Demonstratives in Dutch Cooperative Task Dialogues

Paul Piwek, Robbert-Jan Beun and Anita Cremers

08.12.1995

IPO Manuscript no. 1134, Eindhoven University of Technology

In this paper, we present a model of dialogue situations in which the basic dialogue parameters are carefully distinguished. From this model we constructed a dialogue set-up to investigate in what respect the deictic use of the Dutch proximate demonstrative (dit, deze; this) differs from that of the Dutch distal demonstrative (dat, die; that). Central in previous studies on the anaphoric use of demonstratives is the notion of 'force', which consists of the strategies 'givenness', 'noteworthiness' and 'foregrounding'. We show how the three strategies developed for the anaphoric use can be generalized to the deictic use of demonstratives. On the basis of empirical data obtained from recorded dialogues, it was found that proximate demonstratives are used especially in cases where a speaker wants to signal to the addressee extra force to find the intended referent. Distals are often used in cases where the referent is more 'given' in the addressee's consciousness. The results also help us to explain how the use of demonstratives is related to the use of pointing and postmodification in a dialogue situation. Our results differ, however, from those that have been reported in the literature on the use of the English demonstrative. Finally, we show how the parameters of the model of our dialogue situation can be used to provide a generalization of our results.

1. Introduction

In any natural language dialogue system, fundamental decisions about the system's concrete implementation have to be taken in the early stages of development. These decisions concern, for instance, the primitives and components that have to be included as building bricks of a natural dialogue situation and, more specifically, aspects such as the coverage of the language fragment, the basic rules for anaphora resolution and, if any, the mental states of the dialogue participants (such as the system's belief, knowledge and intentions). These decisions are often based on the intuition of the system developer, but nowadays there is a widespread opinion that developing such systems requires at least a profound observation of the language user in his or her particular environment.

Having decided that real data about the user's linguistic behaviour plays an important role in the development process, we are faced with the following complications. Firstly, if there are no data available, we have to design an experimental situation that resembles the actual use of the system. But, since the system is not finished yet, this has to be simulated in what are often called 'Wizard of Oz' experiments where a human operator takes over the role of the dialogue system. Second, if there are already data available, we do not know whether the data are transferable to the specific situation in which the system is used.

If we are unaware of the basic parameters in a dialogue situation, often we do not know how the results can be generalised and transferred to situations other than the specific one from which the data are extracted. Much of the research on anaphora resolution, for instance, is based on textual data without direct interaction from an interpreter (see e.g. (Carter, 1987)). We therefore do not know whether, and if so how, these results should be transferred to an actual dialogue situation. And, even if we consider a real dialogue situation, we have no certainty about the transfer of our results to the actual application. For instance, can we transfer the results from Grosz's task dialogues about installing a pump to a dialogue about an electron microscope? Can we compare results from a telephone conversation to a conversation by E-mail? Some of the results can be transferred, others cannot; knowing when and why this is the case, however, requires careful modelling of a dialogue situation.

The goal of this paper is therefore twofold. Firstly, we would like to present an experimental design in which the basic dialogue parameters are carefully distinguished. Secondly, we would like to show how the experimental design can be applied in the use and function of Dutch determiners in task-related dialogues, with a particular focus on Dutch demonstratives. The experimental design and its parameters are directly derived from a general architecture for user-interface dialogue systems (e.g. (Hutchins, 1989)). The results of the experiments are applied in the DenK system, a generic cooperative interface combining natural language and visual interaction (Ahn et al., 1995).

This paper will proceed as follows. In the next section we discuss the general user-interface architecture and explain how the dialogue parameters presented in this paper are related to this architecture. In section 3, we will discuss the anaphoric use of demonstratives and suggest how these results can be generalized to their deictic use. In section 4, the set-up of the dialogue situation will be discussed in which the hypotheses are tested that follow from section 3. Finally in sections 5 and 6, the main results will be presented and we will try to show how the results can be generalized to other dialogue situations.

2. The Triangle Metaphor

Underlying the combination of linguistic and visual modalities in the architecture of the DenK system is the recognition that humans interact naturally with their environment in two ways: *symbolically* and *physically.* On the one hand, if there is an intermediary agent (human or electronic assistant), humans can interact symbolically and use language to give commands, ask for or provide information etc. On the other hand, physically, one picks up objects, moves or fastens them, or observes them by seeing, feeling, hearing, or smelling. The essential difference between the two types of interaction is that actions of the first type (for instance, speech acts) need an interpreter who can bridge the gap between the symbols and their actual meaning and purpose, while actions of the second type are related in a more direct manner to human perception and action and hardly need interpretation. This is depicted in what we call the 'triangle metaphor' (Figure 1).

Rudimentary though it is, we think that the triangle metaphor contains the basic ingredients for modelling a realistic dialogue situation. At the corners of the triangle, we recognize the basic components: the two participants and the application domain; the sides represent the information flow (or interaction) between these components. It is important to note that the triangle does not reflect the dialogue's spatial properties. Both participants may be 'inside' the domain of discourse and, as a result, their view on the domain may change constantly.

We will first consider the basic parameters more closely and focus on the aspects

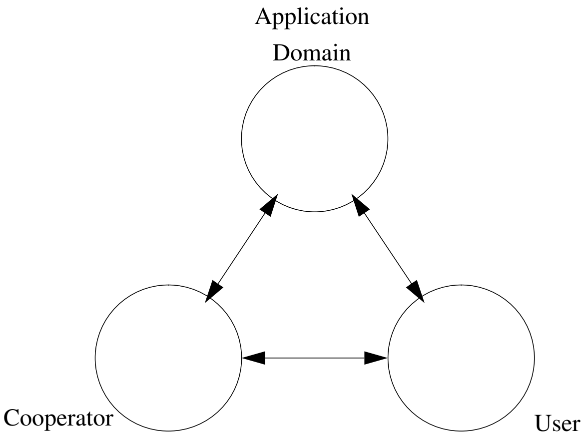

Figure 1
The triangle metaphor: user, cooperator and application domain

relevant for the use of demonstratives:

The application domain. The application domain (or domain of discourse) contains the objects that can be manipulated and observed via direct interaction. It is complete and consistent, i.e. all properties and relations between the objects are determined and these properties can never be inconsistent. This point clearly distinguishes the domain from the mental states of the dialogue participants (see below).

The first distinction we make is between dynamic and static domains, i.e. the ability of the properties of objects to change in time, or not, during the conversation. For instance, in a train table domain, the arrival and departure times of the trains usually remains constant; in a blocks-world domain, blocks may be removed, colours may be changed etc. Consequentely, in a static domain no domain actions can be performed by the participants, and beliefs resulting from observations always remain true during the conversation.

Another property of the domain concerns the type of objects present in the domain. In an electron microscope domain, the objects, such as lenses, apertures and buttons, clearly differ from each other. In other domains, such as a chemistry or a blocks world, objects may be quite similar. Consequently, the referential behaviour of the participants may be dramatically different. In an electron microscope domain, unique names are widely used (e.g. 'the OA aperture', 'the on/off button'); in a chemistry domain or a blocks world, one would expect a more extensive use of focussing principles and relatum objects, such as 'the hydrogen atoms on the left side of the picture' or 'the green block next to the red one'.

The participants. Here we distinguish between the participants' mental states and their actual behaviour. Firstly, a participant must have a knowledge of the application domain, i.e. the objects, their properties and relations. This knowledge need not be consistent and complete (and in fact, in practice it hardly ever is). Therefore it would probably be more appropriate to speak about the participants' beliefs (Bunt, 1989). Beliefs about the domain may be distributed among the participants, but in specific situations, one of them may be the expert on the domain knowledge. Secondly, a participant must have a knowledge of his or her partner's knowledge. Although the level of knowledge nestings may be arbitrarily deep (A knows that B knows that ...), in dialogue it is often only the beliefs about the participants mutual beliefs that play a crucial role (Clark and Marshall, 1981). Thirdly, in order to force agents to do an action, they must have one or more goals or intentions (Cohen and Levesque, 1990). These goals may be explicitly be represented in the agent's mental state, such as the goal to have block A on block B, but may also be driven by the situation of the dialogue (Suchman, 1987). The goal and the task of the participants are often related and determine, for instance, the relevance of specific objects at a particular point in the dialogue.

In order to achieve these goals, people may use different strategies and as a result behave differently in 'equal' circumstances. We will only consider cases in which the participants behave cooperative and follow the Gricean rules of conversation (Grice, 1975). Mutual knowledge, for instance, that an agent is an expert on the application domain has important consequences for the interpretation of 'statements', since all declarative information from the non-expert would be a violation of the quantity maxim (Beun, 1990). In less cooperative situations, agents may, for instance, behave extremely politely to avoid hurting the other person.

Interaction. Here we distinguish between the symbolic interaction between the participants and the 'physical' interaction with the application domain (respectively, observation and manipulation of the domain). Information between the participants may have a different form, content and function. The function refers to the pragmatic aspects of the message, such as the relation with previous information and the goal of the message. The content refers to the truth-conditional aspects of the message with respect to the application domain. The form of the message is related to the surface structure, such as linguistic or pictorial form and is partly determined by the modality and the channel through which the message is transported.

Agents often only have partial access to the application domain. In a telephone or email conversation, for instance, the channel is sometimes completely closed, at least for one of the participants, and the knowledge about the application domain comes from memory or other sources. In case the participants have a shared workspace, they not only perceive the (changing) properties of the objects, but also know that the other can perceive them (see for instance (Sperber and Wilson, 1986, pages 38-46)). A shared workspace also opens the possibility to point at specific objects in order to foreground them. Sometimes objects are impossible to perceive directly but can be manipulated (e.g. apertures in an electron microscope); in other cases, objects can only be observed, but not be manipulated (e.g. the picture of the specimen).

3. The Role of Demonstratives in Referential Acts

Suppose two people, S and H, are engaged in a conversation and S wants to tell H something about some object present in the application domain and/or the mental states of the dialogue participants. For instance, S wants to tell H what properties it has or what H has to do with it. In that case, S will have to make clear to H what object she is talking about. The act of doing so, is called a referential act. A referential act consists of the utterance of a referring expression and/or a pointing act. In this paper, we will only deal with reference to single objects, so only singular expressions are considered.

The referential acts we will examine can be broken down using the schema below. The brackets in the schema indicate that the category is optional; however, at least one of the optional categories must be present in each rule. If a category can be used more than once, it is marked with a star (*). The arrow ($\nearrow$) stands for a pointing act.

referential act = (utterance of a referring expression) ($\nearrow$)
referring expression = (determiner) (descriptive content)

descriptive content = (premodifier)* (head) (postmodifier)*

The following examples of referential acts are taken from real dialogues (Cremers, 1992):

1. $(\text{dat})_{det}$ $((\text{gele})_{premod}$ $(\text{blok})_{head}$ $(\text{daar bovenop})_{postmod})_{descr.cont}$
$(\textit{that})_{det}$ $((\textit{yellow})_{premod}$ $(\textit{block})_{head}$ $(\textit{there on top})_{postmod})_{descr.cont}$

2. $(\text{de})_{det}$ $((\text{groene})_{premod}$ $(\text{glijbaan})_{head})_{descr.cont}$
$(\textit{the})_{det}$ $((\textit{green})_{premod}$ $(\textit{slide})_{head})_{descr.cont}$

3. $(\text{deze})_{det}$ $((\text{blauwe})_{premod})_{descr.cont}$ $(\nearrow)$
$(\textit{this})_{det}$ $((\textit{blue})_{premod}$ $(\textit{one})_{head})_{descr.cont}$ $(\nearrow)$

The components of a referential act (premodifiers, head etc.) can be seen as providing the addressee with a set of clues for identifying the intended referent.

We will look at one specific component from this identificational perspective, i.e. the, deictically-used, Dutch demonstrative. A demonstrative expression (i.e. a referring expression with a demonstrative as its determiner) is used deictically if it is used in a conversation to refer to an object which is present in the application domain and can be perceived. If the object can be perceived visually, the demonstrative expression can be combined with a pointing act.

Besides the deictic use of demonstrative expressions, there is what has been called the anaphoric use. An anaphorically-used demonstrative expression refers back to a linguistically-introduced object. Traditionally, the expression with which the object is introduced, is called the antecedent. An object which has been introduced with an antecedent is said to be part of the linguistic context.

We will examine the identificational function of the Dutch demonstrative by comparing the deictic use of the two main existing types of demonstratives: proximates and distals:[1]

Proximate:	*dit* (this):	neuter gender/singular
	deze (this/these):	rest
Distal:	*dat* (that):	neuter gender/singular
	die (that/those):	rest

Although we focus on the function of the demonstrative, we think it unwise to study the demonstrative in complete isolation from the other components of referring expressions. Such an approach is ill-advised for two reasons.

Firstly, the use of a particular component in a referential act may depend on the other components that are part of the referential act. The choice of a demonstrative may influence the presence of a pointing act and the presence of a pointing act may in turn be related to the amount of descriptive content that is used. Such dependencies, if they exist, arise from the function of each of the separate components. Therefore, we can learn something about the function of each component, and in particular the demonstrative determiner, from these dependencies.

Secondly, a speaker has more determiners available than only the proximate and the distal demonstrative. A speaker, when he constructs a referential act, has to chose a determiner from the whole range of determiners (demonstratives, definite and indefinite article). We will try to identify some of the factors that influence this choice.

1 'Rest' stands for 'common gender/singular and common or neuter gender/plural', where 'common gender' means having one form for male and female.

To start our investigation, we need a framework in which we can formulate a hypothesis about the deictic use of the Dutch demonstrative. Here, we are faced with the problem that, to the best of our knowledge, all quantitative and computationally-oriented analyses of the use of the demonstrative, are concerned with its anaphoric use (e.g. (Kirsner, 1979); (Sidner, 1979); (Ariel, 1988); (Kirsner and Heuven, 1988); (Gundel, Hedberg, and Zacharski, 1993)).

In the remainder of this section, we will first describe how the concepts that have been employed to analyse the anaphoric use of demonstratives can be generalised so that they can be applied to the deictic use of demonstratives. Secondly, we will summarize the findings of the literature on the anaphoric use. Thirdly we will, on the basis of the aforementioned concepts and findings, construct a hypothesis on the deictic use of the Dutch demonstrative.

3.1 Force as the central concept

We will borrow the concept of *force* from (Kirsner, 1979) to give some content to the identificational perspective on the use of demonstratives. Kirsner claims that the distal and the proximate demonstrative can be used to signal different *forces with which the addressee is instructed to find the referent.*[2] He advanced the hypothesis that, in Dutch, proximates signal stronger force (i.e. greater urging that the addressee find the referent) than distals. Kirsner argued that there are basically three strategies which people use to decide on signalling either high or low force; they are: givenness, noteworthiness and foregrounding. All other concepts which have been proposed in the literature can be seen as variations on these three concepts. We will discuss each of them in the next three sections, and present proposals on how to generalise them for application to the deictic use of demonstratives.

3.1.1 Givenness. According to the strategy of givenness, *speakers will direct the addressee's attention strongest to entities that are not given 'in the addressee's consciousness'* (for instance, objects which were introduced rather far back in the discourse). Recently, more sophisticated versions of the notion of givenness have been proposed. The description above suggests that an entity either is or is not given in the addressee's consciousness. In Ariel's Accessibility Theory (Ariel, 1988; Ariel, 1990) and the Theory of Cognitive Statuses of (Gundel, Hedberg, and Zacharski, 1993), the idea of a continuous (Ariel) or discrete (Gundel et al.) hierarchy of givenness levels has been proposed.

In order to check whether people actually apply the givenness strategy, we need criteria to determine the level of givenness of an entity. In the literature, four factors have been proposed which are thought to influence the givenness of referents (see Ariel, 1988; 1990):

1. **distance** The distance between the antecedent and the anaphor. The greater the distance, the less given the object to which the antecedent refers is. Both in texts and conversations, distance can be measured in terms of the number of intervening words or sentences.

2. **unity** Antecedents which are within the same frame/world/point of view/segment or paragraph as the anaphor are more given than those that are not.

2 (García, 1975) uses the word *Deixis* for *the force with which the addressee is instructed to find the referent.* We have chosen not to do so, because it might cause confusion with the more traditional use of the term *deixis* (which we employ when speaking about *the deictic use of demonstratives as opposed to their anaphoric use).*

3. **salience** Some referents stand out more than others and are therefore more given. For example, salience is caused by the topichood of a referent.

4. **competition** If some objects have the same level of salience, this diminishes the givenness of the objects.

The first factor, i.e. *distance*, is very useful for the analysis of anaphoric demonstratives; (Kirsner and Heuven, 1988) use it as their prime criterion and it features prominently in (Ariel, 1988) in support of her accessibility theory. It does, however, also play a role in the deictic use. Even if an object has not been referred to previously in the conversation (e.g. it is part of the application domain instead of the linguistic context), it may be associated with an object which does have a high level of givenness due to distance (i.e. recent reference to the object).

Before we examine how association works in an application domain, let us look at its well-known counterpart in linguistic contexts: bridging, as in '*The book* is not that good, despite the reputation of *the author*'. In the example, the definite description '*the author*' can be used in virtue of the association people have between books and authors. (Grosz, 1977) called this phenomenon *implicit focus*.

Association, as a source of *unity*, does not only occur between objects which are part of the linguistic context. In a situation where two dialogue participants can visually perceive some part of the application domain, the spatial properties of it may actuate associations between objects. Objects may, for instance, be grouped into those which are close to the participants and those that are far away. In fact, proximate demonstratives are usually thought to refer to nearby objects and distals to objects that are far away. This kind of use will, however, not be examined in this study.

We will concentrate on situations where all objects have roughly the same distance to both dialogue participants. In situations, the distance between the participants and an object is irrelevant. What counts is the distance between the objects themselves. If one object is at the center of the attention of the dialogue participants, then part of the attention will rub off onto the objects which are adjacent to it. Following (Sidner and Grosz, 1986) we will describe the object at the center of attention as the object in *focus*. Focus is the highest possible level of givenness that an object can acquire. We propose to describe the objects that are adjacent to it as in *implicit spatial focus*. The focus and the objects in spatial implicit focus together form the *focus area* (see (Cremers and Beun, 1995)).

The factors *distance* and *salience* determine whether an object is in focus. All objects referred to in the previous sentences are candidates for being in focus. If only one object was referred to, then that object automatically will be in focus. However, if more than one object was referred to, the most salient one will be in focus. An object may, for instance, be spatially salient because it has been manipulated as a part of the task.

Salience can also arise through task-independent events, e.g. an object which unexpectly explodes will draw attention even if the object is not relevant to the task at hand.

Another task-independent source of spatial *salience* is the visually perceptible properties of objects. Objects which have an exceptional property, are more salient than those that do not have such a property. Properties are exceptional with respect to the properties that the other objects in the domain have. For instance, in a domain with twenty red blocks and one white block, the white block is more salient than the red blocks. This is not solely due to the fact that the block is white, but also to the fact that the other blocks are red.

According to *competition*, if a number of objects have the same level of saliency, this decreases their givenness. We think that competition should not be thought of as a genuine factor influencing givenness.

Consider the following situation. In an application domain containing three objects, two of the objects have the same level of saliency, and the third object has a slightly lower level of saliency. In that case, according to competition, it is possible that the least salient object is most given. We think that this runs counter to the intuition that givenness is the degree of attention in the addressee's consciousness to a particular object. We think it implausible that the addressee will address more attention to a less salient object, because the more salient objects are equal in saliency.

We will assume that competition has no influence on givenness. If we apply this to the situation which we described that means that the two most salient objects are also most given. It also means that the two objects cannot be distinguished on the basis of their givenness. Therefore, the objects are more difficult to refer to. Note that, here we speak of distinguishability of the objects with respect to a referential act. In the addressee's consciousness the objects are, of course, distinct. Confusion of distinguishability with respect to a referential act and distinguishability in the addressees consciousness, may indeed lead one into taking competition as a factor influencing givenness. Suppose equal level of salience of two objects does mean that the objects are less distinguishable in the addressee consciousness. Then it makes sense to say that the objects are also less given.

We have shown that the factors that influence the givenness of objects in a linguistic context can be generalized and applied to visually perceptible objects of an application domain. The factor of distance did not have to be translated in order to be applicable. Salience could be interpreted in terms of the perceptible actions and events in which objects can be involved and the visually perceptible properties that objects can have. Salience and distance were said to be the factors that determine whether an object is in focus (i.e. has the highest level of givenness). Unity was construed in terms of the spatial relation of adjacency between objects. Objects were said to be in implicit spatial focus, when adjacent to the object in focus. The focus area was defined as the object in focus and the objects adjacent to it.

3.1.2 Noteworthiness. The strategy of noteworthiness means that *a speaker will direct the attention of the addressee strongest to entities that the speaker is most interested in talking about.*

The notion of noteworthiness, or importance, is rather problematic. (Kirsner, 1979), for instance, suggests that the fact that an NP denotes a human, a single object or a named object indicates that the referent is noteworthy. These claims may have been appropriate for the kind of texts which he examined, a book on analytical philosophy and a novel, but are certainly not universally applicable. For example, a bricklayer who tells his assistant which bricks he needs (e.g. 'you have to hand me one long and three short bricks'), is at that moment more interested in the bricks than in the person who hands him the bricks.

Noteworthiness seems to depend on the task at hand. In a novel, which is an attempt to create a psychologically plausible picture of the main characters, these characters are obviously important. A bricklayer's task requires a completely different filling in of the notion of noteworthiness.

Noteworthiness is a sufficiently general concept to be used for the analysis of the deictic use of demonstratives. It does not depend directly on the domain but rather on the task which has to be carried out in the domain (or the shared goals of the participants if there is no explicitly stated task). The task determines which objects in the domain are noteworthy and which are not.

3.1.3 Foregrounding. Foregrounding is the third strategy which speakers can use; *if a speaker wants to urge the addressee strongly to find the referent (because it is noteworthy or not given) then he will, if possible, use more than one means to do so.*

We will try to show that the type of demonstrative (distal or proximate) is one means for signalling strong or weak force. A different means to signal strong force is to put the referent to the foreground. (Kirsner, 1979) suggests that in a text a referent can be foregrounded by putting its referring expression (a) in a main clause instead of a subordinate clause; (b) at or near the beginning of a sentence; (c) in the first sentence of a paragraph.

When objects are present in the domain and can be seen by both participants, there is yet another means to foreground an object: pointing at it.

3.2 Results on anaphorically-used demonstratives

In (Kirsner, 1979) and (Kirsner and Heuven, 1988), the use of the Dutch demonstratives was subjected to a quantitative analysis. In both studies, the hypothesis was that the proximate signals stronger force than the distal demonstrative. This hypothesis was confirmed by the quantitative analyses that were carried out. In both studies, the confirmation of the hypothesis was based on findings to the effect that proximates as opposed to distals are preferred to refer to not given and/or important (i.e. noteworthy) objects.

In (Sidner, 1979), a similar claim on the role of importance has been made for the use of the English demonstrative. However, exactly the opposite claim has been put forward with respect to givenness: English distals are preferred to refer to less given objects than English proximates (Ariel, 1990). Ariel does, however, not present any quantitative evidence for this claims. Such evidence can be found in (Gundel, Hedberg, and Zacharski, 1993). Gundel et al. found correlations between the different forms of demonstratives and different levels of givenness to the effect that distals are associated with a lower level of givenness than proximates.

Gundel et al. examined various kinds of data ranging from transcripts of casual conversations to technical specification documents. Levels of givenness were determined by letting at least two people independently determine the level of givenness of referents. The coders were given general guidelines, like if an object *can be assumed to be part of general cultural knowledge or shared personal experience of speaker and addressee, or if it has been mentioned previously in the discourse in question, the status is at least familiar*[3]. Our approach differs in at least two respects from that of Gundel et al.: we will examine a very restricted dialogue situation and formulate a number of objective criteria to determine the givenness of a referent.

3.3 The hypothesis

We will base our hypothesis on the findings of (Kirsner, 1979; Kirsner and Heuven, 1988) because these studies were concerned with the Dutch demonstrative determiner. We will, however, return to the findings on English and try to show that our results shed some new light on them.

The hypothesis. The deictically used Dutch proximate demonstrative determiner signals stronger force than the deictically used Dutch distal demonstrative determiner:

1. (Givenness) Proximates are preferred over distals to refer to objects which have a low level of givenness.

3 Gundel, personal communication.

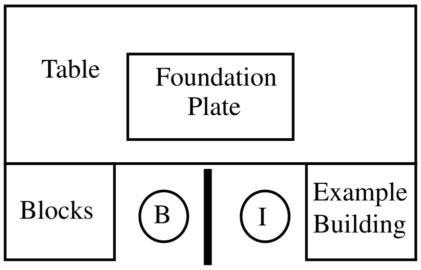

Figure 2
The experimental set-up

2.(Noteworthiness) Proximates are preferred over distals to refer to important objects.

3.(Foregrounding) Proximates are combined more frequently than distals with devices to foreground an object.

We will operationalize this hypothesis in section 4, based on our interpretation of givenness, importance and foregrounding for the application domain and the precise settings of the parameters.

Finally, a cautionary remark, we are only examining one facet of the use of demonstratives. In (Kirsner, 1993), a list of other types of uses of demonstratives is presented, including for example General Local Use (Proximates indicate proximity to the speaker, distals distance to the speaker), Temporal Use (Proximates signal current time, distals remote time; *this summer* versus *that summer*) and Salutatory Use (The distal can be used in greetings like *Ha die Bob!; Here's good ol' Bob again!*).

4. The Dialogue Situation

4.1 Set-up of the dialogue situation

The data from which we will draw our conclusions consist of a set of ten dialogues. These dialogues were collected in a specific version of the dialogue situation which was described in section 2[4]. In Fig. 4.1 the setup of the experiment is depicted schematically. Ten pairs of Dutch subjects participated in the experiment. Half of the subjects was male, the other half female, and their age varied from 20 to 60 years.

The application domain consisted of a number of blocks[5] that could be removed, replaced, and reoriented. The blocks represented some sort of abstract building, placed on a green foundation plate (size 38x38 cm) which functioned as a shared workspace. The blocks had four different colours (red, yellow, green and blue), three sizes (small, medium and large) and four shapes (square, bar, convex and concave). We have chosen these type of objects, because we wanted a simple and non-figurative application domain in order to avoid extensive reasoning on domain specific knowledge by the participants.

4 The transcriptions of the dialogues can be found in (Cremers, 1992). These transcriptions were made on the basis of videotape recordings, because the dialogues involved gestures.
5 In fact, the blocks were samples of the DUPLO-series of LEGO.

The task of the participants was to rearrange the blocks into an example building. One of the participants (the instructor I) had to instruct the other (the builder B); only the instructor could see the example building. Hence, the replica represented the goal state of the application domain which was only known to the instructor.

Both the builder and the instructor could point at and observe all objects present on the foundation plate, but only the builder was allowed to manipulate the objects with his or her hands. Builder and instructor were separated by a screen to avoid communication other than by spoken language and gesturing.

4.2 Force in the dialogue situation

The settings of the parameters, which we gave in the section above, now allows us to formulate inspectable criteria for the three parts of the hypothesis:

Givenness. We judged an object +Given if it fulfilled at least one of the three criteria below. An object that met none of the three criteria was classified –Given. We called an object *in focus* if it was referred to in the most recent utterance, and distinguishable from possible other objects referred to in that utterance, because it was the only object that was manipulated.

1. *Spatial Unity* The object is adjacent to the object which is in focus or the object itself is in focus. In other words, the object lies within the *focus area.*

2. *Spatial Salience* The object in question has an exceptional shape. The convex and concave blocks fulfilled this criterion. They clearly stood out among the rectangular blocks, which were in the majority.

3. *Focussing Expressions* The object lies in an area to which the speaker explicitly directed the attention of the addressee. This is marked by what we will call focussing expressions as in '*Wat nou helemaal naar voren zit*, daar zit die rode dwars' (literally: '*What now entirely to the front is*, there is that red one diagonally').

We expect that distals are preferred to refer to +Given objects and proximates are preferred to refer to –Given objects.

Noteworthiness. We start from the assumption that the task at hand determines which objects are most noteworthy. In the task which we have described, the central activity is the manipulation of blocks. The instructor communicates to the builder which blocks have to be manipulated (e.g. *Remove the green cube*). Therefore, we will assume that if the instructor tells the builder to manipulate a particular block, this block is more noteworthy than the other blocks.

We can determine whether a block is noteworthy by examining whether the utterance in which the referential act to the block occurs is an instruction to manipulate it. Note that objects are not only referred to because they have to be manipulated. Blocks can also be referred to in order to specify other parameters of manipulations. Consider *Put the red block on top of the yellow block.* To provide a precise specification of the manipulation of the red block, the yellow block is referred to. In this case, the red block will be considered more noteworthy than the yellow block.

Foregrounding. Pointing acts are a means to foreground objects.

5. The Results

We looked at those referential acts that were used for *initial* identification of objects in the application domain. In other words, we looked at deictically used referential acts, since objects that are referred to for the first time cannot yet be part of the linguistic context.

The ten dialogues obtained in the set-up of the dialogue situation contained 143 instances of initial referential acts. We excluded 13 referential acts which consisted of a plural and 4 instances of miscommunication and repair. Of the remaining 126 referential acts, 102 (i.e. 81%) contained a demonstrative determiner. The distribution of these referential acts over +Given and −Given entities can be found in Table 1.

In table 1, N' ranges over the following constructions: an adjective (as in *die rode; that red one*, 65%), two adjectives (as in *die vierkante rode; that red square one*, 7%), an adjective and a noun (as in *die rode balk; that red bar*, 14%), two adjectives and a noun (1%). The percentages do not add up to a 100%. The remainder can be found in the table under *no determiner* (2%) and *distal + pointing/proximate + pointing* (11%, as in *dit (↗); this (↗)* and *die (↗); that (↗)*).

In the data, there was one occurrence of what we will describe as a pre–posed post-modifier: *daar achter die gele (↗); back there that yellow one (↗).*

Table 1
The distribution of referential acts over +Given and −Given entities.

referential act	+Given	−Given
distal N′ + pre–posed postmodifier + pointing	1	0
distal N′ + postmodifier	17	3
distal N′ + pointing	18	5
distal N′	24	3
distal + pointing	2	0
proximate N′ + pointing	8	11
proximate + pointing	5	5
definite article N′	8	1
indefinite article N′	5	5
no determiner	4	1
total	92	34
percentage	73%	27%

5.1 Testing the hypothesis

Of a total of 126 referential acts, 73% were used to refer to +Given entities. The remaining 27% were used to refer to −Given entities. The null hypothesis that we will use in the paragraph below on *Givenness* is based on the aforementioned distribution. In section 6.2.1, independent evidence for this (null) hypothesis will be considered.

Givenness. The distribution of proximate demonstratives over +Given and −Given entities is 13 : 16 ($\chi^2_{df=1}$ = 11.67, P < 0.005). Given the aforemenioned null hypothesis, this result is highly significant and confirms our hypothesis: proximates prefer −Given entities.

The distribution of distal demonstratives over +Given and −Given entities is 62 : 11 ($\chi^2_{df=1} = 5,27$, P < 0.025). These data also confirm our hypothesis: distals are used more frequently to refer to +Given entities than to −Given entities.

Noteworthiness. We will examine noteworthiness in the light of the findings on givenness. The results can be found in table 2. An object was described as +Noteworthy if it was referred to in an utterance which described a manipulation of the object; otherwise the object was described as –Noteworthy.

The table tells us that distals are preferred to refer to +Given entities, when compared to proximates: distal and +G = 62 vs. proximate and +G = 13. We do, however, see that proximates, if they are used to refer to a +Given entity, are preferred to refer to +Noteworthy entities: proximate and +G, **+N** = 10 vs. proximate and +G, **–N** = 3.

On the other hand, proximates are preferred to refer to –Given entities: distal and –G = 11 vs. proximate and –G = 16. Furthermore, we see that proximates, if they are used to refer to a –Given entity, are preferred to refer to +Noteworthy entities: proximate and –G, **+N** = 13 vs. proximate and +G, **–N** = 3.

Table 2
The distribution of distals and proximates over +/– G(iven) and +/– N(oteworthy) entities.

demonstrative	+G, +N	+G, –N	–G, +N	–G, –N
distal	33	29	7	4
proximate	10	3	13	3

Foregrounding. Proximate demonstratives are always used in combination with a pointing act. On the other hand, distal demonstratives are preferred in the absence of a pointing act when they are used to refer to +Given entities. The ratio was 21 : 41 (*with pointing : without pointing*, with $\chi^2_{df=1} = 6.46$, $P < 0.025$). Distal demonstratives which are used to refer to –Given entities occur almost as many times with as without a pointing act (5 : 6). This result is statistically not significant.

5.2 Results on postmodifiers

We found that only 6 of the 10 subjects which played the role of instructor in the experiment used postmodifiers (the subjects that fulfilled the role of builder never used modifiers). The distribution of the use of these expressions over the six individuals was as follows: 7, 5, 4, 2, 1, 1. The three most frequent users of postmodifiers never pointed at an object.

Categories of postmodifiers. The postmodifiers that occurred in the data could be divided into three categories.

Firstly, of the total of 21 postmodifiers that occurred, 11 (52%) were used to relate the intended referent to the object in focus. Typical examples of this type of use were: ... *daar achter (... behind there); ... daar onder (... below there); ... wat daar voor staat (... which is in front of it).*[6]

Secondly, 8 (38%) of the 21 postmodifiers were used to indicate the location of the object: ... *op de bodem; ... op de grond (both can be translated as 'on the foundation'); die u nou hebt, pakt (the one you now have, are holding).*

Finally, the remaining 2 (10%) postmodifiers were used to describe the shape of the object: *Dat uh gedeelte met dat uh met die groene ronding (That part um with that um green curve); die blauwe balk van twee bij zes (that blue bar of two by six).*

6 Words like *there* and *it*, see the examples, were not combined with a pointing act. This is consistent with the fact that the object to which they refer is highly given, i.e. in focus.

Identificational purposes of postmodifiers. If a postmodifiers was used in a referential act, the postmodifier always served an identificational purpose. In other words, if the postmodifier had been left out of the referential act, then the referential act would have been ambiguous.

For example, consider the following situation: the focus is a green block and two red blocks are part of the focus area: one in front of the green block and one behind it. Then the postmodifier *behind the green block* clearly has an identificational purpose in the referring expression *that red block behind the green block.*

Postmodifiers and proximates. Postmodifiers were never used in combination with a proximate demonstrative determiner.

5.3 Results on the definite and the indefinite article

Indefinites were only used in sentences which expressed a state. Typical of these sentences were phrases like *Nou ligt er nog een rode hè, op de grond (Now there still lies a red one hu, on the ground); Dan moet uh een groene vierkante, die u nou ziet (Then must uh a green square one, which you now see); Nou zit daar waar u net zat een vierkante rode op de grond, hè (Now there is where you just were a square red one on the ground, hu).*

Definites were used mainly to refer to +Given entities (8 occurrences for +Given and 1 for −Given).

In our data, neither indefinites nor definites occurred in combination with a pointing act.

Finally, there were five referring expressions which lacked a determiner, as in *Even kijken hoor, nou hebben we geel (Let's have a look, now we have yellow).*

6. Discussion

6.1 The hypothesis

All three subhypotheses were confirmed by our findings. This shows that the deictic use of the Dutch demonstrative can be described along the same lines (with the concept of *force*) as the anaphoric use (which was examined in (Kirsner, 1979) and (Kirsner and Heuven, 1988)). It is one more piece of evidence for the claim that the use of referential acts, both in texts and conversations, is governed by general principles of human cognition and perception. In this paper, these principles were formulated as the concept of force and the strategies of givenness, foregrounding and noteworthiness.

The findings on Dutch do, however, not agree with those on English. (Gundel, Hedberg, and Zacharski, 1993) found that in English the proximate signals higher givenness than the distal. For Dutch, we found exactly the reverse. There is, however, a mapping possible between the result from Gundel et al. and ours. Gundel et al. found that whereas proximates were used exclusively to refer to activated objects, distals were used to refer to both activated and (the less given) familiar objects (Gundel et al. define familiar as 'The addressee is able to uniquely identify the intended referent because he already has a representation of it in memory ...' and activated as 'The referent is represented in current short term memory')

We found that proximates are preferred to refer to −Given entities and that they always occur in combination with a pointing act . Distals, on the other hand, are preferred to refer to +Given entities and they do not have to be combined with a pointing act. The mapping from our data to those of Gundel et al. is now fairly straightforward: assume that (1) objects which have been pointed at are activated and (2) +Given objects are familiar.

The possibility of a mapping does not mean that there are no differences. For English, a hierarchy of cognitive statuses was proposed in which activated is a higher status than familiar. Next, proximates were associated with activated and distals with familiar (see Gundel et al., but also (Ariel, 1988; Ariel, 1990)). We think that for the deictic use of the Dutch demonstrative a one-dimensional hierarchy of givenness levels is not appropriate. What we observed can be described better in a procedural manner: if an object is –Given, then a speaker uses extra means to help the addressee to identify the object (in other words, signal extra force to find the referent). He can do this by using a proximate and pointing.

It would be interesting to see whether this pattern is also present in the data on English. This requires making a precise distinction between the level of givenness of an object *immediately before* a referential act, and the level *due to the referential act itself.*

6.2 Outline of a speaker model

In this section, we will outline the beginnings of a computational model for the generation of initial referential acts to objects in the application domain. We will not look at a model for interpretation of such acts, because that would require a method for telling whether a referential act is initial or not. There is, however, work being done in this direction within the DenK project ((Kievit and Piwek, 1995); (Beun and Kievit, 1995)).

We will tailor our model according to a model of the speaker which has been proposed in (Levelt, 1982). In this model, Levelt (following (Kempen, 1977)) draws a distinction between 'the processes involved in the genesis of the *ideas* underlying speech, and the processes involved in the choice of the linguistic *forms* for their expressions' (Levelt, 1982, page 201). This distinction can be applied to the generation of referential acts by looking at them as consisting of two stages: Firstly, the speaker chooses the object to refer to (the *target object*) from other objects present; and secondly, the speaker formulates the appropriate referring expression that will direct the addressee to the target object.

6.2.1 The choice of an object. In (Levelt, 1982) reference to objects in a spatial grid-like network was examined. According to Levelt, in choosing the target object people adhere to the *principle of connectivity.* They chose, wherever possible, as the next node to be described one that had a direct connection to the node that was most recently referred to. In the vocabulary which we have used, this means that people prefer to refer to +Given objects and, in particular, objects that are in spatial implicit focus.

The choice for an object in spatial implicit focus may hinge on a general problem-solving strategy of humans. When people are trying to solve a complicated problem, they tend to break this problem down into parts (Thomas, 1974). In an application domain, a part of the problem may be directly mapped to a spatial area of mutually adjacent objects.

In our data, we also found that people prefer to refer to +Given objects. Of a total of 126 referential acts, 73% were used to refer to +Given entities. The remaining 27% were used to refer to –Given entities.

6.2.2 The formulation of referential acts with minimal effort. The main principle that guides the speaker in formulating the referring expression has been called the *principle of minimal effort* ((Zipf, 1949); (Brown, 1958); (Krauss and Glucksberg, 1977); (Olson, 1970)). It states that speakers try to utter a noun phrase that is as short as possible, while still allowing the addressee to select the referent. Clark and Wilkes-Gibbs extended this principle by saying that referring to objects is a collaborative process (Clark and Wilkes-Gibbs, 1986). Their *principle of minimal cooperative effort* expresses the idea that there is a trade-off between the noun phrase that is uttered first and the possible additions or corrections to this utterance by the speaker or the addressee.

(Cremers and Beun, 1995) argued that in the process of referring to an object there is actually more effort involved than only the effort to utter the noun phrase. In their *principle of minimal cooperative total effort*, they state that the effort to actually identify the target object by the addressee should be taken into account as well.

In formulating the expression and/or gesture to refer to the chosen object, people can make use of the closeness of the object to the one that is in focus. The definite expression can usually be reduced in length, because the target object only has to be distinguished from other objects which are in spatial implicit focus. For example, if there is only one red object located within the current focus area, referring to it by means of 'the red one' suffices, even if there are more red objects located outside of the focus area.

6.2.3 Application of the model. We have already pointed out that the generation of a referential act involves two stages. The speaker first has to determine which object he wants to refer to. In this phase, the principle of minimal effort predicts that the speaker will, if possible, select an object that is +Given. As soon as an object has been chosen, the speaker needs to formulate an appropriate referring expression (possibly accompanied by a pointing act). There are three choices/decisions which the speaker has to make:

1. The choice of the determiner

2. The decision whether to use a pointing act

3. The choice of the descriptive content (head, pre- and postmodifiers)

We assume that people construct a referential act incrementally in a left to right fashion. In this process, descriptive content is added until disambiguation has been achieved. Disambiguation does not have to be achieved with respect to the entire application domain; disambiguation with respect to the focus area is sufficient. In other words, speakers try to minimize their effort.

We will concencrate on those parts of the referential act which we found to be related to the type of determiner: the determiner itself, pointing and postmodifiers.

Determiners. The speaker first determines which determiner has to be used. The choice of the determiner depends largely on whether the speaker wants to signal weak or strong force. The main factor is whether the object is + or –Given. The distal demonstrative is associated with +Given objects, whereas the proximate demonstrative is connected with –Given entities. No hard guidelines for the use of the definite and indefinite articles are given, since their percentage in the data was too low (18%). We will, however, state some observations on their use.

We found that one can refer to a specific object with an indefinite if it is used in sentences which describe a situation, e.g. *and there you see a red block.* This way, the utterance in which the object is referred to can be detached from the description of the action that has to be performed on the object, as in *put it on top of this blue one.*

We think that one reason why a speaker may want to separate the referential act from the description of the action that has to be performed, is simply to buy time. Sometimes, a speaker has already made up his mind about what object has to be manipulated next, but has not yet fully decided on the precise (description of the) action that has to be performed on it. The following subdialogue illustrates this:

A: ... Nou ligt er nog een grond een rode, op de grond.
(Now there still is a floor a red one, on the floor.)
B: Ja. (Yes)

A: ... Nou dat moet, (Well that has to)
even kijken hoor, (let 's see)
dat moet een blauwe worden (that has to become a blue one).

Notice that the fact that the utterance involving the referential act and the action description has been split into two utterances, leaves room for the addressee to give feedback on whether he understood the referential act, before the speaker goes on to describe the action. This enables the speaker to check in advance that the object with which that action will be performed is the right one.

As far as the definite article is concerned, our data only tell us that it is preferred to refer to +Given entities (8 out of 9).[7]

Perhaps, the difference between the definite article and demonstrative determiners cannot be accounted for in terms of the identificational clues they provide. In (Maes and Noordman, 1995), a quantitative analysis of the anaphoric use of demonstrative determiners and the definite article is presented[8]. Their conclusion is that the demonstrative determiner, as opposed to the definite article, signals a predicating reading of the NP. The following example illustrates this:

> Bush$_i$ has decided to raise taxes after all. It is the first time {the / this} president$_i$ has dared to do so. (Maes and Noordman, 1995, page 262)

Maes and Noordman claim that, in this example, the definite article invokes the identify relationship between antecedent and anaphor, whereas the demonstrative determiner invokes the classify relationship. The classify relationship indicates that the referent of the anaphor is a member of the set of objects denoted by the descriptive content of the antecedent. The classify relationship instigates contextually-relevant inferences in which the referent of the anaphor is contrasted with other class members of the descriptive content of the antecedent. For the example discourse, Maes and Noordman suggest that the speaker may intend that the addressee infers that *the speaker is not talking about other presidents* and *other presidents might have increased taxes.*

Pointing. The decision on whether to point or not cannot be considered independent of the demonstrative that is used. If a proximate is used, then pointing is obligatory, whereas if a distal is used, then pointing is optional. Our data do not reveal whether a speaker first determines which demonstrative to use and then decides on the addition of a pointing act, or that he proceeds in the reverse order.

Postmodifiers. We found that proximate demonstratives were never combined with postmodifiers. We will argue that this can be explained by the fact that proximates always cooccurred with a pointing act. We will try to show that postmodifiers and pointing acts are incompatible.

Pointing is a very strong referential device. If it is used correctly, it enables the addressee to single out precisely the intended referent. Let us compare this to the three categories of use of postmodifiers that we discerned.

Firstly, most postmodifiers were used to relate the intended referent to the object in focus (e.g. *dat rode blok daar onder (that red block below there)*). In such cases the

7 For English, exactly the opposite claim has been made in Ariel, 1990 and (Gundel, Hedberg, and Zacharski, 1993). They argue that the definite article signals lower givenness than both the distal and the proximate demonstrative.

8 They examined Dutch texts, but claim that their analysis is also applicable to English.

postmodifier expresses a relation between the intended referent and some other object. That means that the addressee has to identify two objects. Suppose that a speaker points and uses a postmodifier. In that case, the use of a posmodifier only complicates things for the addressee: the pointing act helps him to single out exactly the intended referent, whereas the postmodifier forces him to identify a second object, different from the intended referent.

Secondly, postmodifiers were used to indicate the location of the intended referent (e.g. *dat gele blok op de bodem (that yellow block on the foundation)*). In such cases, the postmodifier does not complicate things, but is redundant, when a pointing act is used as well. Therefore, the principle of minimal effort predicts that speakers will also in this case abstain from using postmodifiers and pointing together.

Thirdly, there was a small number of postmodifiers (2) that were used to describe the shape of the object: *Dat uh gedeelte met dat uh met die groene ronding (That part um with that um green curve); die blauwe balk van twee bij zes (that blue bar of two by six)*. In these two cases, use of a pointing act would again yield a redundant referential act. The first of these two referential acts involved reference to a property of the object which makes it unique in the application domain: its convex shape. In the second instance the description of the size of the object singles out exactly one object in the application domain.

Finally, we would like to make a remark on the first type of use of postmodifier: the use of postmodifiers to relate the intended referent to the object in focus. We found that in all these cases the postmodifier was needed for disambiguation. There is a side-effect to this kind of use of postmodifiers: it increases the coherence of text by relating a newly introduced object to an object which was at the center of attention.

Speakers and the use of postmodifiers and pointing. Finally, we would like to point out that the three strategies of dialogue participants regarding the use of pointing and postmodifiers (exclusive use of modifiers, exclusive use of pointing, or use of both) which we discerned correspond to three different models for the generation of referential acts. One could, for instance, make a model which only caters for pointing acts. After all, there are people who behave this way. This option may make life a little bit easier, since the generation of a pointing act requires little computation, whereas finding the appropriate relations or properties for a postmodifier is not that simple.

6.3 Parameters and applications

6.4 Generalization of the results

In the previous subsections, we have been arguing that the results found by Kirsner on the use of demonstratives in texts can be extended to the use in dialogue situations. But, since we concentrated on a particular setting of the dialogue parameters, the main question of the introduction still has to be answered, i.e. how can the results of our observations be generalized to dialogue situations other than the one described in this paper. A complete answer for all types of (communicative) situations is outside the scope of this paper and is probably impossible to give. To answer at least part of the question, however, we will return to the basic parameters of the dialogue: the application domain, the dialogue participants and the interaction. It should be mentioned, though, that their properties are not independent. For instance, objects that do not exist in the domain are not observable, and auditory perception of static domains is impossibles, since every auditory signal requires some physical change in time.

The application domain. In the blocks domain that we discussed, the givenness criterion of an object was determined by two factors: a. by contrast of the object to other objects due to salient properties of the object, manipulation of the object or reference to the object, and b. by association of the object with another given object. In the application domain association manifested itself in terms of marked spatial areas, either implicitly indicated by the spatial distance to a focussed object, or explicitly indicated by a speech act about the relevant area.

In case of contrast, the salient properties of the objects were limited to static visual properties, such as a specific shape of the blocks. But, since the periphery of the human retina is especially sensitive for temporal change, it is to be expected that autonomous behaviour of the objects (e.g., movement and flickering in a graphical user interface) may cause the same saliency effects. Also, if temporal aspects are included, we may even expect to generalize the results to auditory perceivable domains, such as noises and music (cf. '*this* drum' vs. '*that* violin'). In the spatial domain that we investigated, however, the use of proximates was always accompanied by pointing, which is less usual in a domain that can be perceived auditory only. But, since pointing is a particular type of foregrounding, we may expect speakers to use other means to attract the hearer's attention, such as word stress and specific types of gesturing.

In case of a marked area, the relevant object is located in a specific spatial area that is probably related to the perceptual characteristics of the eye (see e.g. Treisman and Gelade, 1980). Hence, we may expect to transfer these results to other types of spatial environments, for instance, the windows system on a computer screen. We have to make a distinction between spatial and non-spatial relations, however. Objects often form a functional unit which is invisible for the participants. In some cases, these objects may be spatial 'neighbours', but in other cases they may fall outside the scope of the participants' visual focus (e.g., components on a printed circuit board). In our blocks domain the objects were rather simple and non-figurative, and functional units without an observable link between the objects were lacking. So, we have to be careful in generalizing our results to domains where objects become given as a result of relations other than spatial ones. An extreme form is the example of implicit focus given in section 3.1.1 where the introduction in the dialogue of 'the book' makes the author who wrote the book more given. In this case, the use of a demonstrative for reference to the author seems inappropriate (*? The book is not that good, despite the reputation of that / this writer*), unless the writer is already mentioned or foregrounded, for instance, by pointing.

A final remark on the domain properties concerns the relative distance of the objects compared with the distance between the objects and the dialogue participants. Here, we have been concentrating on situations where the objects have roughly the same distance to both participants. Since the use of proximate and distal demonstratives also depends on the distance of the objects to the dialogue participants, it remains for future research how our results can be transferred to situations where the distances between objects and participants is not constant and may therefore influence the use of the demonstratives.

The participants. In the dialogue situation, the builder was the person who manipulated the domain, and the instructor indicated how the domain had to be changed in order to reach the goal state. Since the force criterion discussed in section 3.2 does not depend on the distribution of the knowledge or on the person who manipulates the objects, it may be expected that our results can be generalized to very different types of dialogue situations where agents collaborate in order to perform a mutual task.

Probably, an exception has to be made for situations where the dialogue participants lack a mutual belief about specific relations between objects inside the domain. Consider, for instance, a teacher-apprentice situation where the teacher explains the function of different parts of a car (the shared work space). If the teacher believes that the student is unaware about the parts of the engine, she may not assume, for instance, that the introduction of the engine in the dialogue makes the carburetter more given. In other words, our results are only applicable to situations where the dialogue participants have mutual knowledge about the relations and properties of the application domain.

The interaction. An important property of the dialogue situation was that different types of interaction could be coordinated and carried out simultaneously. Participants could, for instance, at the same time point at an object, observe the domain and perform a referential act. Another important property was that the domain could be mutually perceived. Consequently, if changes happened in the domain, they believed that they were mutually aware about these changes.

An example of a dialogue situation where the referential acts were 'uncoupled' from pointing acts and domain observations is given in (Cremers, 1995). In her paper, Cremers describes the same dialogue situation, but the spoken interaction was replaced by a typed one. In the resulting dialogues, the participants were unable to coordinate pointing acts and domain observation with typing. As a result, she found that the participants did not use the spatial focus area in the same manner as in the spoken dialogues. An important observation was that, in contrast with spoken dialogues, in many cases the determiner was omitted from the referring expression.

Hence, we expect our results to be transferable to situations where the dialogue participants have visual access to a shared application domain, provided that foregrounding can be established simultaneously with symbolic interaction and domain observation.

7. Conclusions

In this paper, a model for dialogue situations was presented based on a general architecture for user-interface dialogue systems. In line with natural interaction of humans with their environment, the model distinguishes between symbolic and physical interaction, for instance, speech acts and direct manipulation. It was used to parameterize a dialogue situation in which the deictic use of Dutch demonstratives was investigated.

Based on an quantitative analysis, it was shown that already existing results on the anaphoric use of demonstratives could be generalized to their deictic use. For that reason, the notion of force, which was developed especially for linguistic contexts, was extended to an application in a visual accessible domain. The results showed that the same cognitive principles, such as focussing, foregrounding and importance, underly the process of reference to objects in both a linguistic context and the visual domain. Another important result was that proximate demonstratives indicate that a speaker wants to signal the addressee extra force to find the relevant referent, while on the other hand, distals are often used in cases where the referent was more 'given' in the addressee's consciousness. These observations also helped us to explain how the use of demonstratives is related to the use of pointing and postmodification in a dialogue situation.

Our findings on Dutch, however, did not correspond equally well to some of the reported findings on the anaphoric use of the English demonstrative. Studies into the English demonstrative are based only on the notions of givenness and importance, and neglect the strategy of foregrounding. We think that this strategy is, however, not limited to the Dutch. This is an interesting area for future research, i.e. examination of the use of English demonstratives on the basis of the dialogue model that we presented and the notion of force, in particular foregrounding.

An advantage of the use of a general model for dialogue situations is that the results of the observed data can be extended to other application domains and other communicative situations. An important limitation, however, is that our results are limited to cases where both dialogue participants have mutual access to a shared domain. Also, our findings are limited to situations which allow simultaneous communication, foregrounding and observation of the domain. It remains for future research how these results can be generalized to dialogue situations where people have partial access to a shared domain and where people have other means for symbolic interaction.

Acknowledgments

The authors wish to thank Bob Kirsner, Tanya Reinhart and Arie Verhagen for their helpfull comments. The research reported in this paper was financially supported by the Cooperation Unit of Brabant Universities (SOBU).

References

Ahn, R., R.J. Beun, T. Borghuis, H. Bunt, and C. van Overveld. 1995. The denk-architecture: a fundamental approach to user-interfaces. *Artificial Intelligence Review*, 8(2).

Ariel, M. 1988. Referring and accessibility. *J. Linguistics*, 24:65–87.

Ariel, M. 1990. *Accessing Noun-Phrase Antecedents*. Routledge, London and New York.

Beun, R.J. 1990. The recognition of dutch declarative questions. *Journal of Pragmatics*, 14:39–56.

Beun, R.J. and L. Kievit. 1995. Resolving definite expressions in denk. *DenK report*, 16.

Brown, R. 1958. How shall a thing be called. *Psychological Review*, 65(1):14–21.

Bunt, H. 1989. Information dialogues as communicative action in relation to partner modelling and information processing. In M.M. Taylor, F. Néel, and D. Bouwhuis, editors, *The Structure of Multimodal Dialogue*. North-Holland, Amsterdam.

Carter, D. 1987. *Interpreting anaphors in natural language texts*. Ellis Horwood Limited.

Clark, H. and C. Marshall. 1981. Definite reference and mutual knowledge. In A. Joshi, B. Webber, and I. Sag, editors, *Elements of discourse understanding*. Cambridge University Press, Cambridge.

Clark, H. and D. Wilkes-Gibbs. 1986. Referring as a colloborative process. *Cognition*, 22:1–39.

Cohen, P. and H. Levesque. 1990. Intention is choice with commitment. *Artificial Intelligence*, 42:213–261.

Cremers, A. 1992. Transcripties dialogen blokken-experiment, 1992. *IPO rapport*, 889.

Cremers, A. 1995. Object reference during task-related terminal dialogues. In H. Bunt, R.J. Beun, and T. Borghuis, editors, *Proceedings of the international conference on cooperative multimodal communication CMC-1995*. Cooperation Unit Brabant Universities, Eindhoven.

Cremers, A. and R.J. Beun. 1995. Object reference in a shared domain of conversation. *IPO manuscript 1089, Submitted for Publication in Pragmatics and Cognition*.

García, E. 1975. *The Role of Theory in Linguistic Analysis: The Spanish Pronoun System*. North Holland.

Grice, H.P. 1975. Logic and conversation. In P. Cole and J. Morgan, editors, *Speech Acts. Syntax and Semantics, Vol. 11*. Academic Press, New York.

Grosz, B. 1977. *The representation and use of focus in dialogue understanding*. SRI International, Menlo Park. Technical Note 151.

Gundel, J., N. Hedberg, and R. Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language*, 69(2):274–307.

Hutchins, E. 1989. Metaphors in interface design. In M.M. Taylor, F. Néel, and D. Bouwhuis, editors, *The Structure of Multimodal Dialogue*. North-Holland, Amsterdam.

Kempen, G. 1977. Conceptualizing and formulating in sentence production. In S. Rosenberg, editor, *Sentence Production: Developments in Research and Theory*. Erlbaum, Hillsdale.

Kievit, L. and P. Piwek. 1995. Services in sherlock. *DenK report*, 13.

Kirsner, R. 1979. Deixis in discourse: An exploratory quantitative study of the modern dutch demonstrative adjectives. In Talmy Gívon, editor, *Syntax and Semantics: Discourse and Syntax Vol. 3*. Academic Press, New York, pages 355–375.

Kirsner, R. 1993. From meaning to message in two theories: Cognitive and saussurean views of the modern dutch demonstratives. In R. Geiger and B. Rudzka-Ostyn, editors, *Conceptualizations and Mental Processing in Language*. Mouton de Gruyter, Berlin and New York, pages 80–114.

Kirsner, R. and V. Van Heuven. 1988. The significance of demonstrative position in modern dutch. *Lingua*, 76:209–248.

Krauss, R.M. and S. Glucksberg. 1977. Social and non-social speech. *Scientific American*, 236(2):100–105.

Levelt, W. 1982. Linearization in describing spatial networks. In S. Peters and E. Saarinen, editors, *Processes, beliefs and questions*. D. Reidel, Dordrecht, Holland, pages 199–220.

Maes, A. and L. Noordman. 1995. Demonstrative nominal anaphors: a case of nonidentificational markedness. *Linguistics*, 33:255–282.

Olson, D.R. 1970. Language and thought: aspects of a cognitive theory of semantics. *Psychological Review*, 77:257–273.

Sidner, C. 1979. *Towards a Computational Theory of Definite Anaphora Comprehension in English*. MIT Technical Report.

Sidner, C. and B. Grosz. 1986. Attention, intentions and the structure of discourse. *Computational Linguistics*, 12(3):175–204.

Sperber, D. and D. Wilson. 1986. *Relevance: communication and cognition*. Harvard University Press, Cambridge, Massaschusetts.

Suchman, L. 1987. *Plans and situated actions*. Cambridge University Press, Cambridge.

Thomas, J.C. 1974. An analysis of behavior in the hobbits-orcs problem. *Cognitive Psychology*, 6:257–269.

Treisman, A. and G. Gelade. 1980. A feature-integration theory of perception. *Cognitive Psychology*, 12:97–136.

Zipf, G.K. 1949. *Human behavior and the principle of least effort*. Addison-Wesley Press.