sense for Aristotle does encode the information that it is a sense *for* Aristotle—it *determines* Aristotle, in fact, on Frege's theory. So why shouldn't a sense, this less familiar thing conjured up by Frege, have been the reference of 'Aristotle' in *Aristotle snored*, after all?

Senses are very controversial things. Many philosophers, including perhaps most of the philosophers in the tradition of the "new theory of reference", probably the most influential in the recent literature on reference, don't give senses much of a role in the study of language, and think that one can do semantics, and explain the phenomena that Frege used senses to explain, using just ordinary individual objects, properties, intensions, and the like—some even doubt senses exist as such. This book shares many basic views and assumptions with the tradition of the "new theory of reference", so

perhaps we can jump onto the wave here also and just forget about senses. But even then, indeterminacies will remain. For example, why should intensions not be referents, given that they seem to be contributed by predicative expressions and (at least in many cases) they seem to encode the information needed to determine the truth or falsehood of the sentences in which those expressions occur? (Just about everybody seems to think intensions are respectable things.)

It is natural to add one guiding idea about reference to the Basic Fregean Idea, one that appears to get rid of the intensions of predicative expressions as the referents of those same expressions. The reference of an expression (when it exists) must presumably be, at least in typical cases, some thing of which it can reasonably be said that some ordinary speaker has at some point (often in the early history of humankind) intended to refer (in the personal sense of 'refer') to it with the expression (or with some ancestor or relative of the expression). Linguistic referents are, we are supposing, determined in part by linguistic conventions. But these are presumably constituted, at least in some typical cases and in part, by the existence of widespread agreements regarding how speakers' referential intentions give rise to referents for expressions or expression uses (agreements which in many cases date in some form or other from the early history of humankind).[2] And if this is so, it is unreasonable to suppose that the referent of an expression is its own intension, for it is unreasonable to suppose that

---

[2] For our main purpose in this book, which is to explain how reference is achieved for certain classes of expressions, we will essentially focus on this aspect of the nature of referential conventions. We will not need to say much about other aspects presumably essential to conventions in general, such as the fact that they are sustained via coordination mechanisms of certain kinds and in virtue of their successful role in the attainment of common aims.

(typically) an ordinary speaker has intended to refer to suitable intensions when she used expressions such as 'Aristotle' or 'snored' (let alone that someone in the early history of humankind intended to refer to the appropriate intensions with the appropriate ancestors of these words). (Though, of course, one is free to give a name to an intension if one knows what an intension is and wishes to give a name to one.) For one thing, intensions are complicated, highly theoretical things, and ordinary people just don't seem to be able to have intentions about them as such, certainly not explicitly, but presumably not implicitly either.

If this plausible idea is accepted, names and verbs do not refer to their own intensions (at least in run-of-the-mill cases). But intensions must somehow be encoded by the real referents, if we want to respect the Basic Fregean Idea connecting reference and truth, and to respect as well the intuition that the truth or falsehood of *Aristotle might not have snored* depends on (something that encodes) the intensions of 'snored' and 'Aristotle'. Fortunately, at least for typical names and verbs, and for the other expressions we will deal with in this book (demonstratives and Arabic numerals in the case of apparent singular terms; and ordinary natural kind nouns and adjectives for sensible qualities, when we turn to apparent predicative expressions), it is reasonable to suppose that our plausible idea and the Basic Fregean Idea connecting reference and truth (as well as the initial desideratum that the reference of a demonstrative or a name should be the thing it intuitively stands for) are jointly satisfied by appropriate things.

On the one hand, it is reasonable to suppose that there is a widespread, if implicit, intention to refer to Aristotle with 'Aristotle' and to something such as the *property* of snoring (or having snored) with 'snored', an intention that may have developed in part via the formation and transmission of earlier intentions to refer to those things. It is just very natural to say that ordinary people have at some point intended to refer to Aristotle

and snoring with some words, in a way in which it is not natural to say that ordinary people can have intended to refer to intensions. That this is natural is due, at least in part, to the fact that it is clear how Aristotle and snoring can have been the objects of ordinary people's intentions; in an intuitively clear sense, both Aristotle and snoring have been the objects of representations in the minds of ordinary people in a way in which intensions have not. Ordinary people did at some point have a perceptual representation of Aristotle that formed the basis of a linguistic referential intention, later transmitted to people who did not get to see Aristotle; and ordinary people did and do have a perceptual representation of snoring that can form the basis of linguistic referential intentions concerning 'snored'. But ordinary people typically have no mental representation of intensions, whether perceptual, linguistic, or of other kinds.

And on the other hand, it is reasonable to suppose that the intensions of 'Aristotle' and 'snored' are encoded by Aristotle and (the property of) snoring, respectively. First, the intension of 'Aristotle' can be said to be encoded by Aristotle, if we accept the largely uncontroversial view that 'Aristotle' is rigid, which implies that its intension is the set of all pairs formed by a possible circumstance and Aristotle *himself*. (And the same can be said of proper names in general, demonstratives, and Arabic numerals.) And second, the intension of 'snored' can be said to be encoded by the property of snoring, as the intension of 'snored' is the set of all pairs formed by a possible circumstance and the set of things that instantiate the property of having snored in that circumstance. (And the same can be said of verbs in general, and of other predicative expressions.) Accordingly, we will suppose that things in general (Aristotle being a paradigmatic example) are the referents of uses of singular terms (when they have

referents), and that properties are the referents of uses of predicative expressions[3] (when they have referents).

This is not to say that predicative expressions refer to properties in just the same way in which singular terms refer to things in general. Presumably singular terms, in an intuitive sense, *stand for* the things they refer to, but the properties referred to by predicative expressions are not things that in the same intuitive sense these expressions stand for. The modes of reference of singular terms and of predicative expressions are presumably different. There are perhaps difficulties spelling out the difference in an unequivocal way, as Frege concluded in his reflections on the distinction between objects and concepts. But I think we can in any case say, without fear of erring too much, that when referred to by a predicative expression, a property is invoked in its capability or potentiality of applying to things, while when referred to by a singular term, a thing in general (including a property) is invoked simply as a thing about which something can be said, its capability of applying to things (if any) not being invoked.

Why study the fixing of reference?

We have now a general idea of what reference must be, in the sense that will concern us here—even if what we have said does not amount, by any stretch, to a full theory of what reference is. So we now have an idea of the sort of thing whose fixing or

---

[3] In taking properties to be the referents of predicative expressions (when they have referents), we are also following the lead of several authors from the "new theory of reference" (and from other traditions). See e.g. Salmon (1981), (1986); Soames (2002).

determination for linguistic expressions or expression uses we want to study in this book. But, it might legitimately be asked, why should a philosopher, or even a philosophy of language specialist, care about how reference gets determined? As I see things, there are two main reasons.

The first, perhaps less important reason, is that a philosopher, and in particular a philosopher of language, should care about how communication is possible, and reference seems to have an important, if perhaps somewhat negative role in the explanation of communication. Frege emphasized that communication, at least in the case of sentences involving names—but suitably analogous points hold for every other expression—would *not* work if the referent of one same name, as used by different people, were not typically the same.[4] Why? Because if people's knowledge of whatever it is that fixes the reference of different uses of one same name directed them to different things, or just to no thing at all, the possibility of communication would be hampered, as people would then interpret other people as talking of different things, or of no thing at all, with the resulting failures of coordination. In order to explain why this situation does not typically (seem to) arise, Frege thought it was enough to suppose that the different senses different people attached to the same name turned out to contain

---

[4] In talking here of the referent of uses of "one same name", I, like Frege, am not presupposing that the notion of two uses being uses of one same name can be understood independently of the idea that they should have the same referent. Rather, I'm using the idea of uses of one same name in an intuitive sense, deferring the question of when two uses are of one same name to ulterior theorization, theorization which should vindicate the Fregean idea in the main text. The considerations in chapter 3 will actually provide the basis for a suitable theoretical elucidation of the idea of two uses being uses of one same name.

modes of presentation of the same thing, the thing commonly referred to by uses of the name made by different people.

Along with the "new theory of reference", I don't think it's plausible to suppose that it is some descriptive sense attached by a speaker to a name that fixes the name's reference. The descriptive senses that Frege had in mind just could not do the job, as well-known arguments by Kripke, Donnellan and others have shown. And in this book I will argue that other descriptive senses that have more recently been thought to do the job in fact cannot do it either. I think, however, and this is not a point always clearly spelled out or perhaps even conceded by "new theorists of reference", that it is plausible to suppose that the existence of shared reference-fixing conventions plays a role in the explanation of the possibility of communication. This might be thought to lead to descriptivism, but it doesn't, because (as I will argue) the existence of observed reference-fixing conventions does not imply the existence of reference-fixing descriptive senses. In general, it is reasonable to suppose that the ability that speakers have of following the common linguistic conventions that fix reference must play a role in the explanation of linguistic communication. When a speaker uses language in order to attempt to communicate that a certain thing (such as Aristotle) has or does not have a certain property (e.g. the property of having snored), which things and which properties these are is something that her hearers will plausibly often guess at least in part in virtue of their having developed an ability to follow the common conventions governing what things are being talked about by means of the expressions used by the speaker. These conventions, as the "new theory of reference" plausibly argued, and as we will emphasize many times in this book, need not be integrated into the semantic structure of the expressions that do the referring. Nevertheless, an ability to abide by the conventions governing particular expressions must be developed somehow by speakers,

on pain of failures of communication—this being the thesis not always clearly spelled out by "new theorists of reference". In this book we will propose a number of reference-fixing conventions as underlying the use of various kinds of expressions in linguistic exchanges, hoping that it will appear plausible to think that the ability to follow these conventions is manifested in relatively simple linguistic behaviors, and doesn't require the attribution of unlikely knowledge to a normal speaker. By contrast with descriptivism, which the "new theory of reference" and the new considerations in this book show to invoke implausibly demanding reference-fixing descriptive associations on the part of speakers, it will appear plausible to think that competence with the relatively simple conventions postulated by the picture of this book does not exceed in an implausible way the capacities that we can reasonably attribute to a normal speaker.

The second, perhaps more important, but at any rate related reason why a philosopher, regardless of specialty, should care about how reference gets determined, arises again from a Fregean view, the Basic Fregean Idea about the connection between reference and truth. On a natural, much reviled but never vanquished conception of human endeavors and accomplishments, truth is something we pursue and very often achieve, of which we often fail to get as much as we want, but that we constantly get more of. Now truth depends on reference: what truths we come to get depends on what it is that is referred to when we express those truths; and if what we think are truths we have got turn out to contain expressions without reference, they will have turned out not to be truths at all. If the natural conception of human endeavors and accomplishments concerning truth is to hold its ground, it must be accompanied by a congenial view or series of views about how reference is fixed, one that makes it reasonable to believe that reference could in fact be fixed that way, consistently with human capacities and

dispositions; and one that makes it also reasonable to believe that our language is not plagued by widespread failures of reference.

There are many puzzles in discussions of reference fixing that constitute obstacles to a view congenial with the natural conception of human endeavors and accomplishments concerning truth. A first kind of puzzles make it hard to understand how reference could take place consistently with human capacities and dispositions. Some of these puzzles exploit the limitations of our epistemic capacities when it comes to singling out things by means of language and thought; others exploit our dispositions to contradict ourselves, which would often seem to get in the way of our ability to single out things by means of language and thought. A second kind of puzzles don't necessarily exploit our epistemic limitations, but make it hard to understand how, even in some cases where our epistemic capacities are not working particularly badly, and where we are not being utterly inconsistent, the world could be such that our reference-fixing conventions manage to single out appropriate items in it.

When we consider puzzles of these two kinds, we are no longer exclusively concerned with guessing what conventions may govern the assignment of reference to expressions, or with the role that referential intentions play in such conjectured conventions. We must engage in more definitely epistemological and metaphysical philosophical tasks. A tempting thought is then that such tasks go beyond the legitimate area of concern of the specialist in reference fixing or in metasemantics. However, the tempting thought must be resisted, as there is at least a strong reason for the specialist in reference fixing to concern herself with an appropriate resolution of those epistemological and metaphysical puzzles. If the puzzles are not resolved, we will be left with unchallenged reasons to believe that our language, or large and important parts of it, is plagued by widespread failures of reference. And then this will immediately

constitute a challenge to whatever conjectured conventions about reference fixing the specialist may have come up with, for these conventions will not have been shown to be compatible with the natural conception of human endeavors and accomplishments concerning truth. Accordingly, from the perspective adopted in this book, which I take to be the perspective adopted by illustrious classics of the tradition of thought about reference fixing, from Frege to Kripke, the theory of reference fixing, or metasemantics, must concern itself with a relatively broad range of epistemological and metaphysical issues. For only in this way can the specialist in reference fixing be reasonably satisfied that her preferred theory of the mechanisms of reference fixing does not inappropriately disconnect reference from truth.

In this book we will consider a good number of puzzles of the mentioned kinds, as problems for a theory of reference fixing, and we will often provide attempts at solutions that, if correct, will together constitute a series of views about reference fixing that will be congenial with the natural conception of human endeavors and accomplishments concerning truth. Some of these views are briefly previewed in the summary of the book's contents given in the next section.

Roads to reference: a preview

Demonstratives such as 'this', 'that', 'he', 'she', 'it' and 'they' are paradigmatic, and perhaps in many ways the most basic, instruments of linguistic reference. They also turn out to give rise to phenomena closely related to those involving proper names, which are probably the most studied instruments of linguistic reference. But in the case of demonstratives, unlike in the case of proper names, and under the influence especially

of Kaplan's groundbreaking work, a certain kind of descriptivism at the reference-fixing level has not seemed difficult to swallow to originators and fans of the recent advances of the "new theory of reference". Just about everybody writing in this area seems to imply, in some more or less explicit way, that there should be a description, provided by the reference-fixing rule for, say, 'that' (by the Kaplanian *character* of 'that') which fixes the reference of a use of 'that' in a given context, and which is known in some more or less implicit or inchoate way by a competent speaker. But the description in question has turned out to be exceedingly difficult to find, as attested by the review of the literature on reference fixing for demonstratives offered in chapter 2. This literature, as we will see, has correctly identified several difficult problems for demonstrative reference fixing, including a particularly vexing one caused by the frequent existence of *conflicting* referential intentions in the same speaker and occasion of use of a demonstrative. The chapter will propose, in line with similar ideas about proper names to be developed later in chapter 3, that the reason for these difficulties is that the reference-fixing conventions for demonstrative reference do not amount to necessary and sufficient conditions for reference, but only to a list of roughly sufficient conditions for reference and reference failure to take place in selected situations—a fact that is manifested in the existence of a fair number of cases where it is uncertain and presumably conventionally indeterminate whether a use of a demonstrative refers to a particular thing or not. If this is so, the chapter argues, there is little hope that a normal competent speaker could in general and invariably associate (in however an implicit or inchoate way) with a use of a demonstrative an appropriate reference-fixing description that he is aware of; for such a description would essentially amount to a general necessary and sufficient condition for a thing to be the referent of that use, a condition that would yield inappropriately determinate and not really known verdicts of reference

failure in many cases for which the real merely sufficient conditions do not yield a determinate verdict. Thus, the demonstrative descriptivism at the reference-fixing level embraced by originators and fans of the recent advances of the "new theory of reference" is just as wrong as the corresponding descriptivism about proper names presumably is. The chapter ends by developing a picture of the reference-fixing conventions that state the conjectured roughly sufficient conditions for demonstrative reference and reference failure, and argues that the picture squares well with all the elements appealed to in the preceding discussion.

Proper names like 'Aristotle' are also paradigmatic instruments of reference, and it can be said that there is much that we know about them, including much about how it must be that their referents get fixed (when some referents get fixed for them). However, one fundamental worry has subsisted, I think, after all the considerable advances in our understanding of names that have been made in the last forty years or so. Some of these fundamental advances, due especially to Kripke, have pointed to the conclusion that the referents of names, at least in typical cases, are not fixed by the ordinary descriptive information that speakers associate with them. The information that a speaker associates with a name N that seems to come out of his mouth with a certain intuitive referent r, notes Kripke among other things, will often be mostly false of r; and if it is mostly true, it will still be, to all appearances, often insufficiently detailed to pick out r. But then how is it that r comes to be the referent of N in our speaker's mouth? Another of the relatively recent advances in our knowledge of names, again due especially to Kripke, is our current understanding that r will often have come to be the referent of N in the mouth of a speaker s as a result of the existence of a chain of transmission of the name, from some speakers who originally started using N (or an ancestor of it) with the intention to refer to r, down to its use by s. This seems right. But

surely, if this happens it is as an effect of linguistic conventions that speakers should have some kind of familiarity with, however implicit and inchoate. So can't it be held (once account is taken also of other conventions for name reference that would appear to be more easily accessible, such as conventions about the possibility of naming stipulations or "baptisms") that s must after all know in some way some semi-occult descriptive information—the information encoded by his familiarity with referential naming conventions—fixing r as the referent of N (as it comes out of his mouth)? Chapter 3 considers this question, made with incisiveness especially by proponents of so-called "causal descriptivism", and answers it negatively. The problem with this kind of idea is again that our name reference-fixing conventions do not seem to amount to a set of necessary and sufficient conditions for name reference; the conventions at work seem to give roughly sufficient conditions for reference or reference failure to take place in particular cases, but no more. Here again cases of presumable referential indeterminacy will play a role in a number of arguments for the suggestion that only roughly sufficient conditions are provided by the relevant linguistic conventions. The upshot is that no "causal-descriptivist" reference-fixing description known by normal competent speakers, not even one known in a merely implicit or inchoate way, can be plausibly built from those roughly sufficient conditions. The chapter considers also the damaging implications of these and other facts for some recent theories of name reference fixing that are not so epistemically demanding as "causal" and other descriptivisms but that propose necessary and sufficient conditions for name reference. Finally, the chapter develops a picture of the reference-fixing conventions stating roughly sufficient conditions for name reference and transmission, arguing that it explains the phenomena appealed to in the chapter.

There are many other instruments of singular reference in natural language aside from names and demonstratives, mentions of which surface in the discussions of the seminal authors of the "new theory of reference", but which are rarely treated in depth by those authors. Among these instruments one finds Arabic numerals, to which chapter 4 is devoted. In this case, as it will turn out, a certain kind of descriptivism will not seem so implausible as a consequence of what will appear as the probable relevant conventions involved in reference fixing. However, the right kind of reference-fixing conventions and descriptions can be brought to light only after a considerable work of clarification of puzzles and debunking of misleading alternatives. There is a common idea that (at least the complex, decimal) Arabic numerals must have a descriptive semantic structure, in this case one given approximately by the polynomials that detail the contribution of each digit in a numeral as the multiple of the corresponding power of ten—in the sum of multiples of powers of ten that the numeral would appear to codify. Even those who adopt the view that the Arabic numerals are semantically non-descriptive typically think of their referents as conventionally fixed by those same polynomials. But, as explained in chapter 4, there is just too much evidence against these and other ideas which seek to explain reference fixing in this area by appeal to sophisticated mathematical descriptions. The chapter proposes an alternative view, on which the reference of (at least the complex) Arabic numerals is fixed ("merely fixed") by means of simple descriptions which single out the referents of later numerals in terms of the referents of earlier numerals; on the proposal, a numeral (or at any rate at least a complex one) will typically get its referent as the result of the operation of adding one as applied to the number denoted by the preceding numeral in the natural series of Arabic numerals (which is generated by principles independent of semantic interpretation). The possibility of reference for numerals, however, faces a basic

problem: if numerals refer to anything, they ought to refer to numbers, and numbers have a bad press among philosophers. Some metaphysical or epistemological principles are often thought to imply that things with the properties of the numbers just could not exist, which leads directly to the view that arithmetical truths apparently containing reference to numbers via numerals, regardless of their simplicity, just cannot be literally true. What could the numbers be, if they are to be fixed as the referents of the Arabic numerals (when used as nouns) by the reference-fixing mechanism postulated in chapter 4? The chapter argues that some of the features of this reference-fixing mechanism actually point toward the view that the referents of the numerals, hence the numbers, are the finite plural cardinality properties.

Chapter 5 concerns ordinary nouns for natural kinds, substances and phenomena, which were shown by Kripke and Putnam to share many characteristics with proper names, including characteristics relevant to the fixing of their referents. In this area, a major challenge for the natural conception of human endeavors and accomplishments concerning truth and its Fregean roots comes from many recent arguments in the philosophy of biology and the philosophy of chemistry. These arguments cast serious doubts on the ability of the presumable reference-fixing conventions for ordinary language nouns for natural kinds, substances and phenomena to fix or determine real kinds, substances and phenomena that those terms could refer to. As in the case of names, in the case of nouns for natural kinds there is much we had wrong or that we just didn't know about forty years ago, but many fundamental advances have been made in the intervening years, due especially to Putnam and again to Kripke. However, the mentioned arguments from the philosophies of chemistry and biology suggest that many aspects of the Kripke-Putnam picture cannot be right, and in particular make it hard to believe that ordinary nouns for natural kinds, substances and phenomena could refer to

the kinds, substances and phenomena that Kripke and Putnam thought were determined by the reference-fixing conventions they postulated for them. A tempting position, in view of the arguments, is again an eliminativist position, on which ordinary language discourse about natural kinds, substances and phenomena is no more than a picturesque and pleasing fiction, but a fiction nonetheless, false or truth-valueless—the sort of view that drags along with it an undesirable conflict with the natural conception of human endeavors and accomplishments concerning truth. The issues here are complicated and require very extensive discussion, but chapter 5 sketches what I think is a sensible non-eliminativist view about the reference-fixing potentialities of the presumable reference-fixing conventions for ordinary terms for natural kinds, substances and phenomena. These reference-fixing conventions, which again amount to lists of roughly sufficient conditions for reference, are not too different from the conventions somewhat imprecisely envisioned by Kripke and Putnam. But careful attention to their accurate formulation, and to a number of details simply overlooked by these seminal authors, leads to the mentioned non-eliminativist view, on which the referents for typical terms for natural kinds, substances and phenomena turn out to be interestingly different from the referents assumed, somewhat uncritically, by Kripke and Putnam.

The final chapter 6 deals with a more localized, but nevertheless important challenge to the possibility of reference in large areas of discourse. The challenge arises with special force for adjectives for color properties, but if correct it must quickly generalize to all terms for what most philosophers have traditionally considered as secondary qualities, such as sound, heat, etc. There is much evidence indicating that different apparently normal people perceive the same colored thing via qualitatively different experiences. This leads them to make what appear to be incompatible color judgments, judgments which, given that the different people in question seem perfectly

normal, would nevertheless appear to be equally faultless. These facts have provoked a tripartite set of implausible reactions. On a first kind of response, the judgments of some of the apparently normal people are right, and the other judgments are wrong, which is compatible with the reference of a color term being a certain objective property or phenomenon; but this brings along an epistemological mystery about how precisely that property or phenomenon has come to be referred to conventionally by all people. A second kind of response is provided by error-theoretic or eliminativist views on which color language is just not a suitable vehicle of reference and truth. And a third response is provided by less radical but somewhat *ad hoc* and linguistically improbable views on which color judgments (and judgments about sounds, heat, etc.) involve reference to subjective, secondary qualities. Chapter 6 develops an alternative picture on which color terms often refer to slightly different objective properties as they are used by different people, a picture perfectly plausible in view of the lessons of context-sensitive reference, and one which avoids both mystery views and eliminativism.

On reference and meaning

By omission, this brief preview gives an idea of the many topics that are not covered in the book, including topics that fall to a greater or lesser extent under the scope of our self-declared subject, the question of how the conventional reference of linguistic expressions is fixed. Thus, for example, many kinds of expressions, including many expressions of traditional philosophical interest, are simply not covered—including descriptions, logical connectives, quotations, modal and other adverbs, gradable

adjectives in general, 'that'-clauses…, to name but a few.[5] Some of these may not seem to present substantive problems from the point of view of reference fixing, but others do. There is not much justification for this limitation in the scope of the book aside from the obvious one, that not everything can be covered in a single mid-sized work. But the book covers at least most kinds of expressions of central interest to classic theorists of reference, and some feeling of relative completeness can perhaps be gathered from this fact.

Mention must be made of one significant range of issues falling under our broad theme that are nevertheless not covered in the book in a substantive degree of depth. We agree with Kaplan's distinction, recalled in chapters 2 and 3 below, between two senses of 'meaning' relevant in discussions in the philosophy of language. In one sense, 'meaning' has a fairly broad application, encompassing many conventional aspects of linguistic expressions; in particular, 'meaning' in this sense includes linguistic conventions about reference fixing, so long as these are somehow accessible to or followed by competent speakers (even if they are accessible only in some implicit or inchoate way). This book is an extended essay on these conventions and thus on meaning in this sense. But there is a second sense of 'meaning', probably the most common in discussions in the philosophy of language, and certainly the most common sense of 'meaning' as used in this book, in which 'meaning' applies to what is often also called 'content'—the information semantically encoded by an expression or expression use. The book does not deal in a substantive degree of depth with many questions pertaining to meaning in this focused sense.

---

[5] For the author's views on reference fixing for descriptions, quotations, and gradable adjectives in general, the reader is referred to Gómez-Torrente (2015b), Gómez-Torrente (2013), and Gómez-Torrente (2010) and (forthcoming), respectively.

In particular, as noted in the Preface, while the book offers substantive arguments in favor of particular views of the semantic nature of certain expressions (this is especially true for color adjectives and for Arabic numerals—the view of the latter as semantically non-descriptive singular terms, favored here, has hardly any substantive defenses in the literature), in other cases views about the semantic nature of expressions are basically assumed without substantive argument. For example, in the treatment of demonstratives and names in chapters 2 and 3, we will essentially just assume that paradigmatic (uses of) demonstratives and names are semantically non-descriptive singular terms which have as referents (when they do have referents) the objects that the demonstratives and names intuitively stand for, without attempting to argue against all possible kinds of views that go against these assumptions. In particular, I will not attempt to argue against all kinds of views on which (uses of) names and demonstratives have some semantic descriptive structure or refer to properties of certain kinds in the way characteristic of predicative expressions. For example, as mentioned in chapter 3, some philosophers and linguists believe that 'Aristotle' is semantically equivalent with some description broadly similar to 'the thing here relevant actually called "Aristotle"'; and as mentioned in chapter 2, some philosophers and linguists think that a use of (the bare) 'that' is semantically equivalent with a use of 'that thing', where 'that thing' is understood as a quantificational phrase roughly equivalent to a description. Many of these philosophers and linguists think also that the referents of such descriptions are suitable properties, intensions, or the like, which are referred to in the way characteristic of predicative expressions. I will not attempt to argue against such views, on which the descriptions or quantificational phrases in question are rather weak and trivial, and cannot by themselves substantively determine the object that is conventionally being talked about (even if this is postulated not to be

the semantic referent). In these cases, I rest content with arguing, as advanced above, that (uses of) names and demonstratives are not even known by normal competent speakers to be co-referential (let alone semantically equivalent), with strong, truly reference-fixing descriptions of any kind. As noted in chapters 2 and 3, even if the supposition that the objects that demonstratives and names intuitively stand for are their semantic referents turned out to be wrong (which I seriously doubt), the theories of reference fixing of those chapters would presumably have an appropriate version that could supplement in interesting ways the hypothetically right theory of the (weak or trivial) predicative semantic structure of names and demonstratives. Analogous claims hold for other expressions dealt with in the book, such as typical ordinary nouns for natural kinds.

The book also doesn't attempt to contribute to the discussion of the question of whether the reference of an expression (in our Fregean sense of the thing that the expression contributes to truth conditions) exhausts meaning in the focused sense. As made clear in the preview above, the book does heavily incline toward the view that several kinds of expressions have their meanings exhausted by their referents, and several arguments in it imply that the meanings of a number of expressions cannot be constituted, even in part, by certain descriptive contents. But the views in the book are not strictly incompatible with the possibility that there may be, even in the cases where the book rules out certain descriptive contents as parts of meaning, other aspects of meaning that go beyond reference in the sense of truth-conditional contribution.

On this general issue, let me just say that I am reasonably certain that the reference of an expression must in general be at the very least the key part of its meaning. It has often been pointed out that reference must play a key role in meaning if meaning (in any reasonable sense) is to have at least many of the characteristics that it

is often supposed to have, including its intersubjectivity and independence from idiosyncrasies of individual speakers' beliefs, its relative stability through time, and its potential materialization for all kinds of expressions.[6] However, not even the relatively weak thesis that reference must be at least the key constituent of meaning will be defended by means of any general argument in the pages that follow. Even so, the series of proposals we will soon be making amounts to an extended case that the achievement of reference is a crucial driving force in the establishment of linguistic conventions, and in the constitution of the role of language in speakers' lives.

---

[6] Once again, these are all characteristically Fregean themes. We must not forget that Frege after all identified meaning (*Bedeutung*) with what later would be standardly called 'reference' in philosophical English, even if he evidently thought that the specifically semantic contents of expressions went beyond their *Bedeutungen* (not just because he viewed certain descriptive contents as in some sense part of meaning, but also because of the existence of other conventional connotations of at least many expressions). A succinct but magisterial development of these Fregean themes concerning the relationship between reference and meaning can be found in Higginbotham (2006).