The nature of referential intentions

MSc Thesis *(Afstudeerscriptie)*

written by

Rachel Maden
(born 5th of March 1997 in London, United Kingdom)

under the supervision of **Prof Dr Martin Stokhof**, and submitted to the
Examinations Board in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:	Members of the Thesis Committee:
8th of July 2020	Prof Dr Martin Stokhof
	Dr Paul Dekker
	Dr Elsbeth Brouwer

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Abstract

This thesis provides solutions for a number of theoretical problems associated with 'referential intentions', which are frequently supposed to play a part in the semantics of demonstrative expressions (e.g. Perry 2009; Åkerman 2009; Stokke 2010; King 2014; Speaks 2016). Particular issues addressed are the relationship between referential and (Gricean) communicative intentions, the 'problem of conflicting intentions' that arises when a speaker has a number of intentions that may be called 'referential', and the accusation that the very idea of a referential intention is irrevocably circular. I argue that the most feasible notion of a referential intention is characterised by two primary features: it is an intention-in-action rather than a prior intention, and it amounts to an intention to establish joint attention on a particular object.

Contents

Chapter 1

Introduction

Central to the philosophy of language is the phenomenon of *reference*, by which objects in the world are associated with particular linguistic expressions. What is the nature of this association? It seems that in the case of proper names, like 'Joan of Arc' or 'Ariana Grande', they refer by virtue of established conventions matching[1] them to individuals. No matter the exact nature or origins of the conventions, which puzzle philosophers of language to this day, it is clear that a significant part of the communicative utility of proper names comes from the constancy of this name-individual relation. How else are we to know that my utterances of 'Joan of Arc' carry the same meaning as yours do? Without that guarantee, successful communication about particular objects would become rather difficult.

Similar considerations apply to definite descriptions, like 'the smallest country in the world'. In this case, however, the source of their referential power is more transparent; the fact that 'the smallest country in the world' refers to the Vatican City clearly arises from the meaning of 'smallest country in the world' (and in turn its component expressions) in combination with the relevant facts about the world. This rule-governed, compositional nature ensures that any competent speaker of English can figure out that 'the smallest country in the world' refers to the Vatican City given the right information, and be confident that other competent speakers would always come to the same conclusion.

In contrast, there are a significant number of expressions that refer to different objects in different contexts. Moreover, it is precisely this flexibility that makes *them* so useful. They are usually known as 'indexicals'— a term encompassing pronouns like 'I', 'you', 'she', 'it', 'this' and 'that' as well as adverbs like 'now', 'here' and 'today'—and despite their context-sensitivity we have little trouble using them to communicate. In some cases, this is because the variation follows a systematic rule; for example, 'I' always refers to the speaker. In other

---

[1]It is an open question whether proper names refer 'directly' or their reference is mediated by some other kind of meaning, e.g. a cluster of definite descriptions. It is out of a desire to stay neutral on this matter that I use the rather vague term 'matching', since it has no bearing on the topic of this thesis.

cases, the variation seems near-unconstrained; for example, 'this' can refer to just about anything. *This* is a bit of a conundrum for semantic theories, which aim to systematically assign meanings to natural language expressions. Could we come up with a rule for 'this' like the one for 'I'? One theory, known as *intentionalism*, claims that the referent a of use of a problematic term like 'this' is simply whatever the speaker intends it to be, or the object of her *referential intentions*. This thesis investigates the nature of those intentions. Before getting into the details, though, there is more philosophical background to be covered.

Let's return briefly to proper names. Gottlob Frege ([1892] 2009) famously observed that there appear to be two 'components' to their meanings. He was persuaded as such by pairs of sentences like this:

(1) Joan of Arc is Joan of Arc.

(2) Joan of Arc is The Maid of Orléans.

Suppose that the meaning of a proper name is simply its referent: the individual matched to it by linguistic conventions. Then (1) means that the individual known as 'Joan of Arc' is identical to the individual known as 'Joan of Arc'. (2) means that the individual known as 'Joan of Arc' is the individual known as 'The Maid of Orléans'. But given that the individual known as 'Joan of Arc' is the individual known as 'The Maid of Orléans', (1) and (2) have the same meaning. This can't be right—(1) is clearly tautological whilst (2) can be informative. Instead, Frege argued that in addition to their referent, proper names have a "sense" which affects the "cognitive content" of the sentences they appear in. So while 'Joan of Arc' and 'The Maid of Orléans' have the same referent, they differ in sense. Conceived as something like a guide to reference, the sense was in a way privileged for Frege because he thought that complexes of senses make up propositions: the truth-conditional entities associated with sentences. For Frege, then, (1) and (2) express different propositions.

Here I have used proper names as an illustrative example, but Frege actually took the sense/reference to apply to everything from definite descriptions to whole sentences (to name a few). In the case of *indexicals*, however, David Kaplan ([1977]1989a) famously argued that Frege's view could not satisfactorily explain their semantics and developed his own seminal theory as a result. Now, Kaplan accepted that an indexical's "extension", i.e. Frege's "referent", does not exhaust its meaning. But he also thought that applying the notion of sense to indexicals would unjustly conflate two different components of their meaning: what he called "character" and "content". Character was a property of linguistic expressions, while content belonged to *uses* of those expressions in context.

Take the example of 'I', which when used in context always refers to the speaker in that context. We might say that the conventional linguistic meaning of 'I' is this 'recipe' for determining the contribution of its uses, since the recipe is all a speaker needs to know in order to use and understand it. Compare now Kaplan's "characters", formally construed as functions from contexts to contents. Contexts are tuples of basic features that include in Kaplan (1989a)

at least an agent (speaker), a time, a position (location), and a (possible) world. So the character of 'I' is the function that takes a context and returns the agent of that context. But note that this doesn't get us all the way to referent, only to "content", where contents are functions from circumstances of evaluation to referents. Circumstances of evaluation are time-world pairs. However, Kaplan thought that indexicals were *rigid designators*, meaning that their contents are constant functions. This means it is technically incorrect to say that knowledge of an indexical's character is enough to be able to figure out the referent of a use of that indexical, but only slightly. (Of course, the extra step of content is still necessary so that we can incorporate non-rigid designators, like definite descriptions, into the semantics.)

Given these definitions, could it be that character corresponds to Fregean sense? No, because while a referent can be associated with multiple senses, each sense only determines one referent. In contrast, 'I' has the same character no matter who utters it, but can clearly change referent. Nor could it be that Kaplanian content is Fregean sense; 'I' with respect to a context where Joan of Arc is the speaker has the same content as 'Joan of Arc' with respect to the same context (it is the constant function that outputs the individual Joan of Arc). Hence identifying it with sense would eradicate the explanatory power of Frege's theory all together, since we would have no explanation for the difference between these utterances when spoken by Joan of Arc:

(3) "Joan of Arc is Joan of Arc."

(4) "I am Joan of Arc."

The difference, of course, is that (4) is clearly informative in a way (3) is not. So, Kaplan successfully shows that his three components of indexical meaning are irreducible to Frege's two. There is one more key difference between Frege and Kaplan, too: Kaplan viewed propositions (to be found at the content level) as singular in the Russellian sense, i.e. as having objects as constituents. This stands in contrast to Frege's sense-complex view, which views propositions as composed from the senses of individual expressions rather than their referents. Given that this conception of propositions is key to Kaplan's theory of indexicals, and that so much writing on them since is based on his theory, I will also adopt the Russellian view going forward.

1.1 Intentionalism

As of yet, we have seen enough of Kaplan's theory to account for what he called the "pure indexicals". These are expressions like 'I', 'here', 'today', whose contents are fixed automatically by context. The case of "true demonstratives" (hereafter just 'demonstratives'), a category exemplified by the expressions 'this', 'that', 'these' and 'those' but also including certain uses of personal pronouns like 'you', 'she', 'them' and so on, is more difficult because it is not so clear what kind of story can be told about their characters. There are of course

some constraints—for example, 'she' must refer to a female[2]—but otherwise it is difficult to identify reliable rules for determining their contents. In particular, they seem to require some kind of further supplementation outside of Kaplan's narrow notion of context. For Kaplan ([1977]1989a), this supplementation is by an accompanying demonstration, which is paradigmatically but not necessarily a pointing gesture[3]. By the time of Kaplan (1989b), however, he has changed his mind—demonstrations are simply externalisations of speakers' "directing intentions" to refer to a particular object, and it is these intentions that do the content-fixing work rather than the merely auxiliary demonstrations. There is also one caveat: Kaplan only commits to this view for "perceptual" demonstratives, i.e. those where the directing intention concerns a currently perceived object.

Kaplan's discussion of this change and its qualification is brief, and as far as his reasoning is concerned we only know that he considers one advantage of the new view to be the perspective it gives on an example from Donnellan (1966, p. 287) concerning definite descriptions:

> Suppose someone is at a party and, seeing an interesting looking person holding a martini glass, one asks, "Who is the man drinking a martini?" If it should turn out that there is only water in the glass, one has nevertheless asked a question about a particular person.

On this basis, Donnellan argued that definite descriptions have two distinct uses: *referential*, when they are used to identify a particular object and say something about it, and *attributive*, when they are used to say something about whatever object uniquely satisfies the description. The above is meant to be an example of a referential use. Kaplan believed that the same distinction arises for demonstratives, since descriptors like 'man drinking a martini' count as demonstrations for him. We need only imagine that one instead asked "Who is *that* man drinking a martini?" in the same scenario. Then on Kaplan's original demonstration theory, the actual semantic facts correspond to the attributive use: 'that man drinking a martini' can only refer to someone drinking a martini. At first Kaplan liked this result, since he thought a "proper semantical theory" (1989b, p. 583) ought not to confuse referential and attributive uses. But then he realised that his "directing intention" account does not *confuse* the two uses *per se*, but merely gives a different account of how they arise. That is to say, what counts as a referential use on the demonstration account counts as an *attributive* use on the directing intention account. However, Kaplan does not give an example of a referential use on the directing intentions account, although he writes that he still thinks the distinction is "fundamental" (1989b, p. 584), so it is even unclear exactly what we ought to take away from this example in

---

[2]One exception: the convention in English to refer to a few animate objects (e.g. ships) as if they are female. Some other languages are still more complicated, since grammatical gender does not always not track natural gender.

[3]One particularly important alternative is the *verbal* demonstration, e.g. 'chair under the tree' in the complex demonstrative 'that chair under the tree'.

support of the directing intention account. For now, I leave this issue aside and turn to another influential treatise on indexicals.

Perry (1997) draws a distinction between *automatic* and *intentional* indexicals. It is clear from these names that Perry is in agreement with Kaplan (1989b) about the semantic significance of speaker intentions for some indexical expressions (demonstratives among them), but the two differ in their conception of a context of utterance. What Kaplan calls "context" is equivalent to Perry's "narrow context": a small number of objective features of the utterance like agent, position and time. Perry's "wide context" consists of those features as well as anything else that may be relevant in determining the reference of an indexical (gestures are given as an example). Thus Perry ends up with a fourfold distinction of indexicals, where Kaplan's true demonstratives are closest to Perry's wide intentional indexicals. The motivation for having narrow intentional indexicals is that some expressions (e.g. 'now' and 'here') are significantly constrained by narrow features of context but still dependent on speaker intentions to fix their exact extension, while wide automatic indexicals have their reference fixed irrespective of speaker intentions but by factors outside of Kaplan's limited context (e.g. the extension of 'yea'[4] depends on the width between the speaker's hands). Since Kaplan does not consider these factors, it is difficult to draw exactly faithful equivalences between the two, but the table below gives an idea:

	Automatic	**Intentional**
Narrow	'I', 'actually'	'here', 'now'
Wide	'tomorrow', 'yea'	'this', 'that','there'

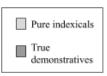|	
☐	Pure indexicals
■	True demonstratives

Figure 1.1: Kaplan vs. Perry on indexicals

Despite the support for taking speaker intentions as semantically significant amongst the most prominent theorists of indexicality, it is by no means universally accepted as the right move to make, and a number of alternative accounts of demonstrative reference[5] have been proposed. We might call these 'anti-intentionalist' approaches, and they usually involve some kind of *salientism*. For example, Gauker (2008; 2019) develops a theory inspired by the earlier

---

[4]'Yea' is a rather archaic English word meaning something like 'to the extent demonstrated', e.g. in an utterance of the sentence 'The fish I caught yesterday was yea big' while holding one's hands out so that the distance between them is the length of the fish.

[5]In fact, the debate does not stop at demonstratives. There is an equivalent disagreement over whether even the content of 'pure' indexicals is affected by speaker intentions, mostly motivated by so-called 'answering-machine' examples; an utterance of 'I'm not here right now' always comes out false on Kaplan's account but intuitively can be true. Theorists in favour of giving intentions this kind of content-determining role include Predelli (2002) and Åkerman (2009; 2010; 2015), theorists in opposition include Corazza et al. (2002) and Romdenh-Romluc (2002). For the sake of simplicity I avoid this debate, except in cases like Åkerman's where theorists explicitly give an account that is meant to apply to *all* kinds of indexicals.

work of Wettstein (1984), claiming that demonstratives refer to the most salient object in the context. Salience is determined by a combination of "accessibility criteria" like relevance to the conversation or being in direction of a pointing gesture. Mount (2008) puts forward a similar view, arguing that mutually-recognised salience between speaker and hearer fixes demonstrative reference.

Then there are the intentionalists. The most straightforward articulations of this view in the literature are Perry (2009) and Åkerman (2009; 2010; 2015). Both take intentions to be necessary and sufficient for fixing demonstrative reference. Perry's view has the additional feature of fleshing out Kaplan's notion of a 'directing intention' to reflect the fact that one may not always be currently perceiving the object in question. Other intentionalist theorists, like Stokke (2010), King (2013; 2014) and Speaks (2016) argue that a referential intention must be supplemented by an additional criterion related to audience recognition, whether this is actual (Stokke), potential (King) or intended/believed (Speaks).

The above theorists all assume that uses of demonstratives indeed *have* reference, i.e. that there is some fact of the matter about what each use of a demonstrative refers to. They also assume that demonstratives *themselves* refer, rather than as a result of being employed by speakers *who are referring*. This distinction is usually drawn as being between *semantic reference* and *speaker reference*, where the former is a relation that holds between (uses of) linguistic expressions and objects in the world but the latter is an *act* performed by speakers *using* linguistic expressions to communicate[6]. (From now on I will adopt the corresponding terms 'speaker referent' and 'semantic referent' when necessary but simply use 'referent' when it is clear which interpretation is intended, e.g. 'the referent of a use of a demonstrative' always means semantic referent.) However, there are some who deny that demonstratives' referential power extends beyond their use in speaker-reference. In particular, Kent Bach (e.g. 2005; 2008a; 2008b; 2017) argues that demonstratives technically do not semantically refer but simply function as speaker-referential devices, while Schiffer and Neale (forthcoming) suggest that we simply define the semantic referent of a use of a demonstrative as the object of the act of speaker reference it is employed in. These 'quasi-intentional' views can be seen as intermediate in the intentionalist/anti-intentionalist debate, because speaker reference itself is defined in terms of intentions, but these are Gricean communicative intentions rather than demonstrative-specific 'referential' intentions of the type intentionalists postulate. Although this distinction may not be entirely clear now, much more will be said about both the speaker/semantic reference distinction and Gricean communicative intentions in the next part of this chapter. I have introduced Bach/Schiffer and Neale here to give a more complete overview of the intentionalist/anti-intentionalist debate.

The ultimate goal of this thesis is to defend intentionalism. But rather than proceed in the typical manner (that is, delineating its advantages compared to salientism), I want to ask what kind of thing referential intentions might

---

[6]The differentiation originates in Kripke (1977), published in response to Donnellan's paper on referential and attributive uses of definite descriptions (as discussed above).

most plausibly be. This requires integrating concepts from the philosophy of action, where theories of intentions abound, especially since discussion of the underlying nature of referential intentions is noticeably limited in the existing literature. It is not until after proposing my own account of referential intentions that I return to the main debate and briefly explain how it surpasses existing salientist and quasi-intentionalist positions. The rest of this chapter explains why the implications of this project are important for a number of wider topics in the philosophy of language (§1.1-1.3), then introduces the specific questions that will be answered in the thesis (§1.4-1.5).

1.2   Intention-based semantics

The challenge of explaining the semantics of demonstratives in the first place is part of the wider project of giving a complete *descriptive* theory of meaning, i.e. a characterisation of the meaning of every linguistic expression. This type of theory can be contrasted with a *foundational* theory of meaning, which aims to explain how linguistic expressions come to have the meanings that they do. One strange thing about a particular variety of intentionalism, namely what I shall call 'reflexive intentionalism' as defended by Speaks (2016)[7], is the similarity between the type of referential intentions it supposes speakers to have and the intentions that feature prominently in a hugely influential foundational theory of meaning, namely H. P. Grice's intention-based semantics.

The defining feature of reflexive intentionalism is that it takes the fulfillment of speakers' referential intentions to be partially dependent on their (potentially) being recognised. Speaks considers two ways of specifying the exact content of such an intention, of which I am interested in the first:

> The value of a use of a demonstrative d in a context c is o iff: the speaker intends o to be the value[8] of d in c the speakers intends that his audience take o to be the object that the speaker intends to be the value. (2016, 329)

For comparison, Grice's ideas as presented in a series of works beginning in the middle of the 20th century were based on a notion of "non-natural" meaning or "meaning$_{NN}$", defined for utterances as follows:

> "A meant$_{NN}$ something by x" is roughly equivalent to "A uttered x with the intention of inducing a belief by means of the recognition of this intention". (1957, 384)

Or, written in a way that makes the three-pronged nature of such an intention clearer,

---

[7]Speaks also argues that King's (2014) coordination account collapses into a reflexive theory when taken to its natural conclusions, an assessment I am inclined to agree with and discuss further in §2.1.

[8]Speaks prefers to speak of the 'value' of a use of demonstrative where I favour 'referent'.

"U meant something by uttering x" is true iff, for some audience A, U uttered x intending:

(1) A to produce a particular response r

(2) A to think (recognize) that U intends (1)

(3) A to fulfill (1) on the basis of his fulfillment of (2) (Grice 1969b, 151)

Grice called this type of intention an "M-intention", although it has since come to be known as a 'communicative intention'. For Grice (1968), communicative intentions entirely determine "occasion-meaning", or the non-natural meaning of an utterance on a given occasion. But what makes Grice's efforts a foundational theory of meaning, rather than, say, a foundational theory of communication, is his attempt to explain what I have been calling 'conventional linguistic' (henceforth just 'linguistic') meaning in terms of occasion-meaning. For sentences, the gist is that x has linguistic or "timeless" meaning p iff there is an inclination amongst speakers in a linguistic community to utter x when occasion-meaning p, and that inclination is dependent on a mutual understanding that other speakers are similarly inclined (Grice 1968). Giving the linguistic meanings of individual expression is more difficult, given the systematic syntactic rules that constrain the ways in which they can appear in whole sentences.

Grice's (rather complex) solution, illustrated in his 1968 with the case of adjectives, rests on the idea of having "resultant procedures in [one's] repertoire", analogous in this case to having implicit knowledge of clause- and sentence-forming syntactic operations. A simplified account of, say, the adjective 'tall' goes as follows: 'tall' timeless-means 'of great height' in a linguistic community C iff speakers in C have in their repertoires the procedure to predicate 'tall' of a name n when they occasion-mean that the referent of n is of great height.

What about the occasion-meaning of individual expressions? It's occurred to some theorists, i.e. the quasi-intentionalists, that Grice's theory can be naturally extended to demonstratives (among other referential terms[9]). The most detailed formulation is in Schiffer and Neale (forthcoming), but the general idea is also endorsed by Kent Bach (e.g. 2017). Schiffer and Neale first give a Gricean definition of speaker reference[10]:

S referred to x (in the course of) uttering u iff in uttering u, S meant an x-dependent proposition. (p. 4)

'Meant' here of course indicates Gricean occasion-meaning. This leads to a definition of referring-with an expression,

---

[9]Schiffer and Neale (forthcoming), for example, is a theory of both demonstratives and pure indexicals.

[10]Actually, the following definition is of what Schiffer and Neale call 'primary' speaker-reference as opposed to 'secondary', which are differentiated according to whether the proposition to be communicated is dependent on the object being referred to or not. For example, if I utter 'That dog behind the red car looks lost', then (presuming I am using the words with their usual meaning), I primary-refer to the dog but merely secondary-refer to the red car. Nonetheless, it is enough for our purposes to stick to primary reference

> In uttering u, S referred to x with e, relative to its i-th occurrence in u, iff for some person A and property φ, S intended it to be common ground between S and A that the i-th occurrence of e has φ and, at least partly on that basis, that S referred to x in uttering u. (p. 6)

and ultimately a theory of semantic reference for demonstratives:

> x is the referent of demonstrative 'that' ['this', 'it'] on occasion o iff the speaker referred to x with 'that' ['this', 'it'] on o. (p. 10)

Remember, however, that Schiffer and Neale (and Bach) do not endorse intentionalism about demonstratives, because they do not take demonstratives themselves to have a special 'automatic' referential power. And although Speaks' conception of a referential intention has surface similarities to a Gricean communicative intention, crucially it concerns the *semantic relation* of reference and *linguistic meaning* rather than the *act* of referring and *speaker occasion-meaning*. There are two key things to say about this difference, one of which is delayed until §1.3. The other, to be addressed now, relates to the influence Grice has had on conceptions of the relationship between 'semantics' and 'pragmatics', a longstanding issue in the philosophy of language that is somewhat obscured by the varied applications of those terms between theorists.

1.3   The semantics/pragmatics divide

On one simplistic definition, semantics deals with linguistic meaning, while pragmatics deals with meaning in context. But this would make all philosophical discussion of indexicals a part of pragmatics, since they are defined by their context-sensitivity[11]. On the contrary, Kaplan's and Perry's are now generally thought of as semantic theories. This is because they attempt to constrain contextual contribution by explaining how it depends on linguistic meaning. If we can give a systematic rule describing exactly which features of context are relevant and how, like in the case of 'I', then we are back in the domain of semantics. So it might be better to say instead that semantics deals with meaning dictated by linguistic properties alone, while pragmatics deals with other elements of meaning communicated by speakers in context. This also explains why many theorists in the intentionalist/anti-intentionalist debate, as well as elsewhere in the philosophy of language, use 'semantic value' or 'semantic content' to mean the referent of an indexical in context rather than its Kaplanian character (although I will continue to use 'referent').

Yet when we return to the literature in the intentionalism/anti-intentionalism debate, we find that 'pragmatics' has a still more specific usage. For example, Bach views demonstrative reference as a 'pragmatic' matter rather than a 'semantic' one. His reasoning is that speakers' do not have

---

[11]For a notable example of this definition in practice, see Montague's (1968) foundational work *Pragmatics*.

special 'referential intentions' beyond those that are part of their overall communicative intentions (2017, 58). But why does this make them 'pragmatic'? Gauker is more explicit about his definition; although his first paper on the topic is titled Zero tolerance for pragmatics, he specifies that he is using the term to refer just to "the ways in which the proposition expressed depends on what the speaker intends in speaking" (2008, 359). But why is such importance placed on excluding speaker intentions, especially since Gauker's theory admits a wide range of other contextual features as semantically relevant?

The reason is that Grice's theory gives us as a way to explicate the notion of 'other elements of meaning communicated by speakers in context', namely as occasion-meaning. So it is not that Bach and Gauker are using a different definition of semantics and pragmatics per se, rather they are employing a particular way of making those definitions more precise. And they are far from the only ones: Grice's intention-based approach is the basis of some of the most influential pragmatic theories today (e.g. Sperber & Wilson 1986/Carston 2002; Levinson 2000; Recanati 2004). Thus I will continue to use 'semantics' and 'pragmatics' in accordance with the second definition—semantics is concerned with content dictated by linguistic meaning, pragmatics with other occasion-specific communicated content—but rely on the background understanding that these definitions are often themselves defined in Gricean terms.

Given these considerations, it should be clearer why suggesting that reflexive referential intentions are semantically relevant—as Speaks (2016) does—is particularly controversial; at least on one understanding, it seems to blur the distinction between semantics and pragmatics altogether. In fact, the idea that referential intentions are semantically relevant is so objectionable to Bach that he leaves demonstratives without any kind of semantic content at all: "If I am right about demonstrative reference," he writes, "you can still do truth-conditional semantics if you want, but assigning semantic references to demonstratives is entirely stipulative" (2017, 59).

One possible out for the intentionalists would be to deny that it is important to maintain a strict distinction between these two types of meaning. Indeed, in contemporary philosophy talk of the semantics/pragmatics divide tends to be associated with the minimalism/contextualism debate of the early 2000s, in which minimalists argued that the distinction tracks something fundamental about meaning while contextualists argued it is mostly arbitrary (see Jaszczolt 2012 for an overview). Hence it is tempting to speculate that referential intentionalists are somehow supporters of contextualism. I think it is more accurate, however, to say that the real disagreement is over whether the intentionalism/anti-intentionalism debate has any effect on the minimalism/contextualism debate at all. If not, there would be no need to associate intentionalists with either position. Indeed, worries about pragmatic intrusion are notably absent from the intentionalist literature; if intentionalists are concerned that their theory may have damning implications for the entire paradigm of speaking about linguistic-meaning as separate from occasion-meaning, they have been very quiet about it.

To clarify: if communicative intentions have an impact on semantics as well

as pragmatics, it upends the traditional distinction between the two. But for the intentionalists, referential intentions are not communicative intentions. They are distinct referential intentions with a different type of content—that is, content concerning the referents of the demonstratives we use (see e.g. King's response in his 2014 to Bach's arguments). If this is the case, it is possible that we can allow them into semantics without allowing communicative intentions to follow. But Bach claims that "in using demonstratives we have intentions to refer but not intentions for demonstratives themselves to refer" (2017, 58), where our "intentions to refer" are inseparable from our general communicative intentions and have no metasemantic content. So this point about the separability of referential and communicative intentions, heavily emphasised throughout Bach's writing on the topic, turns out to determine the relevance of intentionalism to the semantics/pragmatics divide. It is also key for the issues discussed in the next section, on reference.

1.4   Theories of reference

A reminder: the Gricean view of the relationship between occasion-meaning and linguistic-meaning gives us a way to relate the act of referring, as is performed by speakers using linguistic expressions, and the semantic notion of reference, as holds between linguistic expressions and objects in the world? The former is a matter of the speaker's occasion-specific communicative intentions, the latter is nothing but a tendency to be used when speakers have particular communicative intentions. In Bach/Schiffer and Neale's view, then, demonstratives are simply so wide-ranging in their potential uses that no such tendency can be identified; they never get to the level of semantic reference (Bach), or it is simpler to identify the semantic referent of a use of demonstrative with what it is being used to refer to on a given occasion (Schiffer and Neale).

But, as we have seen, the intentionalist needs to posit a special type of intention beyond the Gricean method: a *referential* intention, or one for the demonstrative itself to refer. Again, it is not clear that one can give a satisfactory account of the connection between these two types of intention. If it is possible, however, it might give a new perspective on the relationship between speaker and semantic reference in the case of demonstratives (and hence, as discussed in the previous section, the relationship between semantics and pragmatics). If not, it would seem that a Gricean theory of reference is not applicable in the case of demonstratives.

There is also a possibility that the right theory of reference can help the intentionalists strengthen their position. Consider this charge against the intentionalists, as briefly suggested by Bach (2017, 64 fn. 14) and developed at length by Gauker (2019): their theory is circular. Individual differences between intentionalist theories notwithstanding, they all have in common the claim that a speaker S's referential intention plays some part in fixing an object o as the semantic referent of her use of a demonstrative d. Part of the content of this intention, its being referential, is to have o be the referent of d. But since the

intention itself fixes the referent, how are we to understand the content of S's intention without begging the question?

Gauker claims any attempts to describe a non-circular notion of 'reference' as it occurs within the content of S's intention will leave the theory subject to counterexamples. For example, Gauker suggests but quickly dismisses that the intention's content might be related speaker reference instead of semantic reference. He claims that this amounts to relating it to semantic reference, but this seems to me to take a stand on the question of the relationship between communicative and referential intentions without proper justification. In fact, if Gauker were correct that "a speaker intends to use 'that' to refer to an object o if and only if the speaker intends to use 'that' in such a way that his or her utterance of 'that' does refer to o"[12], then there would be nothing to the difference between Speaks and Bach that I have emphasised so much[13]. On the contrary, I will argue that in the case of demonstratives the speaker/semantic reference relationship is not obvious. It is key, however, to solving not only the circularity problem, but reconciling the differences between Speaks' and Bach's theories and thereby making progress towards answers to the wider questions we have seen so far.

1.5   Identifying referential intentions

In the following preliminary chapter, Chapter 2, I select a type of intentionalism that can be used as a basis for the arguments made in the remaining chapters, namely simple intentionalism as defended by Perry and Åkerman. The main argument for this choice is that the move to reflexive intentionalism (and theories like it) is unnecessary to avoid the problems associated with simple intentionalism. Then I move on to attempting to give the most plausible account possible of a simple referential intention.

The discussion up until now has already revealed two problems with asserting that referential intentions exist: the content of referential intentions must be non-circularly specifiable, and they must be separable from speaker-referential intentions. Otherwise, they simply could not *be* as intentionalists claim they are.

With that in mind, we may describe a third 'identification problem': that of reliably determining from a number of apparently referential speaker intentions which is semantically relevant. It has been observed that a speaker using a

---

[12]Gauker does not straightforwardly endorse this view, but thinks it is "prima facie plausible" (2019, 116). Interestingly, it seems to be implicitly endorsed by Perry (2009), who shifts back and forth between talk of speakers and expressions referring without establishing the relationship between the two.

[13]It is possible that Gauker thinks there is no difference; in his 2008 he cites Bach (2005) as a proponent of the intentionalist view, although Bach specifies in that paper that he does not think referential intentions determine (semantic) demonstrative reference (e.g. 25-6). It might also be a misinterpretation; Bach's view in his 2005 is certainly more obscure on the speaker/semantic distinction than in his 2008a, 2008b, and 2017. Since Gauker does not cite Bach in his 2019, as things stand the evidence in favour of either option is scarce.

demonstrative may, for example, have the intention to refer to the object at which she points, but also the *de re* intention to refer to a particular object *o*. Of course these would be linked by the belief that the object at which she points is identical to *o*, but they may come apart in some circumstances and if so it must be specified which is the ultimate semantic arbiter. One case frequently discussed in the literature (e.g. Perry 2009; King 2013; Speaks 2017), for example, is the imagined 'Carnap/Agnew' scenario originating with Kaplan:

> Suppose that without turning and looking I point to the place on my wall which has long been occupied by a picture of Rudolf Carnap and I say "[That]^{14} is a picture of one of the greatest philosophers of the twentieth century. But unbeknownst to me, someone has replaced my picture of Carnap with one of Spiro Agnew." I think it would simply be wrong to argue an 'ambiguity' in the demonstration, so great that it can be bent to my intended demonstratum. I have said of a picture of Spiro Agnew that it pictures one of the greatest philosophers of the twentieth century. (Kaplan 1978, p. 239)

The standard approach to this 'problem of conflicting intentions' is to give a "theory of trumping" (Speaks 2017), or an account of when one intention 'trumps' others to become semantically relevant. Accordingly, I think it necessary to take a stance on a solution before moving onto the other two identification problems, else it's hard to say what we are trying to identify. As it will be seen, however, attempting to achieve even this first aim requires something rarely provided in the context of the intentionalist/anti-intentionalist debate: a theory of intentions with which we can test the claims being made. Nor is this because it is easy to decide upon one; on the contrary, theorists in the philosophy of action have long been puzzled by the nature of intention. I shall also show that the matter is not trivial, since selecting a theory of intentions changes the approach that must be taken to the problem of circularity.

Hence Chapter 3 addresses whether it is possible to develop a notion of referential intentions such that 1) it can be reliably determined which are semantically relevant in cases of conflict (§3.1) 2) their relationship to communicative intentions is established in a way which might assuage the anti-intentionalist's concerns (§3.2) and 3) their content is non-circular (§3.3). This is done by relating each criterion to existing theories in the philosophy of action and arguing for a particular feature of referential intentions must have as a result.

As regards the problem of conflicting intentions, a connection has already been pointed out by Perry (2009) and King (2013), namely that we may think of an intention as elements of a wider plan of referential action. King suggests that this is best understood by way of Bratman's (1987) theory of intentions as part of hierarchically structured plans, where the referential intention is the 'lowest' in the hierarchy. The primary result of this section is a development of Perry's (2009) ideas to independently identify this ultimately referential intention.

---

^{14}Kaplan actually wrote 'Dthat', using his general stand-in for a demonstrative.

With respect to point 2), first I consider Viebahn's (2018) response to Bach's arguments, which suggests that we view referential intentions as "resulting from and embedded in communicative intentions" (2018, 3). This line of reasoning also makes use of Bratman's theory, so gels nicely with prior results, but I will argue that it falls somewhat short in providing evidence for the existence of referential intentions. My idea to improve it makes use of John Searle's (1983) distinction between future-directed "prior" intentions and "intention in action"; the unique characteristic of semantic-referential intentions as opposed to speaker-referential intentions is that, as intentions in action, their conditions of satisfaction are difficult to distinguish from their completion. This also provides a new perspective on point 3), according to which it is *not* an issue for intentionalists that the content of a referential intention mentions reference itself, but it *is* an issue that no particular state of affairs has yet been specified as constituting 'reference'. This new version of the circularity problem is the starting point for Chapter 4.

1.6   Explicating referential intentions

In Chapter 4, I relate the remains of the circularity objection to an issue originating with Speaks (2016) and commented on at length in Viebahn (2018), namely that young children are capable users of demonstratives but likely do not have complex intentions concerning the referents of demonstratives. I argue that both these considerations push the intentionalist to specify a state of affairs linking a use of a demonstrative and its referent, i.e. give a more complete theory of demonstrative reference. For the answer, I turn to literature from child language acquisition. In particular, Diessel (2006) argues that children learn to use demonstratives as a tool for establishing *joint attention*. Hence referential intentions should be thought of as intentions to establish joint attention on a particular object. This leads to the introduction of a new explication of intentionalism: joint-attention simple intentionalism (JA-SI). Chapter 5 compares JA-SI to existing intentionalist, salientist and quasi-intentionalist theories, and Chapter 6 fleshes out some more details of the theory. Finally, in Chapter 7 we return to the issues raised earlier regarding theories of meaning and reference and consider them in light of JA-SI.

Chapter 2

Varieties of intentionalism

Call a speaker's intention to have a certain object o be the referent of her use of a demonstrative a 'simple referential intention'. At the core of all intentionalist proposals is the claim that simple referential intentions are necessary to fix demonstrative reference, but there is no consensus on what *else* is necessary. We'll call the most basic form of intentionalism, according to which simple referential intentions are both necessary and sufficient to fix reference, 'simple intentionalism':

> *Simple intentionalism (SI)*
> The referent of a use of a demonstrative d is the object o such that
> the speaker intends o to be the referent.

SI originates with Kaplan's (1989b) brief remarks on "directing intentions" and is advocated for in detail by Perry (2009) and Åkerman (2009). In my view, it is also the most plausible variant of intentionalism on offer. Therefore I will use it as a proxy for intentionalism in general throughout this thesis, meaning that arguments and suggested developments will be aimed at SI. The purpose of this chapter is to explain my rationale in doing so. I do so by introducing SI's primary competition, *reflexive intentionalism (RI)*, which has been proposed as a way of constraining and improving upon SI. I argue that SI can already be constrained by communicative practices and that RI's supposed primary advantage over SI — avoiding so-called 'Humpty Dumptyism' — is not really an advantage, because the very same characteristics that lead to Humpty Dumptyism are the principal strengths of intentionalism.

   Of course, none of this implies that the insights of other intentionalist theorists will be abandoned after this chapter. On the contrary, many are still relevant to SI, and I shall return to reflexive intentionalism in Chapter 5 to discuss the extent to which its merits are preserved in my own preferred explication of simple intentionalism. For now, though, I proceed with delineating the advantages of SI.

2.1   Against reflexive intentionalism

Reflexive intentionalism is characterised by its Gricean influence. It conceives of referential intentions as fulfilled by their own recognition:

> *Reflexive intentionalism (RI)*
> The referent of a use of a demonstrative d is an object o iff the speaker intends:
>
>   1. o to be the referent
>   2. her audience to recognise her intention (1)

(1) in the above definition is of course just a simple referential intention; from now on I call (2) the 'reflexive referential intention'. I consider Stokke (2010), King (2014), and Speaks (2016) to be the most detailed defences of RI in the literature. All three are motivated by a desire to account for intuitive instances of reference failure for which SI predicts referential success. The following example is paradigmatic:

> I am sitting on Venice Beach on a crowded holiday looking south, with swarms of people in sight. I fix my attention on a woman in the distance and, intending to talk about her and gesturing vaguely to the south, say "She is athletic". You, of course, have no idea who I am talking about. It seems quite implausible in such a case to say that I succeeded in securing the woman in question as the value of my demonstrative simply because I was perceiving her, intending to talk about her and gesturing vaguely in her direction. (King 2014, p. 223)

Much more will be said about this objection to SI in §2.1.1. For now, I want to explain why I have named Stokke, King and Speaks as defenders of RI when in fact only Speaks (2016, p. 329-30) explicitly suggests it, and even he does so alongside another possible alternative. The attribution to Stokke and King is due to 1) a convincing argument from Speaks that Kings *Coordination Account* is equivalent to an entirely speaker-based theory when taken to its natural conclusions and 2) the fact that Stokke's account, which includes an audience-based *Uptake Constraint* in its referential criteria, is 'halfway' from King's to Speaks', so the same argument can be applied. Finally, I claim that RI is the most viable of the two speaker-based alternatives proposed by Speaks, hence the best way of developing his, Stokke's, and King's positions.

   Stokke's position is that SI must be supplemented by an Uptake Constraint requiring that reference cannot be achieved unless the speaker's audience is "in a position to" recognise her simple referential intention (2010, p. 388). What it means to be in such a position is left deliberately vague, in favour of spelling out the details of Stokke's intention-sensitive character-content framework. But King (2014, p. 225) is more clear on the particulars of his Coordination Account (edited slightly for consistency):

Coordination Account (CA)
The referent of a use of a demonstrative d is an object o iff:

1. the speaker intends o to be the referent
2. a competent, attentive, reasonable hearer would take o to be the object that the speaker intends to be the referent[1]

So, the key difference between Stokke's and King's theories is that the former constrains reference depending on the *actual* audience's potential for intention recognition, the latter some *idealised* audience's potential. Now we turn to Speaks' argument that the Coordination Account leads to an entirely speaker-based theory, such as RI. I am inclined to agree with his line of reasoning, and also note that its first step gets us to Stokke's account.

Speaks' doubts about the usefulness of the CA are initially motivated by his 'sneaky students' example (2016, p. 312):

> I teach Philosophy 101 in a large auditorium which darkens during the lecture so that the students can better see the slides; in fact, though, it becomes a bit darker than it needs to, to the point where I cannot see the students during the lecture. The students have figured this out, and now, very quietly, exit the room minutes after the lights go down, and return minutes before the lights go back up. In the interim, I'm speaking to an empty room.

Speaks' point here is that demonstratives used during his lecture do not suddenly fail to determined referents simply due to the absence of an audience. The obvious recourse for King is to emphasise that the audience is meant to be a hypothetical one possessing particular qualities. Indeed, it seems that if there *were* any competent, attentive, reasonable hearers present during the lecture, they would be able to recognise the lecturer's simple referential intentions (provided they are made suitably manifest). So this is not a counterexample to the CA. Yet Speaks objects to this response, for reasons illustrated here by his imagined 'sudden blindness' case (2016, p. 315):

> I'm having a beer with a friend at a bar, and, pointing to her glass, say 'That beer looks flat'. Unfortunately, she was struck blind moments before my utterance, and hence was unable to discern the object of my referential intention.

Speaks claims that the demonstrative in this case successfully refers, although his referential intention was not recognisable to the "perfectly attentive and competent" (p. 316) hearer that was in fact present. Anticipating that the CA-theorist will attempt to build normal eyesight into her hearer idealisation conditions, Speaks points out that this would suggest we can use demonstratives in the same way when speaking to visually impaired and unimpaired audiences

---

[1]Original: "1) the speaker intends o to be the value; and 2) a competent, attentive, reasonable hearer would take o to be the object that the speaker intends to be the value."

and still achieve reference. That seems plainly incorrect. Nor would it be possible (I claim rather than Speaks) to fix this problem with Stokke's Uptake Constraint, because that would leave us unable to deal with the 'sneaky students' example where there is no actual audience. In principle, one could at this point suggest a combination of Stokke's and King's theories that has conditions like 'if there is an actual audience, then the Uptake Constraint must be met' and 'if there is no audience, then the idealised-hearer constraint must be met'. Aside from being inelegant, though, I think we could continue in the vein of Speaks' and construct counterexamples for such a theory. We need only imagine that there is an actual audience, say of blind individuals, but the speaker believes there is no audience, so feels free to use pointing gestures alongside her demonstratives. If one has any intuition at all that reference is achieved in such a scenario, it seems much more plausible that the relevant feature in all cases is *the speaker's estimation of* her (believed) audience.

So the upshot of Speaks' argument is that when the CA is extended to fit our intuitions about certain odd cases, it inevitably collapses into a theory based entirely on the beliefs and intentions of the speaker. I contend that the same considerations apply to Stoke's Uptake Constraint. That is to say, if one wants to constrain SI with some kind of recognisability criterion, that criterion ought to be about the speaker's believed audience rather than an ideal or actual one. Speaks suggests that this can be formulated as either an additional intention, as in RI (he calls it "I+I"), or as a justified belief about the audience's ability to recognise the speaker's simple referential intention (2016, p. 330, slightly edited):

> *Intention + Justified Belief (I+JB)*:
> The referent of a use of a demonstrative *d* is an object *o* iff:
>
> 1. the speaker intends *o* to be the referent
> 2. the speaker would be justified in believing that his audience will take *o* to be the object that the speaker intends to be the referent

Speaks does not take a stance on which is preferable; his purpose is merely to show that any intentional account must remain grounded in the speaker's psychological state. He does hint, however, that RI might be preferable in accounting for cases like this one:

> Imagine that we are sitting on a couch in my house, and I think I see something quite surprising—like a bird quickly flying past the doorway. You do not flinch, so I am almost sure that you did not see what I think I saw. But, to be sure, I might ask you: 'Did you see that?' Supposing that a bird really did fly past the doorway, it seems clear that my use of 'that' here succeeds in referring to the bird. (2016, p. 328)

I agree with Speaks' intuition in this case. And I think, as he does, that demonstrative uses of this kind are common and hence ought to be accounted for in any

theory we advocate for. There is also a second type of use that I think separates RI from I+JB, namely in which very little is known about the audience. One example might be a newsreader who uses any number of demonstratives in her broadcast, each time giving various clues to her referential intentions but not knowing whether each individual member of her audience will have the requisite background knowledge, eyesight, linguistic competence, etc. to pick up on them. Again, RI comes out on top here[2]. Moreover, I+JB introduces an unnecessary complication in the form of having to explain what constitutes a justified belief. Thus I consider RI to be SI's primary contender, given its (implicit and explicit) support in the literature and its strength over other speaker-based options. In the rest of this chapter I explain why I think RI is misguided.

2.1.1 Humpty Dumptyism

"When I use a word," says Humpty Dumpty to Alice in Lewis Carroll's *Through the Looking Glass*, "it means just what I choose it to mean—neither more nor less"([1871] 2002, p. 190). A ridiculous claim, surely, yet it seems to be exactly the simple intentionalist's position on demonstratives. If intentions alone determine their referents, then speakers can make demonstratives refer to pretty much anything. We might note, however, that this isn't technically the same as making demonstratives *mean* pretty much anything. For whatever a speaker's intentions are, she cannot change the *standing meaning* (Kaplanian character) of a demonstrative, which stipulates that its uses refer as a function of her intentions.

Nonetheless, theorists (e.g. Wettstein 1984; Gauker 2008; Stokke 2010; King 2014; Speaks 2016) find this devolution into 'Humpty Dumptyism'[3] to be wildly improbable given how easy it is to construct counterexamples, like King's Venice Beach example above. One of King's primary motivations in posing his Coordination Account, which as we have seen above is best understood as RI, is to avoid these unintuitive results. Speaks (2016) is also moved to consider RI by the Venice Beach example, Stokke (2010) by a similar one[4]. I'll call these

---

[2]Well, it comes out better than I+JB. It's not obvious that RI can straightforwardly deal with this kind of case, either, at least not without allowing for intentions that concern a potential audience or intentions that concern multiple future communicative 'events' and audiences, where the ultimate number of events and the character of the audiences cannot be known to the speaker. But the point is that it is in principle possible to have such intentions, whereas it's not so clear that one can have (the right kind of) justified beliefs about potential or unknown audiences.

[3]The term was first coined by McKay (1968), who objected to Donnellan's (1966) claim that definite descriptions have intention-determined 'referential' uses.

[4]Stokke's example is actually quite different, and at least to me rather strange. It concerns a U.S. customs officer who, knowing that the woman he is talking to speaks no English, says, "Now I'm going to explain to you why you are in violation of your visa so that you can't say you haven't heard it." I think it strange because it concerns a different type of indeterminacy; rather than being unable to identify the intended referent of 'you', this woman would not be in a position to know that she should be seeking an intended referent in the first place. This is equally problematic for the non-demonstrative parts of the utterance, and I don't see the benefits of introducing such a complication. Still, regardless of his unusual route, Stokke reaches the conclusion that simple referential intentions are insufficient for fixing reference.

cases, in which reference intuitively fails because the speaker's simple referential intention is not plausibly recognisable, 'cases of intentional obscurity'.

However, I think these theorists are misguided in advocating RI as a strategy to avoid Humpty Dumptyism for two reasons: 1) our intuitions about cases of intentional obscurity can already be explained with reference to norms of communication/reflexive Gricean communicative intentions, so an additional reflexive referential intention is unnecessary, 2) in doing so they lose the ability to explain cases where demonstratives seem to successfully refer either without an audience or without a reflexive referential intention. I now turn to explaining these reasons in detail.

We begin with our intuitions of reference failure in cases of intentional obscurity, since these are what motivate Humpty-Dumpty objections to intentionalism. One explanation for these intuitions, favoured by RI-theorists, is that they track semantic content; reference fails because the speaker does not have a genuine intention for the audience to recognise her simple referential intention. But there is also another explanation, versions of which are told by both Perry (2009) and Åkerman (2015). It goes like this: SI is the correct account of demonstrative reference. But semantic content is not much use to speakers if they cannot use it to communicate. Since speakers' intentions are not accessible to their hearers in the same way external features of context are, it is normal to try one's best to make one's simple referential intention obvious. In fact, it is *so* normal that cases of intentional obscurity 'feel' like cases of reference failure, because they are so at odds with the way we are accustomed to using demonstratives. Thus in a way the Humpty Dumpty appeal is justified; speakers technically *can* refer to whatever we want. But in fact we are so inclined to act in accordance with communicative norms that we generally *do not* form such outlandish referential intentions (Åkerman 2015).

Although this story does not depend on any particular theory of communication, its consequences are particularly clear when it is combined with a Gricean account of speaker occasion-meaning. Suppose a speaker utters a sentence including a demonstrative d with the intention of inducing in her audience the belief that p via the audience's recognition of her intention. Subsumed under this overall communicative intention are a number of intentions concerning individual linguistic expressions of the sentence that are also fulfilled by their own recognition, including the intention to communicate something about an object o using d. If we take this to be an intention to speaker-refer to o using d, we have Bach's (2005; 2008a; 2008b; 2017) theory of reference. And then we see the motivation for his charge that reflexive referential intentions are redundant (e.g. 2017, 65): they don't do anything more than communicative intentions[5]. In particular for our purposes, the fact that a demonstrative utterance is normally accompanied by a corresponding communicative intention, combined with the story told in the previous paragraph, already explains why cases of intentional

---

[5]Of course, Bach also thinks that simple referential intentions are redundant: for him, speaker reference via communicative intentions exhausts the referential power of demonstratives. But this problem applies equally to both SI and RI, and is taken up again in Chapter 3.

obscurity are unintuitive.

Another thing that strikes me as odd about the RI-theorist's explanation of cases of intentional obscurity is its departure from Kaplan's original framework. If the idea is to give a 'rule' for demonstrative reference as for the pure indexicals, it shouldn't be a problem that sometimes the audience cannot identify the referent even when the rule is followed. For example, suppose I make a phone call to the police to anonymously report a crime. Throughout the call I talk about what *I* saw, heard, etc., but the person on the other end has no idea who *I* am. Does this mean that my utterances of 'I' failed to refer to me? I think the answer is 'no', both intuitively and as a result of Kaplan's theory. His narrow notion of context as including just a few basic features of the utterance does not imply that these features are *always* identifiable by the hearer, just that they normally are. I think we are happy to attribute reference in abnormal cases when the speaker is afforded a plausible communicative intention, like in the anonymous-report scenario. Thus we have reason to question the validity of our 'semantic' intuitions in cases like the Venice Beach example — perhaps they are not reflective of how we actually track semantic content in real-world situations where speakers act (reasonably) rationally.

The RI-theorist might object that we should not pick and choose intuitions in this way, deciding which do and do not track 'true' semantic content depending on our preferred theory. But as we are about to see, there are cases in which RI predicts demonstrative reference to fail but it in fact intuitively succeeds. Hence proponents of both SI and RI must explain away *some* intuitions—at least the SI-theorist can appeal to communicative intentions to do so systematically. In fact, it is exactly when an imagined speaker is given a plausible (absence of a) communicative intention that the reference failure intuition disappears. This brings me to our second reason to avoid supplementing SI with reflexive referential intentions: we lose the ability to deal with a significant number of demonstratives uses as a result, as well as any theoretically motivated explanation of why this occurs.

As argued above, RI-theorists are too quick to take communicative failure to have semantic import. Commenting on an example of reference failure due to impossibility of audience uptake, Stokke even writes that "there is a strong sense in which [the speaker] is not engaged in what we take to be earnest communicative practice" (2010, p. 390). But being engaged in earnest communicative practice should not be a condition on determining semantic content. The strongest evidence for this claim is the simple fact that utterances can have semantic content without being used for communication. I want to discuss two types of these 'cases of content-communication divergence': first, self-talk cases, and second cases in which the speaker has no intention of the audience recognising her intended referent. The latter type is illustrated by Åkerman's (2015) incapable-audience cases and Michaelson's (2013) "sneaky reference" cases.

The first type is familiar: speakers may mutter under their breath, talk to themselves, or use demonstratives in their inner speech. This is in fact addressed by Speaks (2016, p. 311), who points out that an audience-relativised restriction in the style of Stokke's (2010) Uptake Constraint can in fact account

for self-talk: if we identify the audience with the speaker, since one is always in a position to recognise one's own intentions the restriction becomes redundant and we are back at SI. As we have seen, however, audience-relativised restrictions can be forced to become speaker-based constraints. Yet self-talk has some rather strange implications for our preferred speaker-based theory, RI. While it is obvious that speakers have privileged access to their own intentions, it is much more doubtful that they have intentions to recognise their own intentions[6]. It is not clear how the reflexive intentionalist can avoid this problem, while the simple intentionalist never even encounters it. In fact, I think this demonstrates that Humpty Dumptyism is in some ways a desirable result: it *does* seem to be the case that demonstratives can refer to whatever a speaker wants when used without an audience.

In the second type of content-communication divergence case, the speaker's utterance *it* is directed at an audience, but she has no intention for that audience to recognise her simple referential intention. These cases, illustrated by examples from Åkerman (2015) and Michaelson (2013), do not seem to give rise to the same intuitions of reference failure as in the Venice Beach example. Åkerman is concerned with our not-infrequent tendency to "talk to pets, babies, or even dead people", like in this scenario (2015, p. 87):

> Consider a father who turns to his two-month old child (who has just had a frightening experience) and says: "Calm down, you're safe now here with me"....it seems wrong to deny that he can, in this very case, intend to use 'you' to refer to the child, 'me' to refer to himself, and so on, and it is plausible that he can also succeed in so referring, even if he is fully aware that the child cannot understand the expressions in this way.

Even more evidence against RI can be found in Michaelson's (2013) examples of "sneaky reference", cases in which the speaker's simple referential intention is genuine but her overall communicative intention misleading. For instance:

> You and I both work for a large investment bank. We've just been paid an obscenely large bonus, and we are both at the office very early one day. We run into each other near a window overlooking the parking lot, which is nearly empty at this hour. There is, however, a small cluster of cars which includes several nondescript vehicles and, in addition, a McLaren F1. You ask me what I decided to do with my bonus, and I respond by pointing out the window at the cluster of cars and saying: "I bought that." In fact, I bought the Toyota Camry parked next to the McLaren, and I intend to refer to

---

[6]As Speaks (2016, fn. 16) points out, this same problem plagues Gricean analyses of speaker meaning; we seem to mean things when talking to ourselves without intending that we come to have particular beliefs on the basis of our recognition of our own intentions. But at least the Gricean has the option of claiming that speaker meaning is fundamentally rooted in communication, so that we cannot be said to really *mean* anything during self-talk. I don't think the reflexive intentionalist would want to do the same for semantic content, i.e. deny that demonstratives can refer during self-talk.

the Camry, not the McLaren. That said, I've decided to have a bit
of fun with you. Even in a situation like this one, I take it, I have
plausibly succeeded in referring to my Camry. (2013, p. 91-2)

So, proposing RI to avoid Humpty Dumptyism seems like overcorrection;
it fails to attribute reference in a considerable number of cases that SI has
no issue with. Of course, the RI-theorist might object that cases of content-
communication divergence are not cases of demonstrative reference. But then
she must explain why they intuitively *seem* to be and, unlike for the SI-theorist
who must explain why cases of intentional obscurity seem *not* to be cases of
demonstrative reference, appeal to communicative norms will be of no use. Thus
introducing RI as a 'solution' to Humpty Dumptyism is misguided; Humpty
Dumptyism is only undesirable when it is at odds with communicative norms,
and if we attempt to restrict a theory of demonstrative reference based on those
norms it loses a lot of explanatory power. The more elegant and fruitful strategy
is to revert to SI.

Chapter 3

Identification problems

3.1 The problem of conflicting intentions

The problem of conflicting intentions can be so put: speakers often have multiple intentions associated with their use of demonstratives that might be called 'referential', and in some (unusual) cases those intentions concern different objects. Crucially, the speaker is not aware of this fact, so her intentions are consistent. This raises a question for intentionalism: which intention, if any, is 'ultimately' referential? Throughout my attempt to answer, I shall refer back to Kaplan's 'Carnap/Agnew case' as introduced in *Dthat* (1985):

> Suppose that without turning and looking I point to the place on my wall which has long been occupied by a picture of Rudolf Carnap and I say "[That][1] is a picture of one of the greatest philosophers of the twentieth century. But unbeknownst to me, someone has replaced my picture of Carnap with one of Spiro Agnew....I have said of a picture of Spiro Agnew that it pictures one of the greatest philosophers of the twentieth century. (Kaplan 1978, p. 239)

As is clear from his conclusion, Kaplan is still supporting the demonstration account at this point, and in fact he doesn't mention the case at all in *Afterthoughts*. It is fervently discussed in the intentionalist literature, however. The problem is that the speaker described plausibly has both an intention to fix the picture of Carnap as the referent and an intention to fix the picture in the direction of the pointing gesture as the referent. Intentionalism does not *prima facie* decide between these which is the determiner. It's also worth noting from the outset that, contra Kaplan, neither answer might be obviously 'correct', in the sense that we generally intuit one picture to be the actual referent and hence that result is a clear goal for potential explanations to aim at. While Kaplan thinks it evident that the picture of Spiro Agnew is the referent, Perry (2009)

---

[1]Kaplan actually wrote 'Dthat', using his general stand-in for a demonstrative.

has suggested that fleshing out the case in different ways can manipulate our intuitions in this regard[2]).

Our difficulty in pinning down a target solution to the problem suggests an answer given by King (2014): conflicting intentions lead to reference failure, since one simply cannot determine what the 'true' object of the speaker's intention was. This is a tempting position to take, as it dissolves the problem rather quickly and explains our shifting intuitions, but I think that it cannot be right. The problem is that we often have incorrect beliefs concerning the objects of our intentions, yet they are rarely seen as able to obstruct reference. In fact, supposing that they are would seem to lead to the result that whenever we believe something false about an object we lose the ability to refer to it. For example, suppose we are out in a crowded place and, thinking I catch sight of your brother, confidently point and say 'That's your brother!' Since I intend to say something true, don't I intend the referent of my use of 'that' to be your brother? But if it turns out to simply be a lookalike, we don't want to claim that reference has failed; such cases are perfectly ordinary. And in this example it seems *obvious* that my intention to have my use of 'that' refer to the person I see is the ultimately reference-fixing or 'primary' intention.

This last observation suggests another potential recourse for the simple intentionalist. One could say that in both the Carnap-Agnew case and the brother example just described, the 'secondary' intentions deriving from the speaker's (false) beliefs are semantically irrelevant. But in the Carnap-Agnew case, the secondary intention happens to also be a *communicative* intention, and so it is more easily confused with the primary referential intention. This fits with what was argued in the previous chapter: that theorists should be wary of overstating the relevance of communicative success for fixing semantic content. And in fact a number of theorists (e.g. Perry 2009; King 2013) have proposed that we make use of the relationship between beliefs and intentions to solve the problem. Their idea is to analyse a referential intention as a kind of *plan*, from which multiple intentions can be derived based on their role in achieving the plan.

A typical foundation for this kind of theory is Bratman (1987), whose theory of intentions views them as hierarchically structured entities connected via a process of *means-end reasoning*. The inputs of means-end reasoning are intentions and other mental states, the outputs subsidiary intentions that help to fulfill those further up the chain. For example, in the Carnap/Agnew scenario, reasoning from my intention to have my use of 'that' refer to the picture of Carnap and my belief that the picture of Carnap is behind me lead to my intention to have my use of 'that' refer to the picture behind me. On the view King (2013) supports, the primary intention is that lying at the end of this kind of chain, the final step in executing in the plan.

While this seems promising, Speaks (2017) points out that it is not readily generalisable without saying more about what characterises the secondary intentions. For the hierarchical structure we choose seems to presuppose that a

---

particular intention is primary and *then* gives the 'right' account of the reasoning process ending in that intention. This may not be so clear in the examples we have seen so far, but I think we can modify the Carnap-Agnew example to make it harder to decide which intention lies where in the hierarchy[3].

Imagine that we are playing a game in front of the wall of pictures - I walk along with my back to the wall and try to guess what each one depicts, pointing behind me as I go. 'That is a picture of a Founding Father', I say, then 'That is a picture of a two-time Nobel prize winner'. However, unbeknownst to you, I glanced through the window of the same room last week and saw that the third picture along was one of Rudolf Carnap. Unbeknownst to *me*, it has since been replaced with one of Spiro Agnew. Now, when I move to next picture, point behind me and say 'That is a picture of one of the greatest philosophers of the twentieth century', it seems that there are two ways of describing my process of means-end reasoning. On the first, my wider referential plan is to refer to the picture of Carnap (since that leads to my winning the game), and combined with my belief that the picture of Carnap is behind me this gives rise to my intention to refer to the picture behind me. On the second, my plan is to refer to the picture behind me (since that is the point of the game), and combined with the same belief this gives rise to my intention to refer to the picture of Carnap. To decide between these explanations, we need an independent justification for saying that a particular intention is at the end of the chain of reasoning. I think we can find such a justification in Perry's (2009) work on "directing intentions".

Previously I named Åkerman (2009) and Perry (2009) as proponents of simple intentionalism. In fact Perry goes further than defending SI in general, developing a theory that gives some details about the content of a referential intention. He aims to develop and improve upon Kaplan's (1989b) minimally described directing intentions theory. On that model, a demonstrative is supplemented by a directing intention δ to give its complete character $that[\delta]$. The referent of a use of $that[\delta]$ is then simply the object of δ. But Kaplan's notion of a directing intention was essentially tied to demonstrations and only applicable to "perceptual" demonstratives: it is defined as "the speaker's intention to point at a perceived individual on whom he has focused" Kaplan (1989b, p. 582).

It is clear, however, that demonstratives can be used without (actual or merely intended) demonstrations. Perry also points out that one need not be *perceptually* focused on an object to have one's use a demonstrative refer to it, but one must have *some* kind of "connection" to or way of thinking about it. In non-perceptual cases, this connection might be via memory or definite description. Moreover, Kaplan's formulation of the directing intentions theory does not decide between possible referents in case of conflicting intentions. Hence to remove the limits of Kaplan's theory and extend its scope, Perry suggest that we identify δ not with the directing intention itself, but with the *type* of directing intention, where types are distinguished according to the connection they exploit. It is acknowledged that the choice of connection influenced by the speaker's perception of her audience (Perry 2009, p. 196), but this ought not

---

[3]See also Speaks' (2017, p. 729) 'sleeping student' example in the same vein.

to be confused with an endorsement of something like reflexive intentionalism. Perry is resolutely against theories that allow semantic content to be influenced by an audience's ability to figure out what that content is: "we need a concept very close to what is said, but one that seals off the forensic issues" (2009, p. 192).

I think that Perry's 'connections' provide an independent metric for determining primary referential intentions, in the sense of intentions at the bottom of hierarchically structured plans of action. They are those intentions that are active at the time of utterance, that concern the object in the way that one has in mind upon uttering the relevant demonstrative. In the modified Carnap/Agnew case then, the primary intention is dependent upon the speaker's mental state at the time of utterance. If she, upon speaking, thought of the picture of Carnap *qua* the picture and intended her demonstrative to refer to *that*, then the picture of Carnap was the referent. But (and this seems like the more likely scenario given the nature of the game), if she thought of the intended referent as the picture behind her and merely believed that it also was the picture of Carnap, the referent was the picture behind her.

Skeptics may say that this conclusion is insufficiently 'objective', as it is ultimately all dependent on a fleeting and elusive mental state of the speaker. I would respond that if we are willing to grant simple intentionalism, we are already in the realm of the fleeting and elusive, and that it is hardly surprising that its development should remain grounded in speakers' mental states. And in fact the relevant states may not be so elusive; in my experience using and understanding demonstratives, speakers are often readily prepared to swap them out for definite descriptions in cases of misunderstanding. The following conversations sound perfectly natural, for example:

(5) A: Look, that's your brother!
     B: *(looking around)* Who?
     A: The guy standing by the postbox.

(6) A: That is a picture of a Founding Father.
     B: Which one?
     A: The one I'm pointing at, look!

I take these descriptions to be verbalisations of something like Perry's 'connections', and the fact that they come so readily to speakers to support the claim that there is indeed a singular way one is thinking about an objection when one forms an intention to refer to it. I only hesitate to base the theory itself on these hypothetical scenarios, i.e. claim that willingness to swap out a particular definite description for one's use of a demonstrative fixes which intention is ultimately referential, to allow for the possibilities that speakers may not be able to literally verbalise an appropriate description or may want to for some reason obscure their referential intentions. Nonetheless, one should not be put off but rather encouraged by the fact that Perry's theory gives complete control over to the speaker's particular type of connection; it would seem that we are ordinarily happy to privilege it.

3.2 Communicative and referential intentions

To reiterate Bach's challenge (e.g. 2017): so-called 'referential' intentions are actually a part of one's overall communicative intentions. It was argued in the last chapter, however, that on certain occasions speakers have (simple) referential intentions but discordant or absent communicative intentions. Still, it does not immediately follow that referential intentions cannot *ordinarily* be reduced to communicative intentions. Cases of content-communication divergence are certainly atypical; perhaps (Bach might argue) we should not take them so seriously. We do find an alternative motivation for the existence of separate referential intentions in Viebahn (2018), but I think it does not quite address Bach's concerns. I first explain his account then argue for a replacement.

Viebahn's idea is not to deny that referential intentions normally result from communicative intentions, just to deny that the former supervene on the latter. Instead, we suppose that one starts with a communicative intention and achieves it by way of fulfilling what Viebahn calls an "encoding intention", defined as an intention "to encode a message in a string of expressions" (2018, p. 94)[4]. Encoding a message like this can be a way to make clear your intention to communicate that message to your audience, and this fact is what provides the relationship to communicative intentions. In the specific case of demonstratives, one could say that the intention to speaker-refer *with* a demonstrative can be fulfilled via recognition of one's having already fulfilled an intention to have the demonstrative refer (cf. Schiffer and Neale forthcoming). This answers Bach's questions about the content of each intention: they must have distinct content because (the communicative) one presupposes the fulfilment of the other, but not the other way around. What's more, we again we see a preference for Bratman's theory; Viebahn supposes that this move from communicative to encoding intentions happens via a process of means-end reasoning. It also straightforwardly covers the cases of content-communication divergence discussed in the previous chapter; one can have an encoding intention without a corresponding communicative one (or any communicative intention at all).

So Viebahn's purports to have a simple answer to Bach's doubts about referential intentions. However, I think his response does not quite get to the heart of Bach's criticism. Viebahn has shown that if speakers have referential intentions concerning demonstratives, those *could* have distinct content from their communicative intentions concerning demonstratives. But we haven't explained why we should believe that we have separate referential intentions in the first place[5]. I take this question to be separate to the question of what semantic work referential intentions do that communicative intentions do not. We have an answer for that: they fix reference in cases of content-communication divergence.

---

[4]In addition to the inadequacies of this explanation that are presently described, one might be wary of its commitment to a particular kind of mentalism: it assumes that propositional messages arrive in speaker's heads ready to be encoded into linguistic strings.

[5]Viebahn (2018) actually does attempt to do this: he claims that referential intentions are necessary to fix the semantic values of polysemous expressions. Since I want to stick to discussion of demonstratives, though, and given that demonstratives are not polysemous expressions, I think it's worthwhile to seek an alternative account.

But there is a separate question, also raised by O'Rourke (2019), regarding what work do they do *in their role as intentions* that communicative intentions do not. I think this is at least partially what Bach has in mind when he claims it is "implausible" to suppose that we have referential intentions in addition to communicative ones. If non-communicative referential intentions exist, there should be evidence of that fact.

To elaborate: whatever your theory of intentions, it is accepted that they are themselves *intentional*, i.e. aimed at some state of affairs. On Bratman's theory, they are further specified as "conduct-controlling pro-attitudes" (1987, p. 52), meaning that they commit us to taking particular actions in particular situations. (Alternate theories may take a weaker position when it comes to commitment, but none would deny that intentions at least *influence* action.) Now, in the case of communicative intentions, the state of affairs at which they are aimed is familiar, e.g. for the Gricean it is that a reponse is induced in your audience by way of their recognition of your intention. And the conduct-controlling outcomes of such an intention are clear: they lead us to make our intention recognisable, e.g. with pointing gestures. In the case of referential intentions, there is a state of affairs to aim at: a simple referential intention to have one's use of a demonstrative *d* refer to an object *o* is fulfilled iff *d* refers to *o*. But can we point to any particular behaviour as evidence of the controlling influence of this kind of intention? If so, it would give independent reasons to suggest that referential intentions play a role in our use of language that communicative intentions do not.

We have one immediate example from our conclusions about the problem of conflicting intentions in the last section: referential intentions guide us to think about an object in a particular way upon uttering the demonstrative, so that the reference is fixed by exploiting that connection. But this doesn't seem to be a fact exclusive to referential intentions. That is to say, it's likely that speakers also intend to exploit ways of thinking about an object when merely speaker-referring to it.

Another possibility is to look to examples of content-communication divergence not as evidence of the semantic significance of referential intentions, but of their conduct-controlling properties. This brings us closer to our goal, but needs some refinement. For speakers do not, say, talk aloud to themselves *to achieve the end* of having a particular demonstrative refer. Rather, speakers talk aloud to themselves and *somewhere along the way* both form and fulfill the intention to have a demonstrative refer. What part of the process does this referential intention control?

I think the reason this question is so difficult to answer is because simple referential intentions are stipulated not to be intentions to *do* something, but intentions to *have something happen*. According to Viebahn, one encodes a message in a string of expressions with the intention that that encoding expresses a proposition equivalent to the message. But it is not *you* that does the proposition-expressing, it is the choice of encoding. Perhaps, then, we can say that referential intentions guide us to use combinations of words that express the propositions we want to express (see Viebahn 2018, p. 14). This might

work at the sentence/proposition level, but cannot be the case at the demonstrative/object level. Else we would be trying to fulfill our intention to have a demonstrative refer to a particular object by using a demonstrative that refers that object. And we have supposed that a demonstrative refers to a particular object iff we intend it to refer to that object. So the only conduct-controlling result of such an intention would be that it guides us to fulfil it, which is explanatorily redundant.

Thus perhaps it is not best to look at the different conduct-controlling properties of referential and communicative intentions in order to motivate their separate existence. Instead, I think the intentionalist should say that the best evidence for referential intentions is their *lack* of conduct-controlling properties, at least when it comes to future conduct. And in particular, I think referential intentions ought to be explained as what Searle (1983) calls *intentions in action*.

Searle's theory views intentions in general as "causally self-referential" (p. 85); it is part of their conditions of satisfaction that they play a part in effecting the relevant state of affairs. In the case of *prior intentions*, i.e. those formed in advance of the action that fulfills them, this notion is cashed out rather like Bratman's "conduct-controlling pro-attitudes"; prior intentions are satisfied iff they cause us to perform an action that brings about their satisfaction. These stand in contrast to intentions in actions, which are spontaneously formed and fulfilled without prior existence of or deliberation over the intention. (Searle's paradigmatic examples are bodily movements, like raising one's arm to scratch an itch). Prior intentions have satisfaction conditions like this:

> (I perform the action of raising my arm by way of carrying out this intention.) (Searle 1983, p. 92)

Intentions in action, on the other hand, have satisfaction conditions like this:

> (My arm goes up by way of carrying out this intention in action.) (1983, p. 93)

I hope the resemblance between this latter condition and the form of a simple referential intention is clear. In both, there need be no consideration of one's own experience in bringing about the intention's fulfillment; it is simply fulfilled by the outcome of that role. For Searle, this is a feature of the fact that intentions in action have *presentational* content - a particular type of representational content that obtains when the conditions of its veridicality are directly presented to the agent[6] But for our purposes it is enough to observe 1) that the satisfaction conditions of an intention in action are in a way directly given by the action itself, rather than being first conceived of and then aimed at (as with a prior intention) and 2) these satisfaction conditions concern the outcome of the action rather than the experience of bringing about that action *and* its outcome. Or, as Searle puts 2), "the intentional content of the experience of acting and the intention in action are identical" (1983, p. 93).

---

[6] Searle says that visual experience is a paradigmatic example of an experience with presentational content.

My point in introducing Searle's framework is that if we think of referential intentions as intentions in action, we see why communicative intentions cannot do the same job. The evidence for the plausibility of referential intentions shouldn't be ostensible semantic reference. The evidence for the plausibility of referential intentions should be semantic reference that ostensibly results from intentions, but is spontaneously achieved without concern for one's own future experience of controlling the outcome. This simply doesn't fit with standard ideas of communicative intentions, which involve the intention to induce a belief in an audience via our own reasoned efforts.

For illustration, we can turn to Schiffer and Neale's (forthcoming) quasi-intentional account of demonstrative reference as grounded in communicative or speaker-referential intentions. First they give a Gricean definition of speaker reference[7]:

> S referred to x (in the course of) uttering u iff in uttering u, S meant an x-dependent proposition. (p. 4)

This leads to a definition of referring-with an expression,

> In uttering u, S referred to x with e, relative to its i-th occurrence in u, iff for some person A and property φ, S intended it to be common ground between S and A that the i-th occurrence of e has φ and, at least partly on that basis, that S referred to x in uttering u. (p. 6)

and ultimately a theory of semantic reference for demonstratives:

> x is the referent of demonstrative 'that' ['this', 'it'] on occasion o iff the speaker referred to x with 'that' ['this', 'it'] on o. (p. 10)

Given Schiffer and Neale's definition of referring-with, the intentions underlying an utterance of a demonstrative are only meaningful in terms of their role in the future outcome of having one's overall communicative intention fulfilled. But this does not explain what intention is spontaneously formed and fulfilled in the very moment of uttering the relevant demonstrative, which is not a product of prior deliberation.

We can compare referential intentions-in-action with those underling bodily movements. One might, say, reach out one's arm *en route* to a number of differing end goals; perhaps to stop someone else in their tracks, to try out a new dance move, or simply to see how long one's arm has gotten. But nonetheless these cases all involve instances of the same action (and thus intention-in-action), which has nothing to do with its role in a larger outcome. Different uses of 'that' referring to the same object, I think, should be thought of in a similar way.

Nonetheless, there is no use denying that communicative and referential intentions are closely linked. Searle's view was that prior intentions generate intentions in action when the time to act comes; perhaps we should think of Schiffer

---

[7]Actually, 'primary' speaker reference. See Chapter 1, fn. 8.

and Neale's referring-with intentions as giving rise to semantic-referential intentions as the relevant utterance is carried out. This would certainly go some way to explaining why the two concepts appear so closely intertwined, to the point that some see them as identical.

3.3   Circularity

The circularity objection to intentionalism was first raised by Gauker:

> Some thought must be given to the content of the pertinent speaker intention. In speaking of the case of demonstratives, I have described it as the speaker's intention to *refer* to something. But what is the relation we call "reference" in thus characterising the speaker's intention? Either it is the very same reference relation that we seek to explicate in terms of speaker's intention, in which case the theory is circular, or it is some other relation, in which case we require some other theory of demonstrative reference in addition to the one we are expounding. (2008, p. 362)

The accusation is repeated in more detail in Gauker (2019) and touched upon by Bach (2017, fn. 13), who remarks that intentionalism "smacks of circularity or perhaps even an infinite regress".

Before addressing this kind of objection, it's worth investigating to what extent it really challenges intentionalism. The underlying presumption, as Gauker admits in the excerpt above, is that the theory somehow attempts to explain what reference *is*, or what it *means* to say that a use of a demonstrative d refers to an object o. Lewerentz and Marschall (2018) deny this presumption: intentionalism is intended as a *metasemantic* theory, where a metasemantic theory is one that aims to explain what makes it the case that certain semantic facts hold. Hence it cannot be that the theory is also tasked with explaining what it means for those semantic facts to hold:

> In order to develop a metasemantics for demonstratives, we need to start with facts such as the fact that in a particular situation, demonstrative d referred to o. We then go on to ask: 'What makes it the case that this fact holds?' and try to give an answer. To even state the question, we already need to know what it means for an expression to refer to something. (2018, p. 1671)

Once it is understood as a metasemantic theory, intentionalism need not worry about making use of the concept of reference in its explanation because it still provides new information about the mechanisms underlying semantic facts, if not the facts themselves. The problem with their response, however, is that becomes somewhat pointless in the light of their second possible interpretation of the circularity objection, which serves as a challenge to intentionalism whether it is taken as a metasemantic theory or not. That second interpretation, which

Gauker (2019) clarifies to be his actual concern, claims that intentionalism entails a circular *requirement*. As Lewerentz and Marshall put it,

> For a speaker to be acquainted with the reference relation, this relation surely has to be instantiated somewhere. But the relation is instantiated only if someone has a referential intention, and the latter requires the relevant acquaintance! (2018, 1674)

It may be that this is also what Bach is hinting at with his mention of an "infinite regress"; we need some way of independently specifying the satisfaction conditions of a referential intention, else there is no way to explain its initial formation.

Lewerentz and Marschall's response to this version of the circularity objection is to argue that speakers are familiar with the concept of reference-in-general from their encounters with other directly referential terms (e.g., possibly, proper names), and this acquaintance gives rise to the concept that figures in their referential intention. But, as Gauker (2019, p. 118) points out, given the wildly different referential behaviour of proper names and demonstratives, it is not clear that the former could provide a good model for understanding the other.

The intentionalist need not resort to this explanation to avoid circularity, however. For given what was argued for in §3.2, it is the very nature of referential intentions as intentions in action that their conditions of satisfaction present themselves as the resulting action is executed. Thus there is no need to explain the content of a referential intention in terms of other intentional states, such as recognition of the the referential nature of proper names. That is not to say, however, that the intention comes out of nowhere. In particular, it would be ridiculous to say that we are *born* competent users of demonstratives. So there is still something to the circularity objection: it challenges the intentionalist to explain what the basis of this ability to form and satisfy referential intentions is. The circularity is merely shifted to the level of the action; we must now explain how the first referential intention-in-action might have come about.

Chapter 4

The nature of referential intentions

4.1 Circularity revisited

Although I showed in the last chapter that the circularity objection in its typically expressed form need not be fatal for intentionalism, it raises questions about the mental capacities underlying our supposed referential intentions that cannot be ignored. For example, Speaks (2016, p. 304) points out that young children are competent users of demonstratives but are unlikely to have sophisticated metasemantic intentions:

> ...[intentionalism] requires speakers to have intentions about objects being the semantic values of expression tokens. And while I have intentions of this sort, it is at least not obvious that my two-year-old daughter, who is a competent user of demonstratives, does. (2016, p. 304)

There is little discussion of this issue in either the intentionalist or anti-intentionalist literature (Speaks flags it as cause for concern then promptly moves on), with the exception being Viebahn (2018). Viebahn's idea is that while young children might not have referential intentions *per se*, they may be able to form equivalent intentions that are good enough to enable communicative success, i.e. "deferential intentions" to use an utterance with the same meaning as another (older, more competent) speaker has previously used it. Thus we can explain young children's apparent competence with demonstratives:

> ...consider a case in which I decide to tell my son that a certain plant on our terrace is a sunflower, form an encoding intention and (while pointing at the plant in question) utter, 'That is a sunflower.' Then my son might utter the same sentence to tell someone else that the plant in question is a sunflower. Here, a deferential intention seems

> to do the job of securing the sunflower in question as semantic value
> of my son's use of 'that'. (2018, p. 18)

I'm not sure, however, that this proliferation of intentions is the right defence strategy for the intentionalist. Its most glaring flaw is that it fails to account for the full range of demonstrative competence in young children. Research into early demonstrative use, which begins at age one, reports spontaneous use rather than simple mimicry of adult use (e.g. Clark 1978). Moreover, unless we restrict deferential intentions to operation at the sentence level, it is not clear that they are even less complex than simple referential intentions; an intention to use a demonstrative 'with the same meaning' as another demonstrative still seems to presuppose a certain metasemantic awareness.

On a related note, it's worth noting that the problem at hand would seem to apply to any theory that invokes anything other than very basic intentions. The Gricean theory of meaning, for example, rests on the claim that speakers have intentions to have their intentions satisfied by way of their own recognition, and young children appear to be successful communicators. Does it follow that they are aware of the concept of intention satisfaction? It seems to me that there is something wrong in the formulation of the question: it largely depends on how we spell out the notion of "awareness". The objection only seems to have force if we spell it out in a particularly strong manner, namely requiring that the agent in question knows the concept by name, or can describe the conditions under which it holds. But such strict standards would suggest that young children have a more limited range of possible intentions than their behaviour shows. Gricean communicative intentions, for example, would be excluded.

But even more damning for the circularity objection is that it doesn't seem to apply to intentions in actions at all. Consider, for example, Searle's model intentions in action, i.e. those underlying bodily movements. By 6 months of age the average child has passed many motor milestones — learning to roll over, reach for objects, and sit up unsupported (Bayley 1995) — but has only just begun to understand basic words (Bergelson and Swingley 2012). If we accept Searle's theory, this discrepancy makes sense because the kind of representation that enables intentions in action is tied to the action itself rather than any grasp of an associated concept. The content of the intention causing such a movement is 'presented' by the movement itself. Thus there is no point asking whether one can conceive of, say, rolling over before one does it; doing it is sufficient to acquire the conception.

This implies that whether it is plausible to attribute referential intentions to young children depends on their plausible level of competence or familiarity with demonstrative reference, rather than their being able to conceive of the relevant intention's content in a particular way (i.e. as being about the referents of expressions). Equally, however, intentionalism tells us that competence with demonstrative reference just *is* the ability to have referential intentions in the first place. So we are trapped in another version of the circularity objection, left to explain an initial action that becomes the basis or model for future reference-events. To resolve it, we might turn to examining empirical evidence from the

study of language acquisition.

4.2   Early demonstrative use

Young children's competence with demonstratives, as anecdotally considered by Speaks, is not only borne out in the relevant literature but considered a hallmark of language acquisition. Clark (1978) reports that they are often amongst the ten first words infants produce, and their arrival appears to grow out of competence with pointing gestures. This finding give us a starting point to investigate the development of abilities underlying demonstrative use: their commonality with proto-communicative actions.

On the basis of their early appearance, their relationship to pointing gestures, their universal existence across languages and their etymological originality[1], Holger Diessel (2006; 2013) has argued that demonstratives deserve to be treated as a unique grammatical class serving a unique function: they serve to coordinate the interlocutors' shared attentional focus, or establish *joint attention*. Importantly, joint attention must involve mutual recognition on the part of each agent that attention is coordinated; this distinguishes it from mere shared attention. We can think of joint attention as a three-place relation between oneself, another agent, and an object and, if we follow Diessel, track its development from gaze-following to pointing to use of demonstrative expressions[2]. The use of *deictic* pointing in particular appears to set us apart from non-human primates, who have been observed pointing in captivity but tend to do so *imperatively*, i.e. to request objects or actions from humans (Liebal and Call 2012). In Diessel's view, this reflects that they recognise a cause-effect relationship between the gesture and the desired result, but never develop an understanding of its dependence on others' psychological states. For humans, however, this understanding is reflected by our use of pointing to establish joint attention.

Much is known about the emergence and development of joint attention: as documented in the work of George Butterworth, infants start to follow their caregivers' gazes by 6 months of age and reliably shift their attention in response to both gazes and pointing by 18 months (Butterworth 1995; Butterworth and Grover 1988). Theorists are more divided, however, when it comes to identifying the faculties enabling joint attention. Baron-Cohen (2005), for example, posits the existence of an innate specialised "shared attention mechanism". Tomasello et al.'s (2005) constructivist model argues that young children first recognise the relationship between their own intentional mental states and actions, then extrapolate from that knowledge to understand others'.

---

[1] That is to say, the development of demonstratives cannot be traced back to any particular content word(s). In this way they stand in contrast to other closed-class or 'function' words (see Diessel 2006, p. 474-5).

[2] In a related but not entirely relevant point, some theorists have claimed that joint attention is the fundamental mechanism underlying language acquisition altogether. Tomasello (e.g 1995; 2003), for instance, argues that joint attention plays a critical role in early word learning.

For our purposes, however, it is enough to note that understanding joint attention appears to be a prerequisite for competence with demonstratives. And the aforementioned theories do have a common thread: they understand joint attention as closely tied to the development of a full-fledged theory of mind, or the ability to attribute propositional attitudes to others[3]. Hence they maintain that joint attention presupposes an understanding of others as agents with their own attention and intentions. The *attention* part of this claim explains how infants view gaze as reflecting others' perceptual focus, the *intention* part how they grasp that others want them to share in that focus. How basic this understanding might be is also a matter of debate; Tomasello (1995), for example, argues that the progression of joint attentional capabilities from gaze-following to manipulating others' attention via linguistic expressions (amongst other activities) reflects the development of an increasingly sophisticated intention-specific theory of mind, moving from the recognition that others have attention/intentions to the understanding that those may differ from one's own. But it suffices to say here that some kind of recognition of others' intentional capacities is necessary to participate in joint attention, an understanding of which in turn enables the use of pointing and ultimately demonstratives.

4.3   Joint-attention simple intentionalism

Suppose, then, that we take demonstrative reference to develop out of more basic joint-attention establishing actions like pointing. This explains how infants can begin to form referential intentions. But could we also extend this solution to address Gauker's original circularity objection? That is, could we suppose that referential intentions amount to intentions to establish joint attention beyond the acquisition period? This would mean endorsing what we might call JA-SI:

> *Joint-attention simple intentionalism (JA-SI)*
> The referent of a use of a demonstrative d is an object o iff the speaker intends d to establish o as the subject of the interlocutors' joint attention.

I think JA-SI is well-posed to answer Gauker's concerns, although this claim does not come without more argument. In his 2019, Gauker considers a number of ways that the intentionalist could avoid circularity, namely by substituting 'reference' in the content of the intention with some other notion. Most relevantly, he considers what might happen if we substitute 'drawing attention to' for 'reference':

> So we might say: a speaker refers to o in uttering 'that' only if the speaker intends *to draw the hearer's attention to o* by uttering

---

[3]On some understandings, this simply constitutes "theory of mind". But I prefer to reserve that term for a general, basic ability to recognise that others have intentional mental states that may differ from one's own, rather than the "full-fledged" theory of mind that appears to emerge at around 4 years of age and is reflected by children's performance on Wimmer and Perner's (1983) False Belief Task.

'that'[4]. The trouble is that the answer is subject to clear counterexamples. The speaker may very well know that the hearer's attention is already focused on the referent of 'that' and intend to draw the hearer's attention to something else. Suppose that A is holding a knife and looking at it as if to decide whether it is suitable as a weapon. But B sees that there is a big man with a gun coming round the corner; the knife will not protect A from him; it would be better for A to run. So B says, "That will not help you against the big guy over there with a gun." So the speaker intends to draw the hearer's attention to the big man with a gun by uttering a sentence in which 'that' refers to the knife that the hearer is holding. The referent is not what the speaker intends to draw the hearer's attention to.

Before considering its implications for the topic at hand, I want to note this argument as Gauker formulates it is rather uncharitable towards intentionalism. For it presumes that the relevant intention operates at the sentence level, rather than governing the speaker's use of 'that' specifically. I think it is plausible to say that B has an intention for A's attention to be on the big man with a gun by the end of the utterance, but also an intention concerning the use of 'that' specifically to have the A's attention bypass the knife before reaching its final target. This consideration alone does not dissolve the objection, however; it is still difficult to explain why B would have an intention to have A focus her attention on the knife when it is stipulated that A's intention is already so focused.

Let's consider what changes if we substitute 'drawing attention to' for 'establishing joint attention on'. The result is promising; although A's attention is already fixed on the knife at the beginning of the utterance, B's is not and hence 'that' can still be employed to establish its place in the *shared* attentional focus. There is nothing in the definition of establishment of joint focus that precludes this formulation; what matters is that focus is *coordinated*, whether it is the speaker or the hearer whose focus is first fixed on the relevant object. In response to this argument, one might try to modify the example so that both A and B are already contemplating the knife. But at some point B's attention must switch to the man with a gun (e.g. if B remembers that he is coming around the corner), since she forms an intention to alert A to his existence, so it could still be argued that B's use of 'that' serves to re-coordinate joint attention on the knife. So I think the current account is well-equipped to deal with this example in particular.

Nonetheless, the spirit of Gauker's objection remains. He believes that attempts to substitute reference with *any* other concept will end up leaving the theory subject to counterexamples, thus intentionalism has no choice but to entail a circular requirement (2019, p.116). Indeed, it is not too difficult to think of counterexamples to the idea that referential intentions are joint-attention-establishing intentions. And these examples are *not* restricted to cases

---

[4]Gauker equates speaker reference and semantic reference in the case of demonstratives.

of content-communication divergence, i.e. those that I have set aside for now. For example, it is not at all uncommon to use demonstratives even when joint focus is already established and the interlocutors are aware that it is established. Suppose that we are at the Grand Canyon, looking out on the view together, and I say 'That is the most beautiful thing I've ever seen'. It cannot be that my utterance of 'that' functions to establish shared focus on the view, as we are both already attending to it.

Considering this example in the light of Diessel's view, he does specify that demonstratives only function to initiate joint focus "in the simplest case" (2006, p. 470). But the only other possible uses considered are "to direct the addressee's attention from the current referent to a previously established referent or to differentiate between multiple referents that are already in the shared attentional focus" (ibid). These aren't quite what we want. But perhaps we should pay closer attention to Diessel's way of using 'referent' here. It implies that whatever is already established as being in the shared attentional focus becomes the default 'referent', even if nothing has yet been said. Thus we could say that the Grand Canyon case illustrates something like an *anaphoric* use of 'that', sharing a previously supplied referent. As in other cases of anaphora, this takes the matter of reference-determination out of the speaker's hands. Hence as long as we maintain that anaphoric uses are a special case, it does not preclude that referential intentions are joint-attention-establishment intentions.

No other clear counterexamples immediately come to mind. Although it would be premature to conclude that none exist, I think for now we are justified in advocating JA-SI as a viable solution to the intentionalist's circularity problems. The next chapter is dedicated to its other virtues, particularly in comparison to alternative accounts of demonstrative reference.

Chapter 5

JA-SI in comparison

This chapter looks at how JA-SI compares to other contemporary theories of demonstrative reference. §5.1 discusses my previous comments on other intentionalist proposals and how the development of JA-SI sheds new light on them. §5.2 shows that JA-SI is not only well-equipped to compete with intentionalist alternatives, but also with contemporary salientist theories. These accounts, which have heretofore been only briefly mentioned, claim that demonstrative reference is fixed not by speaker intentions but by some measure of what object is most salient in the context. As we shall see, JA-SI both eliminates the motivation for proposing this type of account and outstrips it in terms of the data that can be explained. Finally, in §5.3 I compare JA-SI to the quasi-intentionalist position, which views demonstratives as tools for Gricean speaker reference rather than as having their own semantic-referential power. I conclude that although JA-SI is a better general theory of demonstrative reference (because it does not depend on communication), this does not exclude the possibility of combining the two theories to explain communicative cases.

5.1   Intentionalists

5.1.1   Reflexive intentionalism

To restate the definition of reflexive intentionalism introduced in Chapter 2:

> *Reflexive intentionalism (RI)*
> The referent of a use of a demonstrative d is an object o iff the speaker intends:
>
>   1. o to be the referent
>   2. her audience to recognise her intention (1)

I argued previously that reflexive intentionalism ought to be discounted for its failure to deal with cases of content-communication divergence. These come in

two types: cases of self-talk, and cases in which a speaker has no intention for her audience to recognise her (simple) referential intention. The latter type was further divided into Åkerman's (2015) incapable-audience cases and Michaelson's (2013) "sneaky reference" cases.

In contrast, JA-SI can deal with these cases much naturally. One always shares one's own attentional focus, so a self-talking speaker need do nothing to have her use of a demonstrative refer to an object other than focus her attention on that object. As for the second type of content-communication divergence case, I think it can be dealt with by delving more deeply into what counts as an interlocutor for the purposes of establishing joint attention. In particular, I think an interlocutor ought to at least understand that they are a part of the conversation. This takes care of Åkerman's incapable audience cases, as occur when speakers talk to (for instance) infants or dead people. They are simply reduced to cases of self-talk, as the 'audience' are not real interlocutors at all.

In cases of sneaky reference, this explanation is slightly more strained. Michaelson's (2013) example was as follows:

> You and I both work for a large investment bank. We've just been paid an obscenely large bonus, and we are both at the office very early one day. We run into each other near a window overlooking the parking lot, which is nearly empty at this hour. There is, however, a small cluster of cars which includes several nondescript vehicles and, in addition, a McLaren F1. You ask me what I decided to do with my bonus, and I respond by pointing out the window at the cluster of cars and saying: "I bought that." In fact, I bought the Toyota Camry parked next to the McLaren, and I intend to refer to the Camry, not the McLaren. That said, I've decided to have a bit of fun with you. Even in a situation like this one, I take it, I have plausibly succeeded in referring to my Camry. (2013, p. 91-2)

I suggested in Chapter 2 that this is a situation in which the speaker's has a genuine referential intention but a misleading communicative one. But can a referential intention concerning the Camry really be considered 'genuine' if it is taken as an intention to establish joint attention? Only if we suppose that the hearer doesn't count as an interlocutor for referential purposes, but does for communicative ones. While this is implausible on its face, I think it should be taken seriously as a possibility. After all, it does give us a new way of spelling out the speaker's misleading communicative intention, namely that its planned operation exploits the hearer's mistaken assumption that she is a referential-interlocutor. And it does reflect the intuition that the speaker is in one sense excluding the hearer from the act of reference.

That intuition might be clearer if we consider a similar case in which the speaker does make a pointing gesture (towards the Camry), but deliberately obscures it from the hearer. At least in my eyes, the hearer's frustration in that scenario would arise from the knowledge that *something* is referred to, but she is not privy to *what*. This cannot be explained as a failed reflexive-referential intention; clearly the speaker never intended her referential intention

to be recognisable to her audience. So the best a reflexive-referential account can offer us is reference failure. At least in my eyes, the Camry/MacLaren case will always be a strange one. But I think that in this regard JA-SI does a better job than RI of explaining *why* it is strange.

However, I also argued in Chapter 2 that reflexive intentionalism ought to be criticised for mistaking communicative considerations for semantic ones. Reflexive intentionalists are motivated to add a recognisability criterion because of supposed cases of Humpty Dumptyism, wherein speakers are afforded implausible control over the meanings of the expressions they use. One striking feature of JA-SI that might initially seem contradictory to that argument is the primacy that it gives to the speaker's audience. At first this suggests that we owe the reflexive intentionalist an apology, but in fact it ultimately proves the redundancy of a recognisability criteria in another way. For if JA-SI is correct, then RI becomes:

> *Joint-attention reflexive intentionalism (JA-RI)*
> The referent of a use of a demonstrative *d* is an object *o* iff the speaker intends:
>
> 1. *o* to be established as the subject of the interlocutors' joint attention
> 2. her audience to recognise her intention (1)

With this reformulation, we can see that reflexive intentionalism need not be redeemed. Because establishment of joint attention involves the audience by definition, the second criterion is made redundant. Part of establishing joint attention with an audience is ensuring that they recognise an attention-establishing event is taking place. Thus JA-RI is *equivalent* to JA-SI.

So we have reached the same conclusion I argued for in Chapter 2—that adding a second criterion to SI is redundant—but via a different route. In that chapter I argued that moving to RI is unnecessary because appeals to a speaker's communicative intentions have the same explanatory power as the second condition. Here I argue that it is unnecessary because the second condition is essentially integrated into the first. This gives more credence to the claim that recognisability criteria are unnecessary to avoid Humpty Dumptyism. In spelling out simple intentionalism as JA-SI, we maintain that speaker's have enough control over their uses of demonstratives to act as they do in cases of content-communication divergence. But we also ensure that there is some constraint on what counts as a genuine referential intention. This gives another perspective on cases like King's 'Venice Beach' example:

> I am sitting on Venice Beach on a crowded holiday looking south, with swarms of people in sight. I fix my attention on a woman in the distance and, intending to talk about her and gesturing vaguely to the south, say "She is athletic". You, of course, have no idea who I am talking about. It seems quite implausible in such a case to say that I succeeded in securing the woman in question as the value of

my demonstrative simply because I was perceiving her, intending to talk about her and gesturing vaguely in her direction. (King 2014, p. 223)

I claimed in Chapter 2 that this case is confusing because the speaker is afforded no plausible communicative intention, and thus it is hard to trust our seemingly semantic intuitions about it. But the current account gives another reason to think so. The case is confusing because 'you' are stipulated as an interlocutor. But the speaker clearly does not have a genuine intention to draw your attention to the woman in question. All RI can say about this case is that it is one of reference failure. But JA-SI can say that it is *either* a case of reference failure or a case of self-talk, so that the speaker's attention is all that matters for fixing reference. This respects both ways in which one's intuition might go when confronting with such a confusing case.

5.1.2 Perry's directing intentions theory

Let's return to Kaplan's original comments on a "directing intention": "the speaker's intention to point at a perceived individual on whom he has focused" (1989b, p. 582). Contemporary intentionalism, in supposing that a referential intention concerns semantic values of demonstratives, at least threatens to move away from this conception. JA-SI seems like a promising route back. Given the close connection between pointing and demonstratives, as discussed from an acquisition perspective in §4.2, we can preserve Kaplan's idea that referential intentions somehow guide their associated demonstrations (although the separation is now enough to allow for cases of reference without demonstration).

This prompts a reconsideration of Perry's (2009) account, which attempted to develop the core idea of Kaplan's theory in a more effective way. Perry advocated for an understanding of directing or referential intentions according to which they are defined by the type of connection a speaker intends to exploit:

> However, I think that we need to distinguish a number of characters for "that", corresponding to the types of connections that one can have to objects one thinks about, and choose to exploit in order to refer to those objects. The reason for this is that when one uses a demonstrative (or an indexical or name), part of what one is doing is opening a second connection for one's audience. The first connection is simply that constituted by one's own remarks. If I say, "He is a fool," you understand that, if my remark is true, someone is a fool. Who? The person I referred to with "he." That's one connection you have to the person about whom I am talking, one way you can think of him. If my remark comes out of the blue, that will be all you can discern about this fellow; you have only one channel of information about him, what I say. But usually you will take my use of "he" to be tied to some additional connection that I have to the person, that you can at least potentially use as the basis of a second connection. For example, you will take it that I am referring to the

person I am looking at, who you may then look at also. (Perry 2009, p. 196)

In §3.1, I described a modified version of the notorious Carnap/Agnew case: Imagine that we are playing a game in front of the wall of pictures - I walk along with my back to the wall and try to guess what each one depicts, pointing behind me as I go. 'That is a picture of a Founding Father', I say, then 'That is a picture of a two-time Nobel prize winner'. However, unbeknownst to you, I glanced through the window of the same room last week and saw that the third picture along was one of Rudolf Carnap. Unbeknownst to *me*, it has since been replaced with one of Spiro Agnew. Now, when I move to next picture, point behind me and say 'That is a picture of one of the greatest philosophers of the twentieth century'.

The suggested solution to this kind of case was to describe the speaker's various intentions as hierarchically structured by means-end reasoning, and conclude that the ultimately referential intention is at the bottom of the chain; the final step in a referential 'plan'. But this needed to be justified with an independent metric for fixing the final step. I advocated for using Perry's theory, but discounted considerations of the connection that the speaker expects the hearer to use, and instead claimed that the way the speaker is thinking about the object in question is enough to settle the matter. If at the time the speaker thinks of the picture of Carnap *qua* the picture of Carnap and intends to refer to *that*, her use of "that" refers to the picture of Carnap, but if she thinks of the picture of Carnap *qua* the picture behind her and intends to refer to *that*, her use of "that" refers to the picture behind her (of Agnew). (And, given the nature of the game, the second situation is more likely.)

Now that we have JA-SI, a theory that prioritises the hearer, should we reconsider? I think there is no need to. For imagine we wanted to modify the account above to incorporate the fact that a speaker may intend to draw a hearer's attention to an object via a connection of a type different to her own. For example, in the modified Carnap/Agnew case, the speaker thinks of the picture of Carnap as the picture behind her, or fixes her attention on it as such, but expects that the hearer thinks of it as the picture she sees, or fixes her visual attention on it. Thus the speaker's intention to establish joint attention on the picture must *somehow* involve a *de re* conception of the picture. But suppose that we constructed a hierarchical chain of the speaker's conflicting referential intentions taking into account both her own and the hearer's connections to the objection. The speaker's wider referential plan is to establish joint attention on the picture of Carnap. Combined with the belief that the picture of Carnap is behind her, this gives rise to an intention to focus her own attention on the picture behind her and the hearer's attention on the picture she (the hearer) sees. But the speaker cannot succeed in drawing the hearer's visual attention to the right picture without indicating that it is behind the speaker. So the chain continues: the speaker has an intention to focus her own attention on the picture behind her and the hearer's attention on the picture behind her. In other word, she intends to establish joint attention on the picture behind her.

My point here is that the speaker's intention to focus her hearer's intention can always be traced back to her own connection to the object. Thus nothing is lost by maintaining, as I originally did, that the speaker's way of thinking about the object at the time of utterance determines which intention is ultimately referential.

Introducing JA-SI also raises the question of how Perry's "connections" are related to joint attention. After all it is certainly not standard in linguistic or psychological research to think of joint attention as a possibly non-perceptual phenomenon, since it is typically studied as a result of pointing or gaze coordination. Nonetheless, I see no reason why one cannot apply the idea of 'ways of thinking' about an object to attention. To make the further claim that this type of attention can be *joint* even in non-perceptual cases, we must suppose that speakers are generally not only aware of their own myriad mental connections to various objects, but also of their interlocutors'. This might be implausible for early language users (and, accordingly, early use of demonstratives is typically to refer to currently perceived objects). But adults' competence in this respect is reflected in the ease with which we conjure up and employ different types of 'connection'. I need only say to you, 'Imagine a room. Is there furniture in that room?', and we will both understand that 'that room' refers to *the room you are currently imagining*. This shared understanding is enough to ensure that joint attention is established.

5.2   Salientists

5.2.1   Gauker's accessibility criteria

Gauker's idea is that a use of a demonstrative refers to whatever object is most salient in the context. He defines salience according to a number of 'accessibility criteria', where the referent is "the unique thing that *adequately* satisfies the criteria and satisfies them better than anything else" (Gauker 2008, p. 364). Prior reference is one such accessibility criterion, relevance (to the goals of the conversation) is another, as is being in the direction of the speaker's pointing gesture (if there is one).

It's natural at this point to wonder how it is determined which object "satisfies [the criteria] better than anything else". Rather than supposing that the accessibility criteria impose some kind of ordering on the available objects in a context, Gauker claims that identifying the most salient object is a matter of an "all-things-considered judgement" (p. 366), i.e. an overall appraisal that may weight different criteria differently depending on the context. We are encouraged to believe that all-things-considered judgements may still track objective matters, such as when someone is asked whether they like vanilla or chocolate ice cream better (p. 367). Although I think this is a questionable claim, the more problematic aspect of these "all-things-considered judgements" is Gauker's assumption that they will always determine a unique object. It seems to me that, as you might conclude that you like vanilla and chocolate equally, there

might be two or more objects that are judged to be equally salient in a context. Gauker's theory gives no indication of how his theory might deal with such scenarios. JA-SI, of course, has recourse to speaker's intentions.

Nonetheless, let's suppose for now that Gauker's all-things-considered semantic judgements always determine a unique referent for a use of a demonstrative. (Gauker does allow for the possibility that his list of accessibility criteria is not exhaustive (2008, p. 365), so perhaps we could simply add uniqueness.) It would still be unnecessary to posit this kind of theory. Gauker is primarily motivated by a firm belief that intentions can never enter into a semantic theory. But the primary reason he gives for holding this belief does not hold up in the face of JA-SI.

That primary reason[1] is that speakers' intentions are inaccessible to their hearers, thus intentionalism rests on an ineffectual notion of the proposition expressed by an utterance. For in order to play a useful role in an account of linguistic communication, Gauker argues, the proposition expressed by an utterance "has to be one that a competent hearer could assign to it using a method of interpretation that the hearer could reliably employ on the basis of features of the situation that the hearer could normally be aware of" (2008, p. 362). Accessibility criteria are supposed to be exactly these "features of the situation".

I won't deny here that the criterion given by Gauker is a desirable one (cf. Åkerman 2015). Instead, I want to point out that the notion of the proposition expressed according to JA-SI is perfectly capable of meeting it. Gauker illustrates his criticism of intentionalism with the following case:

> Suppose that Harry and Sally are at a department store and Harry is trying on ties. Harry has wrapped a garish pink-and-green tie around his neck and is looking at himself in a mirror. Sally is standing next to the mirror gazing toward the tie around Harry's neck and says, "That matches your new jacket." As a matter of fact, Sally has been contemplating in thought the tie that Harry tried on two ties back. At first she thought she did not like it, but then it occurred to her that it would look good with Harry's new jacket. We can even suppose that in saying "that" what she intended to refer to was the tie two ties back. But under the circumstances, Harry is in no position to realize that the tie she intended to refer to was the tie two ties back and therefore is in no position to take Sally's intention into account in identifying the reference of her demonstrative "that". The

---

[1]At least, the primary reason in his 2008. Gauker's 2019 paper provides some further anti-intentionalist arguments, the first of which (the circularity objection) has already been discussed extensively. The second argument claims that intentionalism is psychologically implausible, because it is too much to ask that a speaker anticipates their interlocutors' conclusions about the speaker's intentions. I do not think this position is tenable in light of the evidence that young children are competent at initiating and participating in joint-attention events (see Chapter 4), although Gauker disagrees (2019, p. 123). The third argument is that intentionalism fails to explain uses of demonstratives in thought, because these are unlikely to be accompanied by referential intentions. But according to JA-SI these can be glossed as anaphoric uses (see §6.2).

only thing one could reasonably expect Sally's demonstrative "that" to refer to is the pink-and-green tie around Harry's neck. (2008, p. 363)

This kind of complaint echoes the accusations of 'Humpty Dumptyism' discussed in §2.1.1 and more recently in §5.1.1. My response in §2.1.1 — that this kind of case demonstrates a lack of a plausible communicative intention rather than true intention-reference mismatch — would not satisfy Gauker, because it implies that the proposition expressed can be totally obscure to a competent hearer. But, as I wrote in §5.1.1, JA-SI gives another way of getting to the same result, and in particular one that is compatible with Gauker's view of the proposition expressed. For JA-SI stipulates that the proposition expressed by an utterance containing a demonstrative is dependent on the speaker's intention to establish joint attention on an object. In the tie case described above, Sally clearly does not have a genuine intention to draw Harry's attention to the tie two ties back, because her actions are not sufficient to achieve this aim (and she knew they would not be sufficient to achieve it before acting).

What's more, in all cases in which a speaker *does* have a genuine referential intention, the proposition expressed by a demonstrative utterance will meet Gauker's expectations. In order to have her utterance establish joint attention on the intended object, the speaker must ensure that the hearer's attention is either already on that object or will be directed towards it. And of course the latter will only happen if the object is made sufficiently salient to the hearer, i.e. by satisfying one or more of Gauker's accessibility criteria. On the joint-attention simple intentionalist view, these are not semantically relevant *per se* but *clues* to the speaker's semantically relevant referential intention. Pointing, for example, is obviously a signal that the speaker intends to draw the hearer's attention to a particular object. So, if accessibility criteria are "features of the situation that the hearer could normally be aware of" (by Gauker's own assertion) and accessibility criteria normally make the hearer aware of the speaker's referential intention (by the argument just given), then referential intentions are also features of the situation that the hearer could normally be aware of. Thus JA-SI is compatible with Gauker's notion of the proposition expressed.

This argument doesn't exactly refute salientism; the fact that referential intentions *could* play a suitable role in determining the proposition expressed on a given occasion doesn't prove that they *do*. But it certainly impugns Gauker's motives for rejecting intentionalism. Moreover, we can use the connection between accessibility criteria and referential intentions to cast doubt on his own theory. Consider how the tie example would seem if we afforded Sally a genuine referential intention, but took it as a case of self-talk. That is, suppose that, instead of speaking to Harry, Sally had muttered or even merely thought to herself, "That matches your new jacket". Nothing has changed in terms of Gauker's accessibility criteria, but of course it would now be ridiculous to insist that Sally's use of 'that' actually referred to the pink-and-green tie around Harry's neck. JA-SI can explain this fact: accessibility criteria are only relevant insofar as they serve as clues to a speaker's referential intention, and such clues

50

are unnecessary in cases of self-talk because one can always use 'that' to refer to whatever object one's attention is fixed on at the time. The salientist cannot account for this, at least not without employing a notion of salience that partially depends on interlocutors' subjective mental states. In fact, this is exactly what Mount (2008) does[2]. And, as I am about to argue, it doesn't do much better than Gauker's version in terms of competing with JA-SI.

5.2.2 Mount's mutually recognised salience

Mount's version of salientism is unique in that it explicitly rejects the idea that salience can be defined without reference to interlocutors' mental states. Rather, an object is salient for Mount if it is "the focus of perceptual or cognitive attention" (2008, p. 155). And the referent of a use of a demonstrative is the object that is "mutually recognised as maximally salient by conversational partners", i.e. "all interlocutors have focused their attention of it, and are aware that they have all focused their attention on it". This, of course, is identical to the definition of joint attention (a connection Mount does not make), so the starting point for our comparison is obvious and Mount's proposal reaps many of the same benefits that JA-SI does in terms of the cases it can explain. Nonetheless, the fact that JA-SI takes the speaker's intention to establish joint attention rather than joint attention itself to be semantically relevant is significant—and, I claim, it is preferable.

As Speaks (2017, p. 717-8) remarks, one glaring flaw with Mount's theory is that a number of objects may be in a given group of interlocutors' joint attention when one of them uses a demonstrative. On an intentionalist account, the referent is distinguished by the fact that the speaker intends to refer (according to JA-SI, establish joint attention[3]) on one object in particular. Now, Mount does concede that interlocutors tend to operate under a general principle of "charity for the speaker" (2008, p. 159), according to which they pay particular attention to the speaker's state of mind when trying to interpret an utterance of a demonstrative. She glosses this process as follows: "In trying to interpret utterances so that they are true (or at least reasonable), hearers try to infer what the speaker must be thinking about. They may recognize what is most likely salient to the speaker, and this recognition itself may make the object

---

[2]There are also hints of a similar move in Gauker (2019), although it is not fully worked out. For example, Gauker writes that "what establishes the reference of a demonstrative in thought may be just the relevance to what the thinker is doing in having the thought" (p. 126). This suggests that he thinks the accessibility criterion of relevance (to conversational goals) can be tweaked to apply to cases of self-talk. But I struggle to see how this idea could ever really get off the ground. It relies on the improbable assumption that thoughts are usually had on the way to achieving some further goal, thereby ignoring cases in which one simply thinks for the sake of thinking or a thought simply pops unwilled into one's head. I cannot think of how Gauker could account for these cases without admitting the significance of subjective mental states.

[3]One might (should) ask: how can one intend to establish joint attention on object if it is already in the joint attention, albeit amongst other things? My idea is to view this as a kind of *re*-establishing or reconfiguration of joint attention, which demands that this time the focus is *uniquely* on that object.

salient to them as well" (2008, p. 160).

However, this explanation clearly gets the order of reference and interpretation the wrong way round. It is inaccurate to characterise such a process as 'interpretation' when Mount herself claims that the proposition expressed is by a demonstrative utterance is dependent on the interlocutors' ultimately focusing their attention on the 'right' object. If that is the case, then the process of figuring out what is salient to the speaker and making it salient to oneself is not a matter of interpretation but *part of the process of reference*. This has the rather strange consequences that 1) reference is not immediately achieved upon utterance 2) the process of reference is dragged out until the last interlocutor catches on (and may never be achieved if that person is not paying sufficient attention). Both of these depart significantly from standard views of linguistic communication, and, while it is not impossible that they are true, showing that they are requires much more argument than Mount provides. To avoid this requirement, i.e. to have her notion of 'interpretation' indeed count as interpretation, Mount would have to admit that the principle of charity for the speaker is driven by the fact that the speaker's intention to establish salience (joint attention) is what fixes reference, rather than the salience itself. That is, she would have to adopt JA-SI.

5.3 Quasi-intentionalists

To reiterate, quasi-intentionalists are those who do not believe that demonstratives have inherent referential power beyond their use as speaker-referential devices. In Bach's view (2005; 2008a; 2008b; 2017), this is best dealt with by conceding that demonstratives do not semantically refer. In Schiffer and Neale's (forthcoming) view, we should simply identify the semantic referent of a use of a demonstrative with the object the speaker is referring to. For my purposes here, this difference is not significant enough to discuss the two views separately.

In §3.2, I argued that there is reason to think that referential intentions exist separately from the kind of communicative intentions invoked by quasi-intentionalists, and in §7.3 I talk about the incompatibility between JA-SI and the Gricean project in general, which attempts to explain *all* semantic meaning in terms of communicative intentions. But here I want to point out that the adoption of JA-SI does not exclude the possibility that quasi-intentionalism provides an accurate description of the speaker-referential function of demonstratives.

Let's return to Schiffer and Neale's definition of referring with a demonstrative:

> In uttering u, S referred to x with e, relative to its i-th occurrence in u, iff for some person A and property φ, S intended it to be common ground between S and A that the i-th occurrence of e has φ and, at least partly on that basis, that S referred to x in uttering u. (forthcoming, p. 6)

What kind of thing might this property φ be? For Schiffer and Neale, it essentially amounts to the standing meaning of an expression, which in the case of demonstratives is exhausted by the fact that they are typically used as speaker-referential devices as well as any constraints on gender, number, etc. For example, the relevant property of 'she' is that the referent of a use of 'she' is an object o iff o is female and the speaker refers to o with her use of 'she'. But of course if one has a different view of the semantics of demonstratives, one can simply switch out the property to reflect that view. For example, JA-SI suggests that for demonstratives φ is the fact that their reference is decided by the speaker's intention to establish joint attention on a particular object.

So, adopting JA-SI does not mean that quasi-intentionalists are wrong in claiming that demonstratives are used to speaker-refer in the way described. Rather, they have misidentified the way in which speaker reference and semantic reference depend on each other in the case of demonstratives. It is not that demonstratives semantically refer due to their function as speaker-referential devices, rather their function as speaker-referential devices is guided by their *even more fundamental* function as devices for establishing joint attention (more on this in §7.3). And once we perceive the relationship in this way, we are able to explain how demonstratives can semantically refer without playing a concordant role in communication, i.e. explain cases of content-communication divergence and particularly self-talk, without losing the desirable result that in ordinary cases the two types of reference will coincide.

Chapter 6

Loose ends

In this chapter, I clarify some details and consider some possible extensions of JA-SI.

6.1 Phi-features, deictic proximity, and complex demonstratives

Until now I have avoided discussion of what are known as syntactical markers for number, gender and case, or *phi-features*. Just within the category of Kaplanian true demonstratives are 'those' and 'these' as well as 'this' and 'that'; taking into account demonstrative uses of pronouns we must also consider 'she', 'him', 'they', etc. And then there is the matter of distinguishing the *distal* demonstratives—that' and 'those'— from the *proximal* ones—'this' and 'these'. Given a joint-attentional account, ought these differences to be afforded semantic import or thought of as merely a communicative aid? Affording them semantic import would give the following definition for 'she', for example:

> *Joint-attention simple intentionalism (JA-SI)*
> The referent of a demonstrative use of 'she' is an object o iff 1) o is female[1] and the speaker intends her use of 'she' to establish o as the subject of the interlocutors' joint attention.

If phi-features and proximity were taken to be merely communicative aids, however, all demonstratives and demonstrative uses of pronouns would have the same 'meaning' but different conventions governing their use. One argument in favour of that choice is that it is in keeping with the spirit of intentionalism, and particularly with the version of intentionalism laid out here, which does its best to explain demonstrative reference entirely in terms of the speaker's mental state. What is important is not whether the intended referent is *actually* female, but whether the speaker believes that using a female-marked pronoun is a way

---

[1]Setting aside certain exceptions as in fn. 2.

to establish joint attention on that referent. This is a matter of the speaker's beliefs about the relevant audience's beliefs about the individual in question, not actual facts about the world.

To add support to this option in the case of the proximal/distal distinction, it should be noted that the rules governing it are not universally agreed upon. Although a number of factors encourage speakers to prefer one over the other—not just distance (e.g. Luz and van der Sluis 2011; Coventry et al. 2014; Maes and de Rooij 2007) but salience (Maes and de Rooij 2007), ownership (Coventry et al. 2014) and presence/absence of a coincident pointing gesture (Piwek et al. 2008)—there is by no means a consensus in many cases. To me, this suggests that mastering the distinction is more a matter of learning to make oneself understood than learning a definite semantic rule.

However, there are also definite downsides to the idea that phi-features are not actually semantically relevant. One result is that a number of sentences which intuitively appear to be tautologies, e.g. 'She is female', turn out not to be. And while in the case of English the choice may large amount to a matter of theoretical preferences, other languages with more complex systems of demonstrative pronouns strain the limits of the theory. For example, in German the proximal/distal contrast is not built into individual expressions but added to neutral pronouns via the *adverbial demonstratives* 'hier' and 'da' (Diessel 2005). To properly analyse this fact without supposing that these expressions serve as semantic restrictions, we would need a corresponding account of how the resulting demonstratives 'das hier', 'das da', etc. are formed via the normal compositional syntactic rules but somehow avoid the normal semantic consequences.

Then again, other languages may provide support for the idea that these restrictions are *not* semantic, too. In particular, languages in which gendered pronouns (and hence their demonstrative uses) are decided upon according to a grammatical system of gendered nouns, there is no semantic (in the sense of world-describing) information carried by e.g. the use of a feminine pronoun. Given these considerations, it is hard to claim that one alternative will be universally superior. We may have to concede that it largely depends on one's purposes in giving the semantic theory.

Another related and heretofore unmentioned consideration is the extension of JA-SI to complex demonstratives, e.g. 'that chair'. The 'extra' material seems more significant here; it is certainly standard to analyse complex demonstratives as inheriting the semantic restrictions of the individual expressions they are composed of. But given our new account of demonstrative reference, we might question whether the standard account is correct. Remember, after all, that Kaplan was intrigued by the possibility that his directing intentions account could explain a demonstrative version of Donnellan's 'martini glass' example, in which the speaker successfully uses 'that man drinking a martini' to refer to a man drinking water out of a martini glass. If we were to take the description as merely an extra tool to aid the establishment of joint attention (and provide a compositional theory accounting for this), the speaker's utterance would have *actually* referred to the man drinking water. If we were to take it as an extra

semantic constraint, this would be a case of reference failure.

The first option recalls the worries about 'Humpty Dumptyism' discussed in §2.1.1. Surely one's use of 'that chair' cannot refer to, say, an elephant? But perhaps restrictions can be found in terms of what kinds of referential intention one might reasonably form. If one's interlocutor is known for mistaking elephants for chairs, it is certainly possible that they will understand your utterance of 'that chair' to refer to the elephant. So perhaps we should not insist that your utterance of 'that chair' *actually* didn't refer to anything, just for theoretical harmony.

Then again, one might prefer to analyse the chair/elephant mix up as a case of idiolectical variation, in which the word 'chair' takes on a different meaning in the interlocutors' shared language. And Donnellan's martini example can easily be explained in other ways, e.g. with Gricean communicative intentions. All in all, I again find myself rather undecided on the matter. It may be that for formal semantic purposes, it makes more sense to analyse the extra material of complex demonstratives as semantically restrictive. But this does not mean that the initially objectionable alternative, i.e. the idea that complex demonstratives have no more semantic import than their simple counterparts, is without its own merits.

6.2   Secondary uses of demonstratives

I suggested in §4.3 that some uses of demonstratives might be thought of as 'anaphoric', particularly those in which joint attention is already established on a particular object so uses of a demonstratives 'automatically' refer to it. This, then, would mean that their utterance is not be caused by a 'basic' or 'primary' referential intention, i.e. one to establish joint attention. This gives a particular primary status to demonstratives uses that occur when joint attention has not yet been established, meaning that either at least one interlocutor has her attention focused on something else or the attention is merely shared attention (not mutually recognised). But of course this far from covers the full range of demonstrative use. One advantage of JA-SI is that it can explain what are *usually* called anaphoric uses of demonstratives in terms of derivation from their basic use, rather than having to suppose that demonstratives have a number of different, unrelated uses.

For example, take the following utterance:

(7)   There are plenty of leftovers, but that's because we cooked too much.

One could say that the demonstrative gets it reference anaphorically because the antecedent has already brought the presence of leftovers to the interlocutors' joint attention. But in fact the antecedent cannot fully specify the referent in this way, because of utterances like:

(8)   There are plenty of leftovers, so that's what we're eating for dinner.

In (5), 'that' refers to the *presence* of leftovers, in (6) to the leftovers themselves. So perhaps we should not be too hasty to say that anaphoric uses of demonstratives do not trade on the existence of referential intentions; they might rely on them to distinguish between a number of objects in the joint attention. If so, *true* anaphoric uses of demonstratives are quite rare, only occurring when there is a single object in the interlocutors' joint attention. And we can see how their distinction from primary uses might be flexible and highly situation-dependent, rather than clear-cut and unrelated.

It is also worth considering that in cases like (4) and (5), hearing the content of the complete sentence will likely have a retroactive effect on the hearer's interpretation of the demonstrative. This would suggest that clues to the speaker's referential intentions need not always be present at the time of utterance, and hence that interlocutors' joint attention may not ever be truly coordinated (by the time the hearer has figured out enough to fix her attention on the leftovers, the speaker may have switched her attention elsewhere). Given this fact, can we say that the speaker has a genuine referential intention in such a case? If her use of a demonstrative alone is not sufficient to establish joint attention on the object, it would seem irrational to intend that it does. But then again, given the fast pace and second nature of making and understanding such utterances, perhaps the difference simply does not matter enough for speakers to rationally modify their intentions in such a way. The speaker is likely to never find out whether the hearer's joint attention was fixed on the intended object at the intended time or not (as the hearer will simply catch up almost immediately after via contextual clues), and thus has no reason to rethink her mode of intention formation. 'Irrational' seems too strong an accusation. Rather, we could accept that both speakers and hearers arrive at communicative scenarios with a number of (general and occasion-specific) background expectations that ameliorate their possibly imperfect reasoning, to the point that attempting to perfect that reasoning is simply not that important.

Another important thing to note is that uses of demonstratives in self-talk (whether it is spoken aloud or merely in thought) will always be anaphoric uses. Perry (2009, p. 197) points out how easy it seems to be to refer without an audience:

> Look at something. Refer to it as "that." I'm now looking at the Sprite can on the table, and I just said, out loud, "That is empty." Now I'm not looking at it. I'm remembering taking a full can of Sprite from the refrigerator yesterday. I just said out loud, "That can is now empty." I was referring to the can I referred to as the can connected to me through this memory, not as connected to me through perception.

According to JA-SI, one can always have a use of a demonstrative refer to whatever is already in the joint attention. So switching between referents in the way Perry describes is exactly as easy as switching one's attention between objects.

57

Chapter 7

Conclusion

7.1 Summing up

In attempting to give a coherent theory of referential intentions, we have ended up with a whole new theory of demonstrative reference: JA-SI. We would do well to revisit how we got here, and what it reveals about the deficits of other varieties of intentionalism.

As we saw in the introduction, intentionalism about demonstratives was born out of Kaplan's (1989b) 'directing intentions' theory, which aimed to explain the relevant intentions in terms of their relationships to demonstrations, or ways of demonstrating individuals (pointing gestures, definite descriptions, etc.). But more recent defenders of intentionalism simply took a referential intention to concern *reference* itself. As we saw in Chapter 2, this initially leaves intentionalists vulnerable to accusations of 'Humpty Dumptyism'. Theorists like King, Stokke and Speaks respond by adding some kind of (actual or potential) recognisability criterion to their semantic theories, leading to reflexive intentionalism. But I argued that this kind of move is made redundant by existing theories of *communicative* intentions, and that it leaves us with no explanation for cases of so-called content-communication divergence. Thus if one is to be an intentionalist, one ought to be a simple intentionalist.

Now that we have seen JA-SI, it is even clearer why reflexive intentionalism is misguided (see §5.1.1). Recognisability need not be added as a separate condition or additional intention, rather it can already be found within the content of a simple referential intention. For if referential intentions are joint-attention-establishing intentions, recognisability to one's audience is obviously relevant for forming a genuine referential intention. But recognisability is still not *necessary* to establish joint attention. Particularly, cases of self-talk are explained not by the implausible notion of recognising one's own reflexive intentions, but by the straightforward truth that one always shares joint attention with oneself.

Next, in §3.1, I turned to the problem of conflicting intentions. I agreed with other theorists that a Bratman-style theory of intentions as hierachically

structured entities can help to illuminate cases in which speakers' referential intentions appear to conflict. In response to a challenge by Speaks that identifying the 'final' referential intention in such a structure presupposes one's preferred theory, I suggest we take inspiration from Perry's (2009) theory of directing intentions and consider in what way the speaker is thinking of the object at the time of utterance. As already discussed in §5.1.2, JA-SI can not only benefit from this idea but can do so without appealing to the way the speaker intends the hearer to think of the object, as this will always be equal to the way the hearer is thinking of it when forming a joint-attention-establishing intention.

In §3.2, I argued against the claim that referential intentions can be reduced to communicative intentions. I proposed that referential intentions are thought of as intentions in action rather than prior intentions (unlike typical accounts of communicative intentions), according to the distinction made in Searle (1983). This point has since been strengthened by the introduction of JA-SI; we would expect intentions to establish joint attention to be typically present-directed.

In §3.3, I argued that once we adopt the framework of intentions in action accusations of circularity in intentionalism are significantly weakened; there is no need to suppose that speakers have prior conscious awareness of the satisfaction conditions of referential intentions before they can form them. But the objection still remains in a sense, as was pointed out at the beginning of Chapter 4: referential intentions ought to be formed on the basis of some initial referential action. Using evidence from the study of child language acquisition, I argued that this action is the practice of establishing joint attention. This led to the presentation of JA-SI in §4.3, and to the comparisons to other positions in the intentionalist/anti-intentionalist debate that make up Chapter 5. §5.1, which has already been mentioned above, compares JA-SI to other intentionalist accounts while §5.2 compares it with salientist accounts and §5.3 with quasi-intentionalism. Chapter 6 addressed some further details of the possible development of JA-SI, namely extension to demonstratives marked for number/gender/case/proximity and complex demonstratives as well as anaphoric uses of demonstratives.

In the rest of this chapter, I address a potential issue with using Searle's intentional framework (§7.2), then move to discussing the implications of JA-SI for wider problems in the philosophy of language that were raised in Chapter 1.

7.2   Intentions in action

In employing Searle's framework, the theory laid out here risks inheriting some of its criticisms. In particular, some have raised concerns that positing intentions in action, which are self-referential, leads to circularity; their conditions of satisfaction pertain not just to the relevant (future or current) action but also to the intentional quality of that action (Wilson 1989; Ginet 1990). Thus although we solved the problem of *referential* circularity solved in §3.3., there may be a problem of *intentional* circularity no matter how we specify the content of the relevant intention.

Mirroring Lewerentz and Marschall's (2018) comments on referential circularity and my discussion of them in §3.3, it should be noted that intentional circularity cannot be avoided by appealing to theoretical usefulness without complete explanation. That is to say, no one can deny that Searle's theory provides an *explication* of intentional behaviour, even if this explication is partially in terms of the explication. Nonetheless, again as with the problem of referential circularity, the possibility of explanatory failure lurks in the question of intention formation, and specifically how the *first ever* intention could be formed if it necessitated familiarity with intentional action. But in this case it seems even more unfair to ask the philosopher of action to explain how one becomes familiar with the notion of behaving intentionally. For in the study of child development, the question of how infants move from reflexive to controllable behaviours is by no means resolved. To address it here would be to delve into a number of questions about the historical development of human consciousness and controllable behaviour, which are well beyond our current purview and perhaps even the purview of philosophy in general. The existence of a possible grounding in psychology or biology, however, should help assuage worries that self-referential theories of intention are threatened by this type of circularity.

7.3 Intention-based semantics

Intention-based semantic theories attempt to reduce all semantic phenomena to Gricean communicative intentions. On the surface, this is not compatible with JA-SI; the semantic referent of a demonstrative is determined not by the speaker's communicative (speaker-referential) intentions, but her intentions to establish joint attention. As noted in §3.2, this is not to deny that communicative considerations are usually at the forefront on occasions of reference. In fact, the joint-attention-establishing function of demonstratives can provide a new way of understanding their use as speaker-referential devices (see §5.3). But there is still nothing about referential intentions themselves that grounds them in the inducement of certain propositional attitudes in one's audience.

This may not be a bad thing, however. Holger Diessel, whose 2006 paper provided a foundation for linking demonstratives and pointing gestures in §4.2, has more recently extended his theory to one of language evolution. That is, in support of Arbib's (2012) hypothesis that language developed from the use of gestures, Diessel (2013) argues that demonstratives provided one necessary link between the two (Diessel 2013). While Diessel's work is still exploratory, it points to a particular understanding of language *not* as ultimately based on occasion-specific proposition-communication[1], but occasion-specific joint-attention-establishment.

---

[1] I don't mean to imply, of course, that Grice's theory is intended to say anything about the evolution of language. Rather, Diessel supposes there is a commonality between child language acquisition, language evolution, and adults' linguistic capacities. Grice's theory *does* purport to explain the last of these.

Of course, the question of how one gets from establishing joint attention to communicating propositions is a large and difficult one. But it is no reason to suppose that a Gricean theory is superior. In fact, one fundamental flaw of the Gricean theory is that propositions themselves cannot ultimately be grounded in communicative intentions; they must be taken as already 'available' for speakers to exchange on given occasions. This implies a rather strange (although not totally unthinkable) view of the ontology of propositions. A joint-attention based theory, however, would be able to start from something much more simple: a basic theory of mind and the existence of object-directed attention. Where the theory might go from there is, of course, not something that can be decided here, but one goal would be to describe how an understanding of propositional communication arises from these basis concepts.

7.4   The semantics/pragmatics divide

One benefit of Grice's theory is that it provides a way of delineating the boundaries between semantics and pragmatics; the former is determined by 'timeless' meaning, the latter by occasion-specific speaker-meaning. As was just noted in §7.3., JA-SI rules out the possibility of Griceanism 'all the way down', but that doesn't necessitate incompatibility with this kind of boundary-drawing. Rather, I think JA-SI is compatible with this understanding of semantic and pragmatic meaning.

Searching for scenarios in which the literal and implied referent of a use of a demonstrative diverge brings us back to Michaelson's (2013) tricky Toyota-McLaren case:

> You and I both work for a large investment bank. We've just been paid an obscenely large bonus, and we are both at the office very early one day. We run into each other near a window overlooking the parking lot, which is nearly empty at this hour. There is, however, a small cluster of cars which includes several nondescript vehicles and, in addition, a McLaren F1. You ask me what I decided to do with my bonus, and I respond by pointing out the window at the cluster of cars and saying: "I bought that." In fact, I bought the Toyota Camry parked next to the McLaren, and I intend to refer to the Camry, not the McLaren. That said, I've decided to have a bit of fun with you. Even in a situation like this one, I take it, I have plausibly succeeded in referring to my Camry. (2013, p. 91-2)

Plausibly, one could say that I *literally* said that I bought the Camry but *implied* that I bought the McLaren. I am not sure whether this is a desirable result or not; it seems that attempts to separate speaker and semantic reference for demonstratives will lead to strange cases. But at least by explicating referential intentions in terms of something other than Gricean communication, we retain the possibility of dividing up semantics and pragmatics according to the Gricean

conception. It is just that the semantic and pragmatic content of a use of a demonstrative will almost always coincide.

7.5 Theories of reference

As noted in §7.3, JA-SI eliminates the possibility that semantic reference can be entirely explained in terms of Gricean speaker-referential intentions. But could it constitute its own theory of reference, shedding light on the nature of reference outside the realm of demonstratives?

One particularly salient opportunity for extension is to the realm of definite descriptions, Donnellan's (1966) observations about which inspired Kaplan's original comments on directing intentions. Diessel (2006, 476-8) also provides evidence that the development of definite articles across a number of language originates with demonstratives. If we continue to take diachronic/acquisition evidence as relevant to theories of contemporary linguistic meaning, this would suggest that the 'automatic' referential power of definite descriptions (corresponding to Donnellan's attributive use) is somehow derived from the basic motive of establishing joint attention.

Grice's theory of the relationship between linguistic- and occasion-meaning can provide a model here. Except that theory begins at the utterance level, and then explains the linguistic meaning of individual expressions according to the occasion-meaning of utterances they typically occur in. A joint-attention theory of reference, in contrast, would begin and end at the expression level.

Take, for example, a definite description. Uses of the definite description coincide with instances in which speakers intend to establish joint focus on an object satisfying the description. Thus it eventually gains 'automatic' referential power, referring to whatever object satisfies that description. One could even extend this type of theory to all kinds of referential expressions, including proper names. Demonstratives would be something like a special case, in which no typical pattern is ever established, since uses of demonstratives are used to establish joint attention on almost anything. Again, this mirrors Gricean theories of reference; Schiffer and Neale's quasi-intentionalist account of demonstrative reference essentially views them as speaker-referential devices that never achieved full semantic-referential status. As described in §5.3, there is a way to integrate JA-SI into the speaker-referential side of their theory. But if we are not going to fully commit to the Gricean paradigm, it is fair to wonder if such a move is prudent. The alternative is to fully commit to a joint-attentional view, and claim that speaker reference itself is merely a matter of intent to establish joint attention.

Adopting this type of theory, however, would result in an inability to explain cases like Michaelson's (2013) Toyota-Camry confusion. I suggested in §5.1.1 that one way to explain the divergence between semantic and speaker reference in this case is to suppose that the hearer counts as an interlocutor for communicative but not joint-attention-establishment purposes. This leaves a place for the Gricean theory. As I have said before, however, I think the Toyota-Camry

one is a strange one and intuitions are likely to diverge not only on the best way to explain it but even on what there is to be explained. If this type of scenario is the only motivation to keep both a Gricean and joint-attentional notion of speaker reference, then I do not think a very strong case can be made for it. Better to consider that both semantic and speaker reference can be grounded in joint attention alone.

7.6 Final remarks

This thesis set out to give a plausible account of referential intentions, which I think had been achieved, and ended with a tentative proposition about reference *in general*. As for the first aim, I think it has been satisfactorily met, and that recognising the role of joint attention in using demonstratives would serve the intentionalist/anti-intentionalist debate well. Many anti-intentionalists complain that intentionalism over-complicates the matter, generating an ever-expanding number of increasingly complex intentions to explain apparently simple speaker behaviour. But JA-SI grounds demonstrative reference in an activity so simple that infants regularly engage in it: establishing joint attention on particular objects.

Although little of the thesis as a whole turns on the ideas about reference in general presented in the last section, I think it is interesting that an in-depth study of demonstrative use led us there. The extreme referential variability of demonstrative expressions has made them notoriously difficult to fit into a semantic theory in contrast to relatively more stable referential expressions like proper names, which might suggest that they are somehow deviant or particularly complex. But if we take evidence from the study of child language acquisition and language evolution seriously, it would seem that demonstratives are the most *fundamental* referential devices, appearing early in our individual and collective linguistic lifetimes. It is natural, then, to consider that their function (i.e. to establish joint attention) may play a key role in a foundational as well as a descriptive theory of meaning.

References

- Anscombe, G. E. B. 1957. Intention. Oxford: Basil Blackwell.

- Åkerman, J. 2009. A plea for pragmatics. Synthese, 170, 155-67.

- Åkerman, J. 2010. Communication and indexical reference. Philosophical Studies, 149(3), 355-66.

- Åkerman, J. 2015. The communication desideratum and theories of indexical reference. Mind and Language, 30(4), 474-99.

- Arbib, M. 2012. How the brain got language: The Mirror System Hypothesis. Oxford: Oxford University Press.

- Bach, K. 1987. On communicative intentions. Mind and Language, 2(2), 141-54.

- Bach, K. 1992. Intentions and demonstrations. Analysis, 52(3), 140-6.

- Bach, K. 2001. You don't say? Synthese, 128, 15-44.

- Bach, K. 2005. Context ex machina. In: Z. Szabó (ed.), Semantics versus Pragmatics. Oxford: Oxford University Press, 15-44.

- Bach, K. 2008a. On referring and not referring. In: J. Gundel and N. Hedberg (eds.), Reference: Interdisciplinary Perspectives. Oxford: Oxford University Press, 13-50.

- Bach, K. 2008b. What does it take to refer? In: E. Lepore and B. C. Smith (eds.), The Oxford Handbook of Philosophy of Language. Oxford: Oxford University Press, 516-54.

- Bach, K. 2017. Reference, intention, and context: Do demonstratives really refer? In: M. de Ponte and K. Korta (eds.), Reference and Representation in Thought and Language. Oxford: Oxford University Press, 57-72.

- Bayley, N. 1995. Bayley Scales of Infant Development, Second Edition: Manual. San Antonio, TX: The Psychological Corporation.

- Bergelson, E. and Swigley, D. 2012. At 6-9 months, human infants know the meanings of many common nouns. PNAS, 109(9), 3253-8.

- Bratman, M. 1987. Intention, Plans and Practical Reason. Cambridge, MA: Harvard University Press.

- Burge, T. 2009. Five theses on de re states and attitudes. In: J. Almog and P. Leonardi (eds.), The Philosophy of David Kaplan. Oxford: Oxford University Press, 246–316.

- Butterworth, G. 1995. Origins of mind in perception and action. In: C. Moore and P. J. Dunham (eds.), Joint Attention: Its Origins and Role in Development. Hillsdale, NJ: Erlbaum, 29–40.

- Butterworth, G. and Grover, L. 1988. The origins of referential communication in human infancy. In: L. Weiskrantz (ed.), Thought without Language. Oxford: Clarendon Press, 5–24.

- Carroll, L. 2002. Alice's Adventures in Wonderland and Through the Looking Glass. Oxford: Oxford University Press.

- Carston, R. 2002. Thoughts and Utterances: The Pragmatics of Explicit Communication. Oxford: Blackwell.

- Clark, E. 1978. From gesture to word: On the natural history of deixis in language acquisition. In: J. Bruner and A. Garton (eds.), Human Growth and Development. Oxford: Oxford University Press, 85-120.

- Clark, E. and Sengul, C. J. 1978. Strategies in the acquisition of deixis. Journal of Child Language, 5, 457-75.

- Corazza, E., Fish, W. and Gorvett, J. 2002. Who is I? Philosophical Studies, 107(1), 1-21.

- Coventry, K. R., Griffiths, D., and Hamilton, C. J. 2014. Spatial demonstratives and perceptual space: Describing and remembering object location. Cognitive Psychology, 69, 46–70.

- Diessel, H. 2005. Distance contrasts in demonstratives. In: M. Haspelmath, M. S. Dryer, D. Gil and B. Comrie (eds.), The World Atlas of Language Structures. Oxford: Oxford University Press, 170-5.

- Diessel, H. 2006. Demonstratives, joint attention, and the emergence of grammar. Cognitive Linguistics, 17(4), 463-89.

- Diessel, H. 2013. Where does language come from? Some reflections on the role of deictic gesture and demonstratives in the evolution of language. Language and Cognition, 5(2-3), 239-49.

- Donnellan, K. 1966. Reference and definite descriptions. The Philosophical Review, 75(3), 281-304.

- Frege, G. 2009. On sense and reference. In D. Byrne and M. Kölbel (eds), Arguing about Language. London: Routledge, 49–55.

- Gauker, C. 2008. Zero tolerance for pragmatics. Synthese, 165, 359-71.

- Gauker, C. 2019. Against the speaker-intention theory of demonstratives. Linguistics and Philosophy, 42, 109-29.

- Ginet, C. 1990. On Action. Cambridge: Cambridge University Press.

- Grice, H. P. 1957. Meaning. The Philosophical Review, 66, 377–88.

- Grice, H. P. 1968. Utterer's meaning, sentence-meaning and word-meaning. Foundations of Language, 4(3), 225-42.

- Grice, H. P. 1969. Utterer's meaning and intention. The Philosophical Review, 78(2), 147-77.

- Grice, H. P. 1989. Studies in the Way of Words. Harvard: Harvard University Press.

- Hespos, S. J. and van Marle, K. Physics for infants: characterizing the origins of knowledge about objects, substances, and number. WIREs Cognitive Science, 3, 19-27.

- Horwich, P. 1998. Meaning. Oxford: Oxford University Press.

- Horwich, P. 2006. Reflections on Meaning. Oxford: Oxford University Press.

- Jaszczolt, K. M. 2012. Semantics and pragmatics: the boundary issue. In: K. von Heusinger, P. Portner and C. Maienborn (eds.), Semantics: An international handbook of natural language meaning. Berlin: Mouton de Gruyter.

- Kaplan, D. 1978. Dthat. In: P. Cole (ed.), Syntax and Semantics. New York: Academic Press, 221-43.

- Kaplan, D. 1989a. Demonstratives. In: J. Almog, J. and H. Wettstein (eds.), Themes From Kaplan. Oxford: Oxford University Press, 481-564.

- Kaplan, D. 1989b. Afterthoughts. In: J. Almog, J. and H. Wettstein (eds.), Themes From Kaplan. Oxford: Oxford University Press, 565-614.

- King, J. C. 2013. Supplementives, the coordination account and conflicting intentions. Philosophical Perspectives, 27(1), 288-311.

- King, J. C. 2014. Speaker intentions in context. Nous, 48(2), 219-37.

- Kripke, S. 1977. Speaker's reference and semantic reference. Midwest Studies in Philosophy, 2(1), 255-76.

- Levinson, S. 2000. Presumptive Meanings. Cambridge, MA: MIT Press.

- Lewerentz, L. and Marschall, B. 2018. Metasemantics, intentions and circularity. Synthese, 195, 1667-79.

- Liebal, K. and Call, J. The origins of non-human primates' manual gestures. Philosophical Transactions of the Royal Society B, 367, 118-28.

- Luz, S., and van der Sluis, I. 2011. Production of demonstratives in Dutch, English and Portuguese dialogues. In: C. Gardent and K. Striegritz (eds.), Proceedings of the 13th European Workshop on Natural Language Generation. Stroudsburg, PA: Association for Computational Linguistics, 181–6.

- MacKay, A. F. 1968. Mr. Donnellan and Humpty Dumpty on referring. Philosophical Review, 77(2), 197–202.

- Maes, A. and de Rooij, C. 2007. (How) do demonstratives code distance? In: A. Branco, T. McEnery, R. Mitkov, and F. Silva (eds.), Proceedings of the 6th Discourse Anaphora and Anaphor Resolution Colloquium DAARC 2007. Lagos, Portugal: Centro Linguistica da Universidade do Porto, 83-9.

- Michaelson, E. 2013. This and That: A Theory of reference for names, demonstratives, and things in between. PhD thesis, UCLA.

- Montague, R. 1968. Pragmatics. In R. Klibansky (ed.), Contemporary Philosophy: A survey. Florence: La Nuova Italia Editrice, 102–22.

- Mount, A. 2008. Intentions, gestures, and salience in ordinary and deferred demonstrative reference. Mind and Language, 23(2), 145-64.

- O'Rourke, M. 2019. Referential intentions. In: J. Gundell and B. Abbott (eds.), The Oxford Handbook of Reference. Oxford: Oxford University Press, 19-44.

- Perry, J. 1997. Indexicals and demonstratives. In: B. Hale and C. Wright (eds.), Companion to the Philosophy of Language. Oxford: Blackwell, 486-612.

- Perry, J. 2009. Directing intentions. In: J. Almog and P. Leonardi (eds.), The Philosophy of David Kaplan. Oxford: Oxford University Press, 187-289.

- Piaget, J. 1952. The origins of intelligence in children. New York: International Universities Press.

- Piwek, P., Beun, R., and Cremers, A. 2008. "Proximal" and "distal" in language and cognition: Evidence from deictic demonstratives in Dutch. Journal of Pragmatics, 40, 694–718.

- Predelli, S. 2002. Intentions, indexicals and communication. Analysis, 62 (4),310–6.

- Recanati, F. 2004. Literal Meaning. Cambridge: Cambridge University Press.

- Reimer, M. 1992. Three views of demonstrative reference. Synthese, 93, 373-402.

- Romdenh-Romluc, K. 2002. Now the French are invading England! Analysis, 62(1), 34-41.

- Schiffer, S. 1981. Indexicals and the theory of reference. Synthese, 49, 43-100.

- Schiffer, S. and Neale. S. Forthcoming. How demonstratives and indexicals really work. In: S. Biggs and H. Geirrson (eds.), Routledge Handbook of Linguistic Reference. Abingdon: Routledge.

- Searle, J. 1983. Intentionality. Cambridge: Cambridge University Press.

- Speaks, J. 2016. The role of speaker and hearer in the character of demonstratives. Mind, 125(498), 301-39.

- Speaks, J. 2017. A puzzle about demonstratives and semantic competence. Philosophical Studies, 174(3), 709-34.

- Spelke, E.S., Breinlinger, K., Macomber, J., Jacobson, K. 1992. Origins of knowledge. Psychological Review, 99(4), 605–32.

- Sperber, D. and Wilson, D. 1986. Relevance: Communication and Cognition. Oxford: Blackwell.

- Stokke, A. 2010. Intention-sensitive semantics. Synthese, 175(3), 383-404.

- Stokhof, M. 2011. Intuitions and competence in formal semantics. Baltic International Yearbook of Cognition, Logic and Communication, 6, 1-23.

- Tomasello, M. 1995. Joint attention as social cognition. In: C. Moore and P. J. Dunham (eds.), Joint attention: Its origins and role in development, 103–30.

- Tomasello, M. 2003. Constructing a language: A usage-based theory of language acquisition. Harvard: Harvard University Press.

- Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. 2005. Understanding and sharing intentions: The origins of cultural cognition. Behavioral and Brain Sciences, 28(5), 675–90.

- Viebahn, E. 2018. Ways of Using Words: On Semantic Intentions. Philosophy and Phenomenological Research, 100(1), 93-117.

- Wettstein, H. K. 1984. How to bridge the gap between meaning and reference. Synthese, 58(1), 63-84.

- Wilson, G. 1989. The Intentionality of Human Action. Stanford, CA: Stanford University Press.

- Wimmer, H., and Perner, J. 1983. Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. Cognition, 13, 103–28.

- Wittgenstein, L. 2009. Philosophical Investigations. Oxford: Wiley-Blackwell.