



LUND UNIVERSITY

Bare demonstratives, joint attention and speakers' intentions

Brinck, Ingar

2002

[Link to publication](#)

Citation for published version (APA):

Brinck, I. (2002). *Bare demonstratives, joint attention and speakers' intentions*. Lund Philosophy Preprints.

Total number of authors:

1

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

BARE DEMONSTRATIVES, JOINT ATTENTION AND SPEAKERS' INTENTIONS*

INGAR BRINCK

Department of Philosophy, Lund University
Kunshuset Lundagård, SE-222 22 Lund, Sweden

ingar.brinck@fil.lu.se

www.fil.lu.se?name=ingar.brinck

Draft 2002/03/24. Please do not quote without permission.

keywords: demonstrative reference, attention, joint attention, speaker's intention, intentionalism

Abstract

Non-linguistic facts about the context of utterance relate demonstrative utterances to their context and can thereby fill the double role of providing the token meaning of the utterance and determining its reference. The reference of perceptual, bare demonstratives depends entirely on such facts. But contingent, causal, environmental factors are not reliable enough to take on this semantic role. An element that is within the control of the speakers themselves is required. It is argued that joint attention, based in the speakers' mutual co-ordination of their states of attention, is the default strategy. Intentionalists hold that the speaker's secondary or ultimate intentions help fixing the reference. But processing the speaker's underlying intentions is a costly strategy. A more economic one is to grasp the speaker's intentions implicitly by reading her manifest states of attention as displayed in interaction with the hearer, and make use of environmental cues. Unless the speaker explicitly indicates that the hearer should take the more resource-demanding strategy, there is no reason for her to do so. Moreover, if there is no general expectancy of hearers to give the underlying intentions priority, speakers will not in the normal case be using such intentions to get their messages across.

1. DEMONSTRATIVES

There are no particular restrictions on the kind of entities to which demonstratives ('this', 'that', 'these', 'those') can refer. Meaning supposedly determines reference, but in the case of demonstratives, this does not seem to be the case. The meaning of demonstratives is not specific enough. This characteristic makes it difficult to understand how they work. Below I will explore the suggestion that attention and joint attention are involved in determining the reference of a certain kind of demonstratives. I will first make some initial distinctions between different kinds of demonstratives, then focus on a particular subgroup, which I will call bare demonstratives.

As regards demonstratives, and also indexicals ('I', 'here', 'now'), it is common to make a distinction between the type, or linguistic, meaning and the token, or contextual, meaning. This can be done in different ways that each will have different effects at a technical or formal level. I will not discuss the technical issues, but make

* This manuscript was first presented at the conference of the ESPP (the European Society of Philosophy and Psychology) in Salzburg, September 2000.

what I hope to be a fairly intuitive and basic distinction between the linguistic and the contextual meaning of demonstratives.¹

Let the linguistic meaning be a semantic rule that fixes the application of the type expression. It determines to what the expression can be applied. For instance, the linguistic meaning of the indexical 'I' is that 'I' is used to refer to the speaker. The contextual meaning is that which the context of utterance contributes to the propositional content of the utterance. In the case of demonstratives and indexicals this will, on the view that such expressions refer directly, be the referent itself. The contextual meaning is determined by the linguistic meaning together with facts about the context of utterance.

The linguistic meaning of demonstratives is not rich enough to by itself determine the reference of a token expression. It indicates whether the reference is to a single object ('this', 'that) or a set of more than one object ('these', 'those'). It also indicates whether the referent is close in space and time ('this', 'these') or somewhat more distant ('that', 'those'). The conditions for reference expressed by the linguistic meaning are incomplete. Therefore, the reference of a particular token must be determined by the linguistic meaning together with the context of utterance.

Perry (1997, 593f.) distinguishes between pre-semantic and semantic uses of context. Context is used *pre-semantically* to determine the kinds of word, meaning, or syntactic structure an utterance has. When the context is used *semantically*, the meaning of the utterance exploits the context, in the sense that it depends on the context for its completion. Hence facts about the context of utterance are used to determine the propositional content of the utterance.

Another distinction can be made between the linguistic and non-linguistic facts about the context of utterance that determine the propositional content of an utterance. *The linguistic facts* are about the syntactic and semantic properties of particular utterances, and about the syntactic and semantic relations that hold between the utterances. Such facts help determine the reference in the following cases:

(1) Do you see that woman?

(2) Do you remember when Fred spilled coffee on Mary? - Yes, that was a very delicate situation.

¹A distinction between the linguistic and the contextual meaning of demonstratives is made in Kaplan (1989a). This distinction has been developed in several directions by other authors since it was first put forward by Kaplan in 1977. Perry (1977) advances a slightly different view than Kaplan, one which is considerably elaborated in Perry (1997). For my purposes, a basic distinction between linguistic and contextual meaning will be sufficient.

(1) is an example of how demonstratives are used in combination with sortal terms, a combination that results in *complex* demonstratives. (2) is an example of how a demonstrative can be used as an anaphor, by referring to something that has been mentioned previously. In both examples, the linguistic facts contribute to determine the reference of the demonstrative by relating it to other terms in the context of utterance.

I will call uses of demonstratives in which the reference is determined *solely* by linguistic facts *narrative*. The narrative use occurs when demonstratives are used to refer to past or future events that for some reason cannot be identified perceptually — for instance, by proxy — in the context of utterance. It also occurs when demonstratives are used in fiction and in imagination. When a name is introduced by being part of a narrative, the narrative prepares for future demonstrative reference to the object to which the name belongs.² In the narrative use of demonstratives, the demonstrative indicates a non-demonstrative referring term that occurs either in the context of utterance that contains the demonstrative, or in an utterance within the same narrative. This other referring term will in turn contribute to determine the reference of the demonstrative.

The narrative use can be contrasted with the *perceptual* use of demonstratives. The perceptual use occurs when the reference partly or wholly depends on *non-linguistic facts*. Such facts are about those properties of the context of utterance that although they do not pertain to language nevertheless play a semantic role. Among these properties are the relations that hold between an utterance and the context in which the utterance is made, a context which is itself external to any utterance. Examples of non-linguistic facts about an utterance are who the speaker is, when and where the utterance is made, whether or not it is made in combination with a demonstration, and, if it is, the character of the demonstration.

(1) is an example of the perceptual use of demonstratives. (1) contains a complex demonstrative of the form ‘that *F*’ where ‘*F*’ has an individuating function, but where ‘*F*’, not being definite, cannot alone determine the reference. Both linguistic and non-linguistic facts are involved in doing so. Linguistic facts about how ‘*F*’ relates to ‘that’ together with non-linguistic facts about how ‘that’ is related to the context determine the reference.

Bare demonstratives are a special kind of those demonstratives that are used perceptually. Bare demonstratives are used without attached sortals, descriptions, or

²A narrative can be defined as a series of utterances occurring over time and is identified by the theme or interrelated themes that it is about, for instance, talking about the summer holiday, about one’s neighbour, or about the best way to travel from Paris to London.

narratives that assist in pinning down the referent of the demonstrative. The following are examples of utterances that contain bare demonstratives.³

(3) Look at that!

(4) What are these?

(5) Go get those!

The facts about the context of utterance that contribute to determine the reference of the demonstratives in (3)-(5) are non-linguistic.

2. DEMONSTRATIVE REFERENCE

The role that the non-linguistic facts (henceforth NLF) play for demonstrative reference should be made more explicit. Let me compare how NLF contribute to determine the reference of bare demonstratives with how they contribute in the case of pure indexicals ('I', 'now', 'here'). As concerns pure indexicals, the NLF contribute by supplying the referent that will enter into the propositional content of the utterance which contains the indexical. They do so simply because the utterance is related to the context that contains the referent in accordance with the linguistic meaning of the word. The determination of the reference has its basis in NLF about the utterance.⁴

In the case of bare demonstratives, the NLF have to carry a heavier load. They must not only supply the referent that is selected by the linguistic meaning as in the case of pure indexicals. They must as well complete the linguistic meaning. This is due to the fact that, as noted above, the linguistic meaning of demonstratives *per se* is too vague to determine any particular entity. It only indicates the direction in space, and possibly time, in which the referent is to be found (nearby or further away), and whether the referent is one or several. The linguistic meaning says nothing as to the nature of the referent. The role of NLF is to determine the reference by providing a way of relating the utterance to the context in which it is made.

³Sometimes an utterance containing a bare demonstrative also contains an instruction for how to access or find the referent, conveyed by the verb that is used to direct the attention of the hearer to the object. In (3) the instruction is to "look", while in (5) the instruction is to "go and get", implying that the referent is at a distance, but within view. Such instructions seem to play a pragmatic role. They can give a hint about the characteristics of the referent, for instance, the referent of (3) may have especially salient surface properties, e.g. an intense colour or a strange shape. At other times the instructions do not pick out any salient features of the referent, but could be exchanged for other instructions that would work equally well.

⁴Hintikka (1998, 208f) describes how, in the case of pure indexicals, the reference is determined by the parameters that specify the speaker's spatiotemporal vantage point. This means that reference is conceptually determined. In the case of demonstratives, the reference cannot be determined in this way. However, in both cases, the referents are necessarily identified from the vantage point of the speaker.

An account of how reference with bare demonstratives works should describe how the demonstrative is related to the context. I will claim that the reference-determining relation between the utterance and its context is similar across contexts of utterance. This means that there is a certain kind of NLF about contexts of utterance in general that completes the linguistic meaning, and which together with the linguistic meaning selects the reference in particular contexts of utterance.

Provided that the context contains a speaker, an object, a time, and a place, it seems that in the normal case at least three conditions have to be satisfied for an act of demonstrative reference to be successful. These conditions are that the speaker can discriminate the object O, has her attention directed at O, and intends to refer to O by demonstrative D.⁵ Unless these conditions are fulfilled, reference will either not occur, or it will be merely accidental. Yet another condition can be added. It is that the speaker if she wishes can communicate the reference of the demonstrative to the hearer. If the demonstrative in fact is used to communicate, reference will be accomplished only when it is shared between speaker and referent. When it is shared, speaker and hearer will, on the view that demonstratives are directly referring, entertain the same propositional content.

The speaker can direct her attention towards the object in the context of utterance as the result of a prior intention to do so. A prior intention precedes the act of directing one's attention. Or she may have her attention unintentionally drawn to the object, either by the object itself or by some other element in the context. Then her attention is engaged by the context.

Is there any reason why the speaker could not determine the reference herself merely by intending to refer? Suppose that a speaker utters 'that' intending to refer to some entity in the context of utterance. Is there anything that prevents her from referring to the entity in this way? - Yes, there is.

The conditions that the speaker can discriminate the object and that she has her attention directed at it are necessary for the following reasons. In the case of bare demonstratives, there are no descriptive elements involved in the linguistic meaning that can identify or discriminate the object. But the object has to be discriminated for the utterance to attach to it. Whether the application of the demonstrative to the object is based in a causal link or in an intention to refer or in something else, if the object is

⁵This means that the speaker intends to use D in accordance with its linguistic meaning to pick out a particular object O. For the time being, I exclude cases in which the speaker's immediate intention to refer to O in context *c* comes apart from her prior intention to refer to a certain object independently of its spatiotemporal presence in *c*. The following is an example of how these intentions may diverge. The speaker has a prior intention to refer to Bill, but mistakes him for Bob in the context of utterance. Then the speaker will be referring to Bob by way of an immediate intention. Which intention finally determines the reference is discussed in sections 7-9.

neither discriminated nor attended to by the speaker, the demonstrative cannot be applied. The application depends on that the speaker at least has some idea about the nature of the object.

Since the use of bare demonstratives does not depend on a descriptive identification of the object, it depends instead on a perceptual one.⁶ The speaker must at least be able to perceptually discriminate the object in order to refer to it. Discrimination consists in a detection of dissimilarity or similarity between items (objects, properties, events, features, or anything else). Whether two items are discriminated or, on the contrary, perceived as similar, will depend on the context in which they occur and on with what else they are juxtaposed. Discrimination is an entirely contextual process. Identification, on the other hand, involves categorisation as well as long-term memory. Categorisation relies both on discrimination *between* instances of categories and identification of items *within* category boundaries. Discrimination is thus a weaker condition of reference than identification.

Now, if the speaker can discriminate the object, but does not attend to it, she will still not be able to apply the demonstrative to it. She has to focus on the object, or the reference will be either accidental or indeterminate or both. In the case of bare demonstratives, the linguistic meaning is not specific enough to indicate to which object the speaker refers, and there is not any narrative that can help determine the reference. If the speaker refers to an object without attending to it, there is no way that the reference can be fixed. Neither is there a way of telling which the object is, or even whether the object at all is present in the context of utterance. This is because discrimination is not definite in the sense that it, taken by itself, might individuate one and only one object in the absence of attention.

What about the fourth condition, the one of communication? Suppose that the speaker can only refer for herself, and cannot communicate the reference. If it thence is impossible for her to refer for somebody else (of course, it is not necessary that the speaker actually refers for somebody else), it seems to me that her use of the demonstrative is not linguistic. She might be using it in some psychological way, an account of which is beyond the scope of this article. I will therefore focus on acts of reference that satisfy the conditions of reference mentioned above including the condition of communication. The condition of communication should be such that it explains how the referent can be made manifest to a hearer.

⁶Hintikka (1998, 206) makes the interesting claim that there is a conceptual connection between demonstratives and perspectival identification, that is, identification that depends on the speaker's first-hand cognitive relations to her surroundings. The conceptual connection seems to depend on the fact that a demonstration is possible only if the speaker stands in cognitive contact with the demonstratum, for instance, by sight.

3. SALIENCY

The NLF that complete the linguistic meaning of demonstratives are about the way the utterance is related to its context. The way the referent is disclosed by the NLF must make the referent appear in a manner specific enough to determine the reference. Making the referent manifest means making it available for reference. There are several ways in which the speaker can indicate the referent for the hearer, for instance, by directing her gaze towards the intended referent, gesturing towards it, performing some action on it, or depicting it by an iconic, or mimetic, gesture. But the speaker can only intentionally communicate the reference to somebody else if she already has a grasp of the referent herself. As argued in section 2, the reference of demonstrative utterances depends on discrimination and attention. But how?

Sometimes the object is itself salient enough to determine the reference for both speaker and hearer. This happens in case the object appears in strong contrast to the surroundings, as a single blue car among ten red ones, or a dog among a flock of sheep. Then the context will automatically draw the attention of the subject to the object. Ballard *et al.* (1997) explain how, in a particular setting and in relation to a particular action, a subject can focus her attention on an object by a perceptual, deictic strategy that links sensory data with internal cognitive programs and motor behaviour. A certain sensory mode, like vision or audition, is used to identify and pick out the object in the context. The sensory mode that is used is the one which will be the most efficient in that setting, because its activation depends on salient features of the environment. For instance, when you are looking for your car in the parking-lot, your search may be guided by its colour. When you are looking for your child in the play-ground, you might listen for his or her voice. Usually the subject does not deliberately select a sensory mode in order to follow through a particular action. Rather, the sensory mode is engaged by the environment through salient features that automatically attract the attention of the subject.

In the examples mentioned in the last paragraph, the identity of the object is given from the start. The subject has a particular object in mind, and the question is which is the most efficient way to find it. This case can be contrasted with such that do not involve categorisation, or even discrimination, of the object, for instance, when action relies on so-called pragmatic representations of objects. Pragmatic representations do not necessarily entail binding of particular features or a discrimination of the entire object that is used to finish the action (Jeannerod 1994; Ballard *et al.* 1997).

Ballard concentrates on how reference is determined in relation to a single user or individual. But a deictic strategy may also be used in making the referent accessible

to the hearer in communication. Say that you are walking through the parking-lot looking for your bright pink car with a friend who has not yet seen your car. You suddenly utter

(6) That is the one.

Provided that colour is the most salient sensory mode in the context, and that the pink colour diverges sufficiently from the colours of the other cars, you may be quite confident that the hearer will grasp that it is to the pink car you refer. If the speaker does not demonstrate a particular object, the hearer will automatically turn her attention to those features that diverge the most from the rest of the context — if there are any. Of course, in this example, as in the ones above, the category (car) is given. It is implicit in the activity in which speaker and hearer are involved.

Colour is not a feature that in general is used to identify cars. It can help you single out your car in a particular context in which there are no similar objects with a similar colour, but it is not used for finding cars as such. The idiosyncratic and action-oriented nature of bare demonstratives makes them adapted to a particular purpose. But the question is how this way of referring becomes accessible to the hearer. It must be accessible, or the hearer will not be able to entertain the same propositional content as the speaker. In some contexts, as in the example, the selection of referent may be obvious. Yet in others, it may not.

Perceptual saliency is connected to how much effort it takes to perceive a feature. The most salient features are the ones that are picked up with the least effort. We do not need to look for them; they present themselves to us by attracting our attention. They contrast against our expectations on the actual context. Perceptual saliency is also at work when sudden movements or sounds disambiguate demonstrative reference. If the alarm goes on of one of the cars in the street, it is not hard to know what is demonstrated by the utterance

(7) That really is a nuisance.

But in many situations, we do not know what difference or contrast to look for. Then saliency will not do to make the referent manifest to either speaker or hearer.

It is not always evident which contrast is the relevant one in the particular context. There may be several competing features, or no particularly salient features present at all. Then we need something to guide our search for saliency. Moreover, different people may find different features salient, for instance, because of familiarity with the context, or simply because they have different backgrounds and thereby different expectations. Suppose that speaker and hearer find different features salient. Then, even if saliency would suffice to make the referent manifest for either speaker or hearer or both of them, it would not make the *same* referent manifest to them.

Saliency is not reliable enough to constitute the mechanism that determines reference by relating the utterance to its context. Contingent factors in the environment may simplify the use and understanding of acts of reference, but cannot take on a semantic role. We need an element that depends on the speakers themselves, which is within their control, and that does not make them dependent on environmental contingencies. It should be reliable enough to guarantee reference during normal conditions. In the next section, I will describe how attention might take on the role of determining the reference for the speaker. The idea is that the utterance is related to its context by the speaker's attentional state towards the object that the utterance is about. The further suggestion that joint attention makes the referent manifest for the hearer is explored in sections 5 and 6.

4. ATTENTION

Attention consists in an increased awareness of something either external or internal to the attending subject. It can be directed at behaviours, sensations, perceptions, or conceptions, and may be involuntary as well as deliberate. Attention is a complex phenomenon, as asserted by Pashler (1998). It involves many different kinds of processing — at different stages, such as the sensory, perceptual, and conceptual ones, and in different domains, such as cognition, memory, and behaviour. Parasuraman (1998) describes its purpose at allowing for and maintaining goal-directed behaviour. Several functions are involved in serving this purpose. Parasuraman specifically mentions selection of processing (due to computational limitations), vigilance (ensuring that goals are maintained over time), and control of the attentional processes (needed for planning, changing back and forth between different tasks, *et cetera*).

I will start by distinguishing between three kinds of attention, all of which may occur within a single perceptual system: scanning, attention attraction, and attention-focusing (Brinck 1997; 2001). The distinction is intended to capture the different functions of attention in relation to increasing degrees of control and complexity of processing. The claim is that the most complex kind, attention-focusing, is necessary for referring by way of bare demonstratives.

Perceptual *scanning* of the environment is continuous. Information is, as argued by Gibson (1986), registered in an active search directed at discovering possibilities to act. The information guides movements and causes actions "on-line" without interaction with long-term memory. Encoded stimuli have direct pathways to response centres. The latter trigger motor patterns without any intermediary categorisation of the information (Jeannerod 1994).

Attention is attracted by events that are at odds with what is expected on the basis of previous experience. It is often attracted by changes in one's surroundings, like

unexpected sounds and movements. Novelty or unfamiliarity induces a re-orientation of attention. The opposite case is when the subject has formed an expectation to encounter something so far not present. Then attention is geared to detect the expected item when, or if, it turns up. In order to plan for future needs, a subject can, for instance, learn to expect and attend to positive affordances as well as negative values on occasions when there is no immediate need for, nor any danger of them. Then the attention will react to features that indicate the presence of something worth focusing on.

The subject will focus on the event that has attracted her attention if it has a positive or negative value, and if it needs to be further determined before the subject can decide how to act on it. By *focusing* her *attention* the subject perceptually identifies the event. I will call this kind of attention *object-focused*. When attention is focused as a direct consequence of the attention's being attracted by a stimulus, the object of attention controls the attention. The attentional state is *goal-governed*.

Attention-focusing does not always depend on attention attraction. Attention may as well be focused as a result of a prior intention. Then the attentional state is *goal-intended*. It does not arise by being directly prompted by something external to the subject. Instead, the subject has an independent motivation to focus her attention, which is not an immediate function of attention attraction. The motivation may be triggered by different sorts of internal states, like emotions or imagery, that predate the perceptual encounter of the subject with a particular entity.

Of course, goal-intended attention does not exclude that the subject attends to a particular entity as a function of having learnt to attend to it in similar contexts. But for an attentional state to be goal-intended, it must be the case that, for one thing, the subject can have the same kind of attentional state in different contexts, and, for another, that different attentional states may occur in the same kind of context. If not, the situation will be one in which the external environment prompts or triggers the attentional state, perhaps as a result of habituation.

In section 2, several conditions for an act of demonstrative reference were introduced. The second condition is that the speaker has her attention directed at an object. The attentional state of the speaker can now be further characterised. Demonstrative reference requires that the speaker focuses her attention on the object. Scanning and attention attraction are not sufficient. The reason why is that at least discrimination, as it was described in section 2, is necessary for acts of demonstrative reference. And discrimination of the object occurs only at the stage of attention-focusing.

At the earlier stages of attention, the information that the subject reacts to is not processed to the extent that the subject represents it as being about something in particular. Neither is the information kept in memory. The link between stimuli and

action is direct. Then the subject is not able to apply the demonstrative intentionally, that is, to choose whether to apply it or not. Without discriminating the object, she will only be able to apply the demonstrative as a response to stimuli.

Furthermore, in order to be a competent speaker, the speaker must be able to focus her attention on the object in a goal-intended way. The reason is that otherwise she would not be able to attend independently of the context, but her attention would always be controlled by the environment. Then her language-use would be driven solely by factors external to herself. That the speaker is able to attend in a goal-intended way does not, of course, exclude that she sometimes in fact attends in a goal-governed way.

To sum up, the NLF that concern the relation between the utterance and its context are facts about those attentional states of the speaker that occur during attention-focusing. The utterance is related to its context by the attention of the speaker, and the attention, together with the linguistic meaning, determine the reference of the utterance.

5. CO-ORDINATION OF ATTENTION

Attention can also be *subject-focused*, that is, focused on other subjects, and based in an observation of the behaviour of others. By, say, looking at each other, two subjects can detect their respective objects of attention. They do so on the basis of the (spatial) direction of their respective movements taken in combination with a salient item that functions as a target. By co-ordinating their attention on the basis of *behavioural co-ordination*, the subjects can focus on the same thing simultaneously, that is, engage in *mutual object-focusing*.

But mutual object-focusing does not take us all the way to joint attention, at least not the kind of joint attention required for demonstrative reference. The reason is twofold. Mutual object-focusing is behaviourally based, and it is not goal-intended. The following is an example of how it works. Suppose that a group of tourists engage in mutual object-focusing as a result of behavioural co-ordination and contextual saliency. Mutual object-focusing spreads automatically among them, as a function of attention attraction. The process starts when the guide attracts the attention of one of the tourists, and shows her a highly unusual and thereby salient building. The tourist goes on to focus her attention on the building. The other tourists notice the behaviour of the first tourist and turn their attention in the same direction as she has done. Having done so, they will automatically focus on the building because of its saliency. Mutual object-focusing occurs as a result of the tourists' having had their attention attracted by the first tourist and of the presence of a salient object in the direction in which the tourists turn.

Attentive behaviour can spread among a group of subjects without resulting in mutual object-focusing. Say that an unusual kind of bird attracts the attention of one of

the tourists, and she focuses on it. The other tourists co-ordinate their behaviour with hers, looking in the same direction as she does. Suppose that the bird by then has flown away, or is no longer salient in a way that distinguishes it from the surroundings. Then the attentive behaviour of the group as a whole will be co-ordinated, without there being a common object for them to focus on. There is a behavioural co-ordination, but no mutual object-focusing.

One reason for introducing joint attention into the discussion of demonstrative reference is that it serves to disambiguate reference when salient contextual features are missing. Mutual object-focusing cannot disambiguate reference in such circumstances precisely because it is based in behavioural co-ordination. Behavioural co-ordination is a consequence of the attention's being attracted by salient contextual features, like movements or sounds. When these features disappear or fade away, the attention of different subjects cannot be mutually focused. To take on the role as a determinant of reference, joint attention must be based in *attentional co-ordination*. Directing somebody by gaze or in some other way displaying one's attentional state is a way of making one's focus of attention accessible to others in the absence of contextually salient objects.

6. JOINT ATTENTION

Joint attention occurs when two or more subjects, as a consequence of attending to each others' attentional states, focus their perception simultaneously on a common stimulus. Joint attention can be described as *subject-subject* attention. It is not necessary that the subjects share the location in space, but simultaneity does seem crucial. For instance, it is quite possible to refer by a bare demonstrative over the phone. Say that a sound as of an explosion is heard in the vicinity of Anne. Anne is talking to Jane, who is living in another town, on the phone, and Anne exclaims

(8) Did you hear that?

Provided that Jane heard the sound through the receiver, reference is successful. Anne and Jane both attend to the sound, which was salient enough to draw their joint attention.

On the other hand, it seems impossible to keep up the reference of a bare demonstrative when an earlier context of utterance is separated from a later one. If the context of utterance in which the reference is made does not follow immediately upon the earlier context to which it supposedly is related, there will not be any obvious causal connections that relate the earlier use to later ones. Then the speakers will not be able to track the referent perceptually, without using general concepts, by following its trajectory through space and time.

If communication is to be successful, speaker and hearer must focus on the same referent, that is, a common stimulus. It seems reasonable to say that a referent is shared when, in a particular context (a time and a place), it is mutually accessible and accessible in the same way for speaker and hearer. Since the reference of a bare demonstrative is determined perceptually, it also seems reasonable to say that its referent is shared when speaker and hearer can interact with it in a similar way on a perceptual basis. Note that shared referents cannot be constituted by the sense impressions of individual subjects or by proximal stimuli, but must exist in a common world. Otherwise they would not be intersubjective and publicly accessible.

That the referent is determined perceptually does not exclude that it, once determined, can be conceptually categorised. The speaker may use a perceptual strategy to direct the attention of the hearer onto the referent, while the hearer identifies the referent conceptually as soon as she has focused her attention on it.

Several authors have emphasised that joint attention requires not only that the subjects attend to each others' attentional states, but that they attend to those states in a certain manner (Bruner 1998; Gómez *et al.* 1993; Gómez 1998; Tomasello 1998; Tomasello 2001). The subjects should grasp that the attentional states are directed at an object and, by being so directed, are about that object. This condition ensures that the subjects do not only attend to each others behaviour, but to the intention behind the behaviour. Attending in this way may result in attention contact between the subjects. However, this condition, I maintain, is not sufficient for joint attention in the absence of salient objects.

Furthermore, the subjects will have to attend to each other as capable of attending in a goal-intended way, that is, in a way that is not controlled by the object of attention. The latter condition is important in order to distinguish subject-subject attention as it occurs in joint attention from mere subject-attention occurring in mutual object-focusing. Goal-intention provides for the capacity to either direct or follow the attention of the other subject in the absence of salient objects. This means that attention-focusing can be guided by the subjects' mutual attention to each other alone, and not by the environment.

The claim is that, in the absence of a salient referent, shared reference is achieved by joint attention. The account in terms of joint attention satisfies the conditions on reference by bare demonstratives that were put down in section 2. To repeat, these are that the speaker discriminates the referent, attends to the referent, intends to refer to it, and can communicate the reference to the hearer.

The speaker succeeds in communicating the reference by engaging in joint attention with the hearer. She attracts the attention of the hearer by her utterance and a gaze or perhaps a gesture (or both). She then uses her gaze to direct the attention of the

hearer onto the object by alternating the gaze between the object and the hearer. She intends to make the audience follow the demonstration to its target.

The intention to make the hearer follow the demonstration to its target is, it seems, an intention to make the hearer do something. It has an imperative force. The aim of the intention is to share the object of attention. This intention may in turn be a tool for making the hearer do something else, for instance, share the speaker's attitudes or reactions to the object, get the object for the speaker, or participate in an activity together with the speaker.⁷

In section 3, I described how bare demonstratives are used to initiate and sustain action. They do not state something about the referent, but show or display it, inviting the hearer to respond. Bare demonstratives are particularly useful for tasks that exploit implicit knowledge and contextual, idiosyncratic ways of taking the object of attention. The reason is that they can be used as pointers or tags that keep the object of attention in focus over time until the goal of the action is achieved. This holds both in case there is a single subject or in case there are two or more subjects jointly attending to the object.

7. IMMEDIATE AND ULTIMATE INTENTIONS

Many authors have stressed the role that intentions play for demonstrative reference. In short, the idea is that the speaker's intentions determine the reference of the demonstrative utterance. Two general arguments that can be raised against the theory that meaning is determined by intentions are that the theory makes meaning arbitrary, and that it makes meaning private, or at least dependent on the mental states of particular individuals.

Wettstein (1984) argues against intentional theories of meaning on these grounds. He adopts the view that language is a social institution, and that demonstrative reference is determined by publicly accessible features like contextual cues. But which would these cues be when it comes to bare demonstratives? The cues that Wettstein refers to are such as sortal terms, an ongoing discussion between speakers, *et cetera*. Such cues, I agree, are exploited in perceptual or narrative uses of complex demonstratives. But they can obviously not be exploited in *purely* perceptual uses of demonstratives.

⁷The distinction between the intention to make the hearer follow the demonstration to its target and any further intentions the speaker may have with her demonstration reminds of the distinction between illocutionary and perlocutionary speech acts introduced by Austin (1962). The illocutionary act consists in putting forward a propositional content in a certain way, for instance, as an assertion, question, or warning. The perlocutionary act concerns those additional effects the entire speech act may have on the hearer's thoughts or actions.

Furthermore, sortal terms do not always determine the reference. They can be used in combination with demonstratives without serving a disambiguating role. This happens if the sortal term does not add any information to the context of utterance in addition to that which is already given by perceptual means. Any sortal term that is true of a particular referent will not do to determine the reference of the demonstrative. The sortal term must add enough information for the hearer to be able to discriminate between the referent and its surroundings. An utterance of

(9) That dog is Fido.

made together with a demonstration in the direction of a group of five alsatians will not help the hearer pick out the referent. The utterance does not single out a feature of Fido that distinguishes him from the other dogs. This means that in this case the hearer cannot rely on linguistic cues to understand to what the utterance refers.

Sometimes there is an ambiguity concerning the speaker's intention as to which referent is picked out by a demonstrative. A common intentionalist explanation of how demonstratives refer in these cases relies on distinguishing between two kinds of intention, one *immediate* that consists in making the demonstration (for instance, pointing at an object), and one *ultimate* that consists in the intention to demonstrate a certain object. They can occur independently of each other.

The speaker's intention in the context of utterance to direct the attention of the hearer to an object is an immediate intention. The prior intention, if there is one, that makes the speaker produce the immediate one constitutes the ultimate intention. Say that Anne demonstrates John while saying his name. The immediate intention is expressed by her uttering his name and pointing to him. The ultimate intention is the underlying intention to demonstrate John that causes the immediate intention.

The distinction between immediate and ultimate intentions comes into play when the demonstration is ambiguous or fails. It can be maintained that the ultimate intention will fix the reference in such cases. Say that Anne wants to demonstrate John. She thinks that he is the man dressed in blue and thus points to that man in order to refer to John, while uttering

(10) That is John.

But John happens to be dressed in white. Anne's immediate intention lies behind her pointing to the man in blue, Mike, while her ultimate intention is to point to John. Nevertheless, I do not think that the ultimate intention fixes the reference of the demonstrative in this case. Anne tries to fulfil her ultimate intention by pointing, but fails to do so. The immediate intention constitutes the referential act. The man she actually refers to by the demonstrative is Mike.

If I want to deny that the ultimate intention fixes the reference, why do I not simply say that Anne failed to refer? After all, Anne made the hearer direct her attention to somebody else than Anne herself ultimately intended. The reason why is that I believe that the shared reference of demonstratives is determined by the cues that the speaker intentionally makes accessible to the hearer, and that are in fact accessible in the context of utterance. If the speaker wants to make herself understood, she will try to convey information that will help the hearer interpret her. The hearer, for her part, will use the information that is imparted in the context — unless there is something that tells her not to do so. However, if there is a conflict between information that is easy to retrieve in the context and such that is difficult to retrieve, it seems reasonable that the hearer will use the information that is more easily retrieved. The way of interpreting the speaker that demands the least effort, but still gives a pay-off in terms of a referent, will be the most attractive.

In (10) the demonstrative bears relations to NLF about the context of utterance. Suppose the hearer paid attention to the cues that were made accessible by Anne. Anne's pointing gesture together with her utterance will then suffice to determine the reference, although the referent will not be the person Anne had in mind earlier on. In the narrative use of demonstratives, as opposed to the perceptual one, ultimate intentions may be important to grasp the reference. If the ultimate intention is made accessible by a narrative, and the use of the demonstrative occurs as a part of the narrative, the ultimate intention overrides the immediate intention of the speaker.

The following scenarios exploit the possibility of there being a narrative. Suppose that Anne has expressed her wish to show us John. She has also told us that John is a famous tennis-player, always dressed in white. However, when she finally demonstrates him, she happens to forget about the way he is dressed, and points to the man in blue. But from what she has told us, we will nevertheless understand that she does not intend to point to Mike. Alternatively, Anne might have told us that John is a famous tennis-player who once in a while wears white. She has promised to show him to us when he turns up. Later on she points to a mediocre player in white clothes and calls him John. Again the ultimate intention will cancel the immediate one. The most plausible conclusion — granted that we remember what Anne told us about John and her promise — is that Anne pointed to the wrong man, and that reference failed.

Let us have a look at a few more examples. The following one was put forward by Reimer (1992, 385) in order to show that demonstrations can cancel conflicting intentions. Say that you have left your keys next to Anne's on the table. While focusing on and reaching for your keys you utter

(11) These are mine.

Then you look at the keys in your hand, and discover that they are not yours. As it happens, you are holding Anne's keys in your hand. Which keys did you refer to by (11)?

It seems to me that by holding Anne's keys in your hand while uttering (11), you refer to them. Any other account appears *ad hoc* given the NLF about the context of utterance. The immediate intention together with the demonstration fixes the reference. This means that you are saying something false about Anne's keys, although you will notice the mistake yourself as soon as you look at the keys in your hand.

In the pure perceptual use of demonstratives, as in (11), there are no other clues for interpreting an utterance but the means which the speaker uses to draw the hearer's attention to the referent. In the example, the speaker visibly is reaching for Anne's keys and is directing the attention of the hearer to them by making the utterance. There is nothing that indicates to the hearer that she should turn to ultimate intentions in order to interpret the utterance.

Another example put forward by Reimer (1982, 380) is meant to show that intentions rule out contextual cues of Wettstein's kind. The two dogs Spot and Fido are standing next to each other. They look similar. The speaker ultimately intends to say something about Fido, but during the perceptual encounter with the dogs, she happens to mistake Fido for Spot. She nods vaguely in the direction of both dogs with the immediate intention to indicate Spot, and utters

(12) That's your dog Fido.

Did she say something false about Spot, or perhaps something true about Fido? Reimer opts for the first reply, since it was the speaker's immediate intention to demonstrate Spot. But this does not seem right, although the speaker in principle can fix the reference by perceptually attending to Spot. The aim of the speaker is after all to share the reference with the hearer. In the example, shared reference cannot be achieved by the immediate intention alone. It cannot determine the reference in the absence of a proper demonstration. But to say that the speaker said something true about Fido seems equally wrong. She did intend her nod to indicate Spot, although the ultimate intention that caused the nod was to indicate Fido.

Most certainly the owner of Fido will be able to identify her dog, but she will not be able to judge whether the speaker's attempted identity statement was true or not. The reason is that she does not know which of the dogs the speaker was intending by uttering 'that'. Since the nod is vague, it seems reasonable to conclude that the speaker fails to direct the attention of the hearer to any particular dog.

There is no accessible evidence about the speaker's ultimate intention in the context of utterance. The ultimate intention cannot come to the rescue of the immediate intention. Neither is there anything in the example that indicates a wider, narrative use

of the demonstrative. This means that the only evidence made available for the hearer in the context of utterance is the utterance itself. The context as such is not salient enough to make it explicit whether the speaker nodded to Fido or Spot, although, of course, the nod directs the attention of the hearer onto the dogs. The result is that the demonstrative lacks a referent in the context. One might say that the utterance misfired, because the demonstration failed to indicate the demonstratum. From examples (10)-(12) we can, it seems, conclude that ultimate intentions do not play any role in fixing the reference of demonstratives in the perceptual use.

8. DEMONSTRATION AND SECOND-ORDER INTENTION

Kaplan (1989a; 1989b) has provided us with an example in which demonstration and description supposedly come apart. He describes himself pointing to a picture behind him and uttering

(13) Dthat is a picture of one of the greatest philosophers of the twentieth century.

Unbeknownst to Kaplan, the picture of Carnap that he intends to point to has been changed for one of Spiro Agnew. The question is to which picture Kaplan is referring. If we take the pointing to fix the reference, it means that the demonstrative refers to the picture of Agnew. If we take Kaplan's ultimate intention to talk about Carnap to fix the reference, the demonstrative refers to the picture of Carnap.

Obviously, Kaplan does not have Agnew in mind when he makes the utterance. Given that the audience knows something about the background and views of Kaplan, it will probably take the description as an indication that Kaplan at least did *not* intend to demonstrate the picture of Agnew. The narrative in which the example is set, David Kaplan's talking about a great philosopher, makes it appear wrong, or even incoherent, to say that he is referring to the picture of Agnew.

Intuitions tend to go in different directions as concerns this example. The speaker aims to identify one of the greatest philosophers of the twentieth century by pointing to a portrait of the philosopher. The demonstration is not vague as it is in (12). On the other hand, it is based on old, and at the time of the utterance faulty, knowledge. The picture has been changed. Does a demonstration succeed if made without the speaker's having directed her attention to the object she is demonstrating and having checked its identity?

If one believes that the reference of demonstratives in the perceptual use is based in attention, the answer is straightforward. Kaplan did not succeed in referring to the picture of Agnew, since he did not attend to what he was demonstrating. He did not intend the picture that he happened to demonstrate. But neither did he intend to demonstrate anything else by his pointing, since he did not attend to any particular

object. He did simply not succeed in stating a complete proposition by uttering (13), because the demonstrative was not aimed at any object.

The normal reaction when somebody makes a demonstration without perceptually attending to something in the direction of the demonstration is not to take the demonstration as authoritative. A demonstration is not only a gesture, neither is it a gesture accompanied by some prior intention. The object that is demonstrated should also be the one that the speaker attends to while producing the demonstration. There are two reasons for this.

First, for a movement to at all be perceived as a demonstration, it has to be accompanied by an appropriate attentional state of the speaker, one that is perceivable. There must be something that signals to the hearer that the movement should be taken as a demonstration. Second, a demonstration will not be taken as a way to call attention to anything in particular if the person behind the demonstration does not show any interest in something that the demonstration might be aimed at. If a possible target for the demonstration cannot be detected, the demonstration will appear vacuous.

If one, on the other hand, takes the demonstrative in example (13) to be used narratively, the intuitions may differ. Then the demonstration will be taken as a gesture without semantic import. But since the description in (13) is not uniquely determining in the absence of the demonstration and, moreover, there is nothing that satisfies it in the context of utterance, it does not succeed in picking out either Carnap or Agnew. The conclusion from the narrative interpretation of (13) is that Kaplan did not intend to refer to Agnew, since Agnew is not considered to be one of the greatest philosophers of the twentieth century. On this interpretation Kaplan did not succeed in referring.

The major problem with (13) is that the demonstration that accompanies the utterance is faulty in the technical sense described above. (13) is not really an example of demonstration and description going in different directions, but one in which the demonstration is vacuous, and the description consequently becomes incomplete. If Kaplan had perceptually attended to the picture while pointing, then he would have said something false. Agnew is not one of the greatest philosophers of the twentieth century.

Reimer (1992) defends a view that she describes as quasi-intentional and that she holds will account for Kaplan's example. It takes into account both contextual cues and intentions. The idea is that there is a disambiguating secondary intention that plays a semantic role if the actual and the intended demonstrata diverge. In such ambiguous cases the secondary intention serves to connect the demonstratum entertained in the ultimate intention with the demonstrative act that relies on the immediate intention. The connection is made through certain beliefs about the object of the demonstrative act. The reference is in the last instance determined by the second-order intention to

demonstrate and say something about a particular object which the speaker believes is identical to the object that she demonstrates.

Reimer holds that when Kaplan points to the picture of Agnew while intending to point to the one of Carnap, the belief that plays a semantic role is the belief that the picture in the direction of the gesture is identical to the picture Kaplan has in mind. When this belief is conjoined with the ultimate intention to refer to the picture of Carnap, it will result in the secondary intention to demonstrate and say something of the picture that is demonstrated. Thus, Reimer maintains, the picture of Agnew is demonstrated. Since the belief that the actual picture is identical with the one Kaplan has in mind is a false belief, Kaplan utters something false.

Reimer gives a plausible description of what is going on in the mind of the speaker when the actual and the intended reference diverge. In these cases the beliefs that the speaker has about the intended referent do not correspond to those she has about the object that she actually demonstrates. There is a mistaken presumption of identity between the intended and the actual object. Moreover, the speaker's beliefs about the intended object lie behind her desire to demonstrate the object, an object that actually happens to be another one than she had in mind. Reimer's account works as an explanation of where the speaker goes wrong. But that does not mean that it works as an explanation of the general traits of demonstrative reference. I doubt that hearers have recourse to secondary intentions in interpreting the kind of ambiguous demonstrative utterances we have been considering. In fact, I doubt that these intentions even are retrievable from such contexts.

9. COST AND EFFECT

It seems that distinguishing between different kinds of intention is not of much help to understand how reference works in the problematic cases we have considered. Underlying intentions, such as ultimate or secondary ones, are supposed to help when the demonstration is ambiguous or fails. Why don't they?

For underlying intentions to be of any use, there must first of all be something in the context that indicates to the hearer that she should have recourse to such intentions, and not rely on the information that is manifest in or displayed by the context. This indication should be provided by the speaker. If the hearer takes other intentions than the immediate one into account, regardless of whether the speaker noticeably indicates to her to do so or not, the hearer runs the risk of misinterpreting the speaker, or engaging herself in mere guesswork. It is not evident that there always is an underlying intention, nor is it evident that the underlying intention always should override the immediate one.

Suppose that the speaker intends the hearer to turn to the underlying intention. Then, provided that the speaker is rational, and that she wants to be relevant in the sense of maximising effect and minimising effort for both herself and the hearer (Sperber and Wilson 1986), she would try to make this intention available to the hearer. She would supply some kind of information in the context of utterance that would make it possible for the hearer to pass from the demonstration and the immediate intention to the underlying intention.

Such information can be provided in several ways. Suppose Anne again demonstrates Mike while uttering

(14) That is John.

this time while distinctly winking her left eye. The hearer will presumably take the unexpected, but distinct, winking as an incitement to start searching for an underlying intention behind the utterance. Depending on the context, different intentions will be attributed to Anne, and result in different interpretations of the utterance. Anne might, for instance, intend the demonstration to be understood ironically. She knows that John never would wear blue, but only wears white, and by indicating Mike while saying ‘John’, she makes an ironic joke about John and his dress code.

There are a few things worth considering concerning those cases in which the speaker attempts to make her underlying intentions accessible to the hearer. For one thing, it becomes necessary for the hearer to reason about the intentions of the speaker in order to understand what she intends to communicate. Using one’s underlying intentions to convey a message exploits another strategy than the one that only relies on the immediate intention. The former strategy seems to be more costly for the hearer and take more of an effort to process. Moreover, it is only accessible to subjects capable of reading the intentions, as opposed to the attention, of others. Suppose that one can only understand demonstratives if one can attribute underlying intentions to the speaker. That means that less than four-year-old children — who up to that age have not yet developed the theory of mind required for intention-reading — would not understand any demonstratives. That does seem counterintuitive.

I believe that unless the speaker explicitly indicates that the hearer is supposed to take the more resource-demanding strategy, it is not used by the hearer. Moreover, if there is not any general expectancy of hearers to give the underlying intentions priority, speakers will not in the normal case be using such intentions to get their messages across. Joint attention is, I submit, the default strategy for understanding demonstratives in the perceptual use. The cues that determine shared, demonstrative reference are publicly accessible without in every case being linguistic.

References

- Austin, J.L.: 1962, *How to Do Things with Words*. Oxford University Press, Oxford.
- Ballard, D., Hayhoe, M., Pook, P. and Rao, R.: 1997, 'Deictic Codes for the Embodiment of Cognition', *Behavioral and Brain Sciences*, 20(4), 723-767.
- Brinck, I.: 1997, *The Indexical 'I'*. Kluwer, Dordrecht.
- Brinck, I.: 2001, 'Attention and the Evolution of Intentional Communication' *Pragmatics & Cognition* 9(2), 255-272.
- Bruner, J.: 1998, 'Routes to Reference' *Pragmatics & Cognition*, 6(½), 209-227.
- Gibson, J. J.: 1986, *The Ecological Approach To Visual Perception*. Lawrence Erlbaum Ass., Hillsdale, NJ.
- Gómez, J. C.: 1998, 'Some Thoughts about the Evolution of LADS, with Special Reference to TOM and SAM', in P. Carruthers and J. Boucher (eds.), *Language and Thought*. Cambridge University Press, Cambridge, pp. 76-93.
- Gómez, J. C., Sarriá, E., and Tamarit, J.: 1993, 'The Comparative Study of Early Communication and Theories of Mind', in S. Baron-Cohen, H. Tager-Flusberg and D. J. Cohen (eds.), *Understanding Other Minds. Perspectives from Autism*. Oxford University Press, Oxford, pp. 397-426.
- Hintikka, J.: 1998, 'Perspectival Identification, Demonstratives and "Small Worlds"', *Synthese*, 114(2), 203-232.
- Jeannerod, M.: 1994, 'The Representing Brain: Neural Correlates of Motor Intentions and Imagery', *Behavioral and Brain Sciences*, 17(2), 187-202.
- Kaplan, D.: 1989a, 'Demonstratives', in J. Almog, J. Perry and H. Wettstein (eds.), *Themes from Kaplan*. Oxford University Press, Oxford, pp. 481-563.
- Kaplan, D.: 1989b, 'Afterthoughts', in J. Almog, J. Perry and H. Wettstein (eds.), *Themes from Kaplan*. Oxford University Press, Oxford, pp. 565-614.
- Parasuraman, R.: 1998, 'Introduction', in Parasuraman, R. (ed.), *The Attentive Brain*. MIT Press, Cambridge, Mass., pp. 3-15.
- Pashler, H.: 1998, *The Psychology of Attention*. MIT Press, Cambridge, Mass.
- Perry, J.: 1977, 'Frege on Demonstratives', *Philosophical Review*, 86(4), 474-497.
- Perry, J.: 1997, 'Indexicals and Demonstratives' in C. Wright and B. Hale, B. (eds.), *A Companion to the Philosophy of Language*. Blackwell, Oxford, pp. 586-612.
- Reimer, M.: 1992, 'Three Views of Demonstrative Reference', *Synthese*, 93, 373-402.
- Sperber, D. and Wilson, D.: 1986, *Relevance*. Harvard University Press, Cambridge, MA.
- Tomasello, M.: 1998, 'Reference: Intending that Others Jointly Attend', *Pragmatics & Cognition*, 6(1/2), 229-243.
- Tomasello, M.: 2001 'Perceiving Intentions and Learning Words in the Second Year of Life', in M. Bowerman and S. Levinson (eds.), *Language Acquisition and Conceptual Development*. Cambridge University Press, Cambridge, pp. 132-158.
- Wettstein, H.: 1984, 'How to Bridge the Gap between Meaning and Reference', *Synthese*, 58, 63-84.