# 1 The Multidisciplinary Field of Referring

> One of the oldest problems in the philosophy of language, indeed, one of the oldest problems in Western philosophy: the nature of linguistic reference.
>
> (Searle, 1990: xiii)

## 1.1 Introduction

The motivation for this volume began several years ago when we started to notice certain types of referring expressions which did not, in our view, fit comfortably with the established literature on reference. We started discussing these outlier types of expressions in terms of atypical reference. As will be clear later in this volume, we have continued to use the term ‘atypical’ despite it being not entirely satisfactory, for reasons we explain in Chapter 4. For now, we can simply think of typical reference as conventional and atypical reference as divergent uses. Since the field of study related to referring and reference is broad – spanning multiple disciplines, including philosophy, psychology, linguistics, and computer science – understanding it is a multidisciplinary task. Some disciplines combine more readily than others for a variety of reasons. We also find that even within one discipline, there are many different approaches and perspectives. This kind of variation is certainly true of linguistics. While this volume is anchored within linguistics, we build on multiple perspectives to provide an integrated approach to referring in language with a focus on atypical expressions.

At this point you might be wondering what exactly an atypical example of referring is. We want to differentiate these expressions from ones that are errors, or that are related to language development (see Chapter 9). The expressions that caught our attention often appeared to be unusual uses of definite and indefinite articles, or ones that seemed to be playing with the conventions of reference. There is, for example, in the general literature on referring, an assumption that most uses of the definite article, *the*, signal a known discourse referent. In fact, in the previous sentence, we used *the* in the noun phrase (NP)[1] *the definite article* to signal that there is in fact only one such article in the English language and most people interested in language know this. The definite article is called *definite* because of its typical use in specifying an identifiable, or definite, referent. We will return to the concepts of definiteness and identification in various places in this chapter and later chapters (specifically in Chapter 2). For now, and with the idea of the definite article in mind, let us consider examples (1) and (2).[2] The uses of *the* in these two examples seem atypical in the sense that they do not represent the most common or most frequent uses of *the*; in other words, the highlighted expressions are doing more than what we typically think of as the job of the definite article.

(1) That's not to say I'm not feeling some inflation; as a consumer I've begun really noticing that my money isn't going as far as it did, say, about five years ago when **the wife** and I relocated to California.

(2) I'm tired of doing things that I know I SHOULDN'T do and I don't WANT to do but can't seem to have control over. I hate being **the wife** that freaks out in the grocery store or **the mom** who can't play with her kids because I can't pull myself away from the anxiety of everything.

In example (1), *the* could easily be replaced by *my* as in *my wife*. Some of us might not appreciate the use of *the* in this expression because it has an objectifying connotation. There is something about using this expression, *the wife*, to refer to a specific woman that is atypical, or certainly not the norm for referring to people, other than by their position (e.g. *the teacher* or *the lawyer*). The juxtaposition of *the* with a common noun implies a kind of relational position. The use of *the* in example (2), while also an atypical expression, is different from its use in example (1). In (2), the speaker is referring to herself, and the implication is that there are types of wife (*the kind of wife that . . .*), which construes a qualitative description similar to *I am smart* or *I am sensitive*, for example.[3] However, here, it is very clearly a self-assessment of behaviour that the speaker finds undesirable; she is referring to herself, as signalled by the use of *I* in a chain of self-reference (*the wife, the mom, I, myself*). It is as though the speaker steps out of definite reference to herself as

[1] The linguistic term noun phrase or NP will be explained in detail in Chapter 3.
[2] We would like to point out that we are only commenting here on the use of *the*. Referring to close personal relationships using expressions such as *the wife*, *the daughter*, etc. is common in many dialects of English.
[3] Unlike example (1), the phoricity of *the* here is referred to as esphoric reference by Martin (1992: 123), although in this case the expression is not being used to introduce a participant to the ongoing discourse, which is the common use of esphora (see also Du Bois, 1980 for a discussion of esphora).

the speaker to identify a type of referent. Thus, these uses of *the* stand out and, because such uses are typically infrequent, they carry different, or additional, referential meanings.

It is not only definite NPs that are used in atypical ways. The literature on referring is also clear that indefinite expressions such as *someone* are inherently non-specific since there is no discourse referent that can be identified. If you return home and see your front door is open and there appears to have been a break-in, you are likely to say *someone broke in*. The use of *someone* is useful for such occasions because you may know that one or more humans are involved, but you may have no idea specifically which individual or individuals this includes. In examples (3) and (4), we present instances of the use of *someone* where this is not the case.

(3) **Someone** didn't put the garbage in the bin and now there's garbage all over the deck.

(4) I'm guessing **someone** took the car and **someone** took the truck.

Example (3) is one that some parents may recognise as a useful strategy to let a child know that they have made a mistake. These examples illustrate, then, that there are contexts in which a speaker will use an indefinite or general expression despite knowing the specific identity of the referent. In example (4), the speaker is referring to her son and daughter-in-law who had been visiting one afternoon but left, each in one of their jointly owned vehicles. The speaker, not knowing which one drove off in the car and which one left in the truck, opts for *someone*. These examples illustrate how a speaker can have a specific referent in mind and still choose an indefinite expression. The question raised is whether these instances are referring expressions or not. We will come back to this point later in the chapter and in more detail in Chapter 2.

These four examples provide a glimpse into the kinds of expressions that have motivated this volume. In the remainder of this chapter, we will provide an overview of the approaches that have informed the integrated perspective we present in the volume. Towards the end of this chapter, we will set out the aims of the volume and briefly describe the structure of the book.

## 1.2 The Foundations of Reference

The study of reference is a rich field with significant and historical foundations in philosophy which have spread to computational approaches and psycholinguistics as well as general linguistics. In this section, we very briefly review some of the key contributions from philosophy, computational linguistics, and psycholinguistics to give us the background of influences that are still relevant to our understanding of reference and referring.

### *1.2.1 Philosophy of Language*

The earliest studies of referring expressions and reference stem from philosophy. It was then developed outwards and into linguistics via two principal paths: semantics (e.g. Lyons, 1977) and computational linguistics (e.g. Appelt, 1985). Searle's work on reference has led indirectly from philosophy to studies of discourse and pragmatics (e.g. Searle, 1969). The long-lasting influence of philosophy of language on linguistics is not surprising since, in comparison, linguistics is a young science. It is worth noting briefly, as we do here, some of the key features that we have inherited from philosophers.

For philosophy, the central notion has been that of unique reference – the use of an expression to uniquely identify some object or person. Referring expressions generally include 'proper names, definite descriptions ("the tallest man in the world"); demonstrative terms ("this man") and some pronouns' (Evans, 1982: 1). In this view, our examples (2) to (4) would not be referring expressions since they cannot be said to uniquely identify an object or person. Within the field of philosophy, however, there has been some dispute as to what constitutes a referring expression. Part of the problem lies in interpretations of the (non-)existence of the referent and/or truth value of the expression in question; for example, whether an expression referring to something that does not exist can be said to have reference.

The first real treatment of reference is attributed to Frege (1892/1993), although according to Searle (1969: 77), it can be traced back as far as Plato. For Frege, a referring expression is an expression that refers to a unique object, in other words, a definite description. In his view, a referring expression is a type of sign, and by sign Frege (1892/1993: 24) means 'any designation representing a proper name, which thus has as its reference a definite object (this word taken in the widest range)'. One of Frege's greatest contributions, or at least his longest-lasting contribution, is his idea that any given referring expression will have two different types of meaning: sense and reference. Frege (1892/1993) demonstrated the need for this distinction most famously with his now classic *morning star*/*evening star* instance, given here in example (5).

(5) The morning star is the evening star

Frege's claim was that while each nominal expression in (5) has a different sense, they both share the same reference (i.e. both NPs are used to refer to Venus). Reference, for Frege, is the actual object being designated, that is, 'an object perceivable by the senses . . . an internal image' (1892/1993: 26). Sense is part of the meaning of the sign but separate from reference. He also claimed that an expression may have a sense but no reference, as in example (6) from Frege (1892/1993: 25), where there is no unique object to be designated. This expression is not a proper name (definite referring expression) in Frege's terminology. While the distinction between sense and reference may be welcome to some linguists, most of us would still want to consider examples such as (6), and our invented example (7), as referring expressions despite the absence of a uniquely identifiable referent.

(6) the least rapidly convergent series

(7) the largest number in the world

Frege's work led to considerable debate and various developments in relation to reference, but we will jump to Strawson (1950) because his contributions are important to linguistic views on reference. He argued that existence should not be a criterion for defining a referring expression, which led to his theory of presupposition, a standard notion in the field of pragmatics. This work may well have formed the first functional approach to the study of referring expressions. In Strawson's view, a referring expression is 'an expression with a uniquely referring use' (ibid.: 324); by 'uniquely referring use', Strawson intends the use of expressions 'to mention or refer to some individual person or single object or particular event or place or process, in the course of doing what we should normally describe as making a statement about that person, object, place, event, or process' (ibid.: 320). These types of expressions include proper names, singular demonstrative pronouns, singular personal and impersonal pronouns, and 'phrases beginning with the definite article followed by a noun, qualified or unqualified in the singular' (ibid.). However, he insists on the importance of the use of the expression. He makes a three-way distinction among an expression, the use of an expression, and the utterance of an expression (ibid.: 325–326). Both the context of the utterance of an expression and its function are very important according to Strawson. The following two examples, (8) and (9), are given by Strawson to make this point.

(8) the king of France is wise

(9) the greatest French soldier

In (8), Strawson explains (1950: 325) that the context of the utterance *the king of France* is important since if uttered during the reign of Louis XIV, some would say the sentence is true, but if it were said during the reign of Louis XV, many would say it is false. Similarly, the expression *the greatest French soldier*, in example (9), may or may not be uniquely referring, depending on the function of its use in a given context. If uttered in the sentence *Napoleon was the greatest French soldier*, then the expression is not referring since it is not being used to refer to a particular individual; instead, it is being used to say something about a particular individual (Napoleon). However, if the same expression is used in the sentence *the greatest French soldier died in exile*, then it is being used to refer to an individual (Napoleon). The realisation of the significance of context enabled Strawson to see referring not as something that an expression does, but rather as a resource that can be used by a speaker.

Linguistic approaches to reference have inherited a great deal from philosophical studies of reference. Traces of its influences can be found in most, if not all, linguistic theories, and this is most notably relevant in the use of concepts such as definiteness, sense, and reference. Although much of what we glean from philosophical studies of language is important, some of it carries with it concerns and definitions that are not relevant for linguistic studies, as we will see in our discussion of (in)definiteness in Chapter 2. Nevertheless, any study of referring expressions must understand the roots of its development, and very clearly these roots lie in the earliest philosophical studies of linguistic reference. Comprehensive treatments of reference within the philosophical traditions are presented in Bach (1987) and Sainsbury (2005). For more recent contributions to the study of reference from a philosophical perspective, see work by Clapp et al. (2019) on negative existentials, and also Hanks (2019) on speech acts.

### *1.2.2 Influences from Computational Linguistics*

Although we are not including computational approaches to referring in this volume, it is worth briefly considering contributions that have been made from this field. Referring expressions, as a domain of study, have received the greatest attention from the field of natural language processing, which covers various computational linguistic tasks, such as machine translation, speech recognition, text-proofing, automated summarisation, information retrieval and extraction, question answering, natural language understanding and natural language generation, and, increasingly, robotics. In language generation, many approaches were developed to solve the problem of generating referring expressions, for example, from a pragmatic perspective (e.g. Appelt, 1985; Appelt and Kronfeld, 1987; Dale and Reiter, 1995), from a discourse perspective (e.g. Paris and McKeown, 1987), from a syntactic perspective (e.g. Power, 1999), or from the perspective of the computational architecture (e.g. Horacek, 1997).

Researchers working on referring expressions for the purposes of language generation have generally defined the generation of a referring expression as ‘the task of selecting words or phrases to identify domain entities’ (Reiter and Dale, 1997: 68). Domain entities are an infinite set of entities (things) which belong to a particular domain, including everything concrete, abstract, and imaginable. For any given computer generation system, the set of domain entities is usually constrained to a specific context, for example, weather reports (Reiter et al., 2005), automatic route descriptions (Dale et al., 2005), or recipes (Dale, 1989). As it is widely accepted that a referring expression will be realised as a noun phrase, referring expression generation is, then, a procedure for producing an NP that 'identifies an entity . . . to the hearer in the current discourse context' (Reiter and Dale, 1997: 79).

While it is perhaps more common for computational work to take up concepts from linguistics than the other way around, there are two contributions we feel are worth mentioning briefly here. One stems from the work of Dale and Reiter (1995), which focussed on how to incorporate Gricean maxims (Grice, 1975) while avoiding the potential to create false implicatures. Example (10) from Dale and Reiter (1995: 251) illustrates this. Here, a referring expression is generated that provides an overspecification of the referent, that is, it is not as brief as it could be (cf. *Sit by the table*), where the additional information expresses the communicative goal of warning the addressee not to touch the table as they sit down (because of the fresh paint).

(10) Sit by the newly painted table

A key finding in their work was that, at least in computational terms, adhering strictly to Grice's maxim of 'Brevity' is very costly. Psycholinguistic evidence suggests that speakers do not adhere to this maxim strictly when referring (see Levelt, 1989, for example). The conclusion drawn by Dale and Reiter (1995: 247) concerning the task of generating referring expressions is that 'the Gricean maxims should not be interpreted too literally' (ibid.: 249).

Another, perhaps more significant, view of referring expressions that was gaining importance at the same time was the view of referring as action and plan rather than as description, that is, the speaker is planning an expression, rather than describing an object (Dale and Reiter, 1995) in the computational architecture. A major contribution comes from Kronfeld, who argued that 'a computational model of referring must show how the successful use of a referring expression in a given context is due to the solving of a planning problem – given also a goal, various rationality assumptions, and relevant linguistic institutions' (Kronfeld, 1990: 9). Kronfeld understood that referring was not only about identification and that the speaker had to plan the best approach to achieving the desired effects (however unconsciously). Planning is required, according to Kronfeld, because of the need to consider 'the effects on the hearer's mental state' (ibid.: 76).

In summary, computational models, in the very least, offer contributions that help us separate cognitive, psycholinguistic, and pure linguistic behaviour, even if there is some degree of overlap. Both the domains of the philosophy of language and the computational modelling of language come (although via very different routes) to find that there is a kind of separation, or rather a distinction to be made, between the beliefs that a speaker has and the linguistic choices available. Psycholinguistics is the field that tries to maintain the balance between what is going on in the brain and what is happening in the language system. In what follows, we turn our attention to more experimental evidence from psycholinguistics.

### *1.2.3 Psycholinguistic Perspectives on Referring*

Despite the significant contributions made by philosophers, Clark and Wilkes-Gibbs (1986: 3) argue that philosophy offered a limited account of referring. They claim that the main reason for this lies in what they call a 'literary model' of reference, which holds four tacit assumptions (ibid.): (i) reference is only expressed by certain types of noun phrase – proper noun, definite description, or a (certain) pronoun; (ii) noun phrases identify the referent uniquely against the common ground; (iii) speaker intention is met simply by issuing that noun phrase; and (iv) the referring process is controlled by the speaker alone. When compared against what speakers actually do during conversation, Clark and Wilkes-Gibbs (1986) found that these assumptions failed to account for 'real' instances of referring. In what follows, we will very briefly consider what referring looks like from a psycholinguistic processing perspective.

One of the most important contributions from psycholinguistic research into referring expressions has been in relation to the role of collaboration between speaker and addressee, which helps participants to refer successfully (Clark and Wilkes-Gibbs, 1986; Schober and Clark, 1989; Brown, 1995; Brennan and Clark, 1996). The 'collaborative view' proposed by Schober and Clark (1989: 211–232) claims that participants in a conversation actively work together to ensure that understanding takes place and do not proceed with the conversation until they are satisfied that they have mutually understood each other. We will expand on this issue throughout the volume but in particular in Chapter 5.

For successful reference to take place, there needs to be a level of mutual understanding between the speaker and addressee. Evans (1982: 315) suggests that the interlocutors not only need to be thinking of the same referent, but also thinking of it 'in the right way', proposing that this necessitates 'thoughts that are pretty similar on the part of the speaker and hearer' (ibid.: 316). Brown echoes this (1995: 65) but distinguishes between correct interpretation and adequate interpretation. The former requires that the listener must have the same thought, or thoughts that are 'pretty similar' to the speaker, while the latter takes into consideration the relationship between 'the utterance, the context and the listener's intentions in interpreting the utterance' (ibid). The question here, of course, is to what extent the addressee needs to retrieve the intended referent and think of it in a sufficiently similar way to the speaker for successful understanding to take place. Clark and Wilkes-Gibbs' notion of understanding 'a criterion sufficient for current purposes' (1986: 34) goes some way to explaining how closely aligned interlocutors have to be. This criterion states that for a new contribution (i.e. 'a unit of conversation' (ibid: 35)) to proceed, the old, previous contribution must have been mutually accepted.

Clark and Wilkes-Gibbs' (1986: 33) principle of 'mutual responsibility' between speaker and addressee is also important since referring is a collaborative process requiring action on the part of both speaker and addressee in order to allow for successful reference to take place. They suggest that during the referring process, the speaker and addressee may repeatedly repair, expand on, or replace the referent until they arrive at a mutually acceptable version (see our discussion of rephrasing in Chapter 7). So, before the conversation can move on, both interlocutors must be satisfied that they have reached a sufficient level of mutual understanding. In Chapter 7, we consider how this collaboration works even when speaker and addressee are in a more adversarial than collaborative relation. As long as the speaker and addressee are satisfied that mutual understanding has taken place, then it is sufficient.

In ongoing, interactive discourse, it is fairly easy to stop and check understanding, especially if there is any doubt about the identity of the referent. In written modes, the principle of mutual responsibility may be weakened or modified and becomes instead a principle of 'distant responsibility' (Clark and Wilkes-Gibbs, 1986: 34), which means that writers have to ensure that the reader understands their meaning 'to a criterion sufficient for current purposes' (ibid: 36). It could be that the writer needs to provide other pointers to help the reader: for example, certain co-textual or semantic clues, such as the ongoing sequence of references to the same referent or the fact that the heading or title is the discourse topic of the text. Clark and Wilkes-Gibbs (ibid: 36) add, however, that some writers may 'retain a sprinkling of provisional noun phrases, repairs, expansions and replacements apparently to affect a spontaneous style or for other rhetorical effects'. This suggests that the principles of both mutual and distant responsibility may be somewhat violated in order to have a particular impact on the discourse (cf. Fox, 1987). The identification of a referent is not the only function of referring expressions; the speaker or writer might be 'doing something special' (ibid.: 17). This matter is revisited in Chapter 8, where the function of the expression-type is discussed.

When a speaker is intentionally over- or under-explicit, this overspecification or underspecification could be seen as a breach of the 'communicative contract' (Givón, 1993a: 232) or 'conceptual pact' (Brennan and Clark, 1996: 1491), that is, a temporary agreement between interlocutors about how they are to conceptualise a particular entity, since the speaker could be considered to be acting in an uncooperative manner. For instance, as we saw in example (10), the modifier *newly painted* was over-explicit for referential purposes but still a successful reference. The overspecification does not necessarily mean that these expressions are not collaborative. In the case of a written text, the writer might assume that they have put sufficient cues in place for the reader to be able to successfully identify the referent. Thus, the principle of distant responsibility to the addressee remains, at least to some extent, intact (see Chapter 6).

The role of the addressee is central to referring, not just in terms of a shared conceptualisation of the referent, but because for acts of referring to work, the addressee has to implicitly or explicitly agree to the expression. As explained by Brennan (2000: 4), 'referring expressions are provisional until ratified by addressees', who play a crucial role in the successful resolution of a referring expression. In instances where the identity of the referent could be potentially unclear or the intentions of the speaker are uncertain, it is the addressee's role which is the most decisive. We will look at examples of this in some detail in Chapter 7, but for now we will briefly consider evidence of how speakers include their addressee's knowledge about the referent and surrounding context.

A study which is particularly pertinent to Brennan's above mentioned claim is that of Brown's shared map task (1995), in which she reports that addressees are not particularly concerned with the definiteness status of referring expressions. Brown discovered that listeners, when collaborating in a shared map task, interpreted both definite and indefinite expressions as referring. The listener was able to identify the referent successfully in both cases. In the map task, it did not appear to matter whether an indefinite or definite expression was used to describe a shared feature; the listener took it as intending to identify a particular referent (ibid.: 70–72). Brown found that indefinite expressions are treated as referring to 'the one unique example of a palm beach on their map' (ibid.: 70). Wright, in his analysis of a similar map task, supports this observation, stating that 'the occurrence of definite and indefinite articles is not a reliable indicator as to whether a speaker is treating an item as shared or not' (1990: 71). Further, he (ibid.: 78) concludes that '[i]t does indeed seem that more skilled users of the language are able to tailor the referential expressions they use to the informational conditions in which they occur [. . .] [and] that speakers take account not only of information in the referential domain, but also of their hearer's state of knowledge of that domain'. This comment is noteworthy since Wright (1990) is suggesting that it is not the formal marking of referring expressions which is crucial to the mutual understanding of the interlocutors, but rather their judgement and interpretation of the context, as well as how they evaluate each other's awareness of the context of utterance.

Returning to Brown's map task experiment, it must be noted that she concludes that the speaker is simply indicating that 'the epistemic status of the referent is not (or may not be) well-grounded for both participants' by