

Processing definite NPs

Peter Bosch & Bart Geurts

This paper proposes a computational model for the treatment of definite NPs in natural language discourse. The basic idea is that we do not, first, parse a sentence and then, on the basis of the parser output, start the semantic and discourse processing, including the referential resolution of any definite NPs, but rather tackle the task of NP resolution already during the parse process by letting a specialized module (the "NP-Resolver") interact with the parser. The advantage of this approach is twofold: on the one hand, we argue, this interactive form of processing is cognitively more realistic than a classic sequential approach. On the other hand, we get considerably less ambiguity in the course of parsing, because parses that would be syntactically possible, but are implausible for semantic, in particular contextual reasons, are eliminated as early as is at all possible. Related problems that are discussed in some detail in this paper include the notion of definiteness for NPs and the question of representation and referential accessibility of discourse referents.

0. Introduction.

In this paper we consider definite NPs from a *computational vantage point*. The adjective "computational" is here not meant to signify the fact that we use computers in our linguistic work, which is only a trivial practical matter. The point is that there are a couple of important implications for our way of doing linguistics.

One is that the computational approach makes us think of a natural language not only as a set of rules and representations but also as a system of procedures that make use of such rules and representations. This gives us an important advantage over linguistic theories that are interested in competence only: it does not leave a gap between the empirical observation of linguistic behaviour and hypotheses about competence and thus makes our approach in a more direct sense empirical.

Our computational approach does not imply that we blur the competence-performance distinction. On the contrary: we can be quite explicit about the distinction. On the one hand there is competence and there are, e.g., rules of syntax and semantics. On the other there are interpretations assigned to sentences in particular contexts, inferences drawn from these sentences, etc. The computational approach does not oblige us to let the former directly

produce the latter. Performance, i.e. procedures that make use of linguistic rules, is what comes in between and may explain, on the one hand, why certain interpretations that ought to be possible according to competence never actually occur to a native speaker and why, on the other hand, interpretations that, on grounds of competence rules, ought *not* to be assigned to a sentence are still the most natural ones for a native speaker, even though, perhaps, with the feeling that what is said could better be expressed in other words.

The second important implication is due to the fact that we engage in modelling, or simulating, human linguistic understanding on a computing machine. The aim is to make the machine behave like a human language processor, at least with respect to some aspects of the use of language. Our system reads natural language texts (in our case German texts) and shows that it has understood these texts by answering questions that any human could reasonably be expected to answer after having read the same texts. This approach forces us to take in principle all factors into account that influence human text understanding. Even if we believe, as some linguists do, that the human linguistic faculty is an autonomous mental organ, independent of other cognitive faculties, we still have to model also the workings of those other faculties, at least to the extent that they play a role in linguistic understanding. Thus we must gain a reasonably clear understanding of how knowledge of the language interacts in linguistic comprehension with other forms of human knowledge.

Since this is a task that goes beyond the regular and accepted tasks of linguistics, we would like to give a brief illustration. Often when we consider linguistic phenomena we would find that linguistic comprehension interacts with, or subsumes, the processing of other than merely linguistic knowledge. For instance, when we wonder why in sentence (1a) one would usually prefer an interpretation of the pronoun according to which it refers to the photo rather than the pocket, we can, as it were, choose between an explanation via world knowledge or via linguistic mechanisms.

(1) a. John took a photo out of his pocket and put it on the table.

According to world knowledge photos are more likely than pockets to be put on tables, pockets usually being firmly attached to pieces of clothing. But we may also use linguistic regularities to explain the preferred reading: that non-emphatic third person personal pronouns tend to refer to contextually more salient objects rather than to less salient ones, and that referents presented by direct objects are *ceteris paribus* more salient than referents that are mentioned in prepositional phrases.

The choice between the two explanations is not only a matter of taste. Although, as linguists, we ought to be more interested in linguistic mechanisms than in apparent trivialities of common sense knowledge about pockets and photos, we must be aware that human linguistic comprehension may use either. Cf. (1b), where the linguistic rule is not supported by a common sense expectation: photos are as likely to be put on a table as are wallets. The

preference for letting *it* refer to the photo is still there, even though perhaps to a lesser degree than in (1a):

(1) b. John took a photo out of his wallet and put it on the table.

Still, we may not conclude that common sense knowledge only comes in to facilitate particular interpretations, while linguistic regularities are eventually decisive. In (1c) we can see how the same linguistic regularity that dominates the interpretation of (1a) and (1b) loses out against common sense knowledge.

(1) c. John wiped the dust off a photo and put it on the table.

Depending on the strength of the linguistic rule or the preference dictated by common sense, and depending on the information available it may once be the linguistic regularity and another time it may be common sense knowledge that is decisive. Thus if we are serious about wanting to explain, or to model, human linguistic comprehension, there is no way but to investigate the interaction of linguistic knowledge with other knowledge sources.

A purely competence-oriented approach would give us neither any rules of preference, nor would it get involved with the role of common sense knowledge. It would treat (1a-c) all in the same way and blandly inform us that there are no constraints that would forbid either of the two readings, and thus there would be no account for the differences we have noted.

The clearest case though that can be made for the crucial importance of the interaction of linguistic knowledge and other forms of knowledge in linguistic comprehension does not come from the investigation of how humans understand single sentences – which is a somewhat artificial task anyway – but from the understanding of sentences in a discourse context. No sentence is appropriately understood unless it is understood in relation to the knowledge that has already been gained from preceding sentences in the discourse. This is demonstrated most easily by reference to reference: Unless readers of (2) have understood that the pronoun in the second sentence refers to the person that is already mentioned in the first sentence, they cannot draw the inference that Fred took the train to Gloucester this morning. And unless a reader has drawn this inference, he cannot fairly be said to have understood the text in (2).

(2) Fred went to Gloucester this morning. He took the train.

When we think about how to model the required interaction between linguistic knowledge and what may for short be called factual knowledge then a simple and, in a sense, natural picture comes to mind, according to which there is, first, a stage at which the linguistic knowledge is set to process an utterance, and then, second, a stage at which the results of the linguistic analysis are processed with regard to other cognitive faculties.

The currently most generally used architecture of natural language understanding systems corresponds to this two-stage approach, as in the early work

by Woods et al. (1972). Here, in a first stage, the syntactic analysis produces a parse tree, and only in subsequent stages is there a semantic analysis where the contents of the sentence can be related to knowledge already available from preceding text, general knowledge, or common sense. The problem with this architecture is that, on the one hand, it is – as we shall see – hardly true to the facts of human linguistic comprehension and, on the other, that it is computationally extravagant, because it forces us, in order to be sure that we won't miss the syntactic analysis that is actually relevant, to produce *all possible* syntactic analyses – only in order to throw away all but one of them after we have taken account of the semantics and related the sentence to already existing knowledge representations. In this paper we shall develop an alternative architecture that does not suffer from these drawbacks.

1. *Definiteness of NPs.*

We want to concentrate here on the understanding of full definite NPs and related ambiguities, which comes down mainly to the question of reference resolution for definite NPs.

Why should there be a problem of finding referents for definite NPs and not for indefinite NPs? Why are we happy just introducing a referent into our representation and specify that this referent is a man when we process the indefinite *a man* in (3a), but start a whole procedure of trying to figure out what man may be meant when we process *the man* in (3b)?

- (3) a. A man entered a pub.
 b. The man ordered a beer.

Well, because this is part of understanding the definite NP in (3b), and if we wouldn't, we would miss crucial inferences, as already noted in the discussion of (2). – But this answer is of course too general. We need to know in what way exactly the definiteness and indefiniteness of the NPs causes the difference in the processing of (3a) and (3b).

As for (3a), part of the understanding process consists of setting up a model in which there is a man and a pub and where the former entered the latter at a time preceding the current time. In building this model we may make use of linguistic knowledge and of background knowledge we have about pubs and men and men entering pubs. We do not need, however, any knowledge of any specific man, any specific pub, or any situation in which a specific man enters a particular pub. When it comes to (3b), however, this is exactly the kind of knowledge that we need in addition, because (3b) carries a presupposition, connected to the definite *the man*, which must be fulfilled in the specific situation with respect to which the sentence is interpreted. This presupposition is usually taken to be a presupposition of existence and uniqueness: that there is one and only one man in the relevant situation whom the sentence is about. In the interpretation of (3b) then we must look for a suitable model of such a situation. And if we cannot find such a model, or cannot even construct

one – perhaps because it would be in contradiction with other things we know – then the presupposition would fail and the sentence would be incomprehensible. If things go as they ought to, and there is already a unique referent for the definite NP in the model, then we have to extend the model so that what the sentence says about the man is added to the model as a property of that referent. If things are not as they ought to be and there is no unique man in our model, we may, if nothing contradicts this, just fix up our model so that it does contain a unique man and then proceed as before.

This latter comprehension strategy – we shall speak, following Lewis (1979), of “accommodation” – does not really work when the uniqueness presupposition is not fulfilled, because then we are stuck with a model that contains *more* than one man. We may infer that we ought to fix up our model so that there is only one man, but if there are more, we don't know which one to remove and which to keep.

But accommodation works nicely for the existence presupposition in case there is no man available yet. Still, even here it works best when our model is pretty empty, ideally at the beginning of a discourse. When in the middle of reading about, say, postmodern architecture one is faced with sentence (3b) nothing would really literally contradict introducing a referent specified as a man and letting him order a beer. The problem then is rather in finding a way of fitting this referent in with the rest of the model. – Satisfying a presupposition thus is not merely a matter of avoiding contradiction. Avoiding contradiction is only one necessary condition for establishing coherence, but by no means sufficient on its own.

Thus accommodation is nothing to be too lighthearted about: it's easy at the very beginning of a discourse when the model is empty, because coherence is not at issue. But when presuppositions want to be accommodated in an already densely populated universe, a little more is required. We shall return to this problem in a later section.

For a definite NP, coherence with the current discourse model is typically established in the form of reference to an entity in the model. In the case of an indefinite, most clearly a quantified NP, the coherence requirement typically rests on the interpretation of the corresponding VP, which must link up to the existing discourse model. Thus sentence (3a) would be as much out of place in the middle of our piece on postmodern architecture as (3b). But here the problem is not with any presuppositions of the NP. It is rather that we find it hard to link the interpretation of the VP to the current state of our discourse model (cf. Bosch 1983, Ch. 3.3).

But coherence is too big a topic for this paper. Here we are specifically concerned with definite NPs and with how they are coherently linked to an existing model. Thus we shall concentrate on cases where there are no problems with a coherent VP interpretation and shall avoid making special assumptions with regard to that. The more specific question to be discussed now has to do with the requirements the definiteness of an NP imposes on already existing representations of contextual as well as general knowledge.

A currently popular assumption is that definiteness is first and foremost a matter of anaphora, where "anaphora" must be taken in the not too narrow sense of resumption of referents that are already represented in the current discourse model, and not merely of coreference with an antecedent. Thus not only Hawkins' (1978) class of *anaphoric definites* would be covered by this notion, but also his *immediate situation use* and *wider situation use* of definite NPs as exemplified in (4) and (5) respectively.

- (4) Watch out for the dog.
(uttered near premises that are known to be guarded by a watch dog, without any dog currently being noticeable)
- (5) The Prime Minister will be visiting Rio next month.

Also here one could argue that the current model, induced by the knowledge of the respective premises or country, already, as it were "contain" the dog or prime minister. Purely anaphora based accounts of definiteness get into difficulties however with cases of definite reference that are not clearly related to specific models and accordingly don't impose any requirements (that could possibly be *not* fulfilled) on referents already established. Cf. (6).

- (6) a. *The product of three and four* is twelve.
b. *The highest mountain of Corsica* is nearly ten thousand feet.
c. *The first day in April* is a Wednesday this year.

After (4) or (5) "What dog?" or "What prime minister?" are reasonable questions for an uninitiated listener to ask. The analogous questions after any of the sentences in (6), like "What product of three and four?", are plain nonsense.

Löbner (1985) takes uses like those in (6) as the paradigm of definiteness and argues that it is indeed not anaphora that counts for definiteness, but the fact that the NP is *interpreted as a functional concept*. Accordingly, nouns, or modified nouns, that force a functional interpretation are unacceptable when preceded by an indefinite article. Cf.

- (6) a' *A product of three and four...
b' *A highest mountain of Corsica...
c' *A first day in April...

Functional concepts are a subclass of relational concepts, namely those whose referent-defining relation is a function. *Relational concepts* are those that describe their referents in relation to other objects. They stand in opposition to *sortal concepts*, which only *classify* their referents as something of a particular kind.

The obvious problem now for a functional account is to come to grips with the other end of the stick: true anaphoric definites. Since the point is that definite NPs are interpreted as functional concepts, we must construct local interpretations, for anaphorically definite NPs that are functional concepts and thus let the NPs refer unambiguously.

Although Löbner's account becomes somewhat less convincing at this point, his fundamental idea seems clearly correct. The interpretation of an anaphorically definite NP can indeed be turned into a functional concept when we avail ourselves of all currently relevant information, including the focus structure of the current discourse model.

Eventually, the difference between anaphora based accounts and functionality accounts of definiteness boils down to the question whether anaphorality or functionality is the more basic notion. Here we clearly side with Löbner: the fundamental idea of definiteness is unambiguous determination of the NP's interpretation or referent, in the same sense in which a function unambiguously determines its value. Anaphoric definiteness can be handled as a special case of this, but not conversely—not, at least, unless we are prepared to water down the notion of anaphora so as to make it useless for other purposes.

The basic notion of definiteness then that we assume here is the following:

- (i) *The definiteness of an NP implies that the information provided by the NP, together with contextually available information, determines a unique referent.*

Note that this formulation gives a necessary condition for definiteness, but not a sufficient one and thus does not provide a theory of definiteness. A theory of definiteness would also have to specify under what circumstances a definite referent than an indefinite NP must be used. Despite the clear availability of identifying information, we still find indefinite NPs in (7).

- (7) a. When I went to see my old teacher, I met a cynical geriatric with no illusions.
b. A tired and plainly listless Boris Becker lost the last two matches.

But our intention in this paper is not to provide a full-blown theory of definiteness. We only need to know enough about definiteness to be able to devise a model for the interpretation of definite NPs. Thus our task starts once we have found a definite NP. In this context, a principle of the form of (i) should be sufficient.

2. When are definite NPs resolved?

A crucial comment with regard to (i) is that the required contextual information must in principle be available already when the NP is encountered, and not at some later stage, like the end of the clause, the end of the sentence, or even the end of the paragraph. In a sense this follows from our acceptance of Löbner's view that the interpretation of a definite NP is a functional concept, together with the more general (and perhaps also more problematic) assumption that each word is interpreted when it is encountered. If the functionality of the NP's interpretation only comes about after additional text that follows the NP has been read, then the difference between definiteness and indefiniteness is blurred and it becomes unclear what the difference between the following two sentences would be:

- (7) c. The man I met that day turned out to be my old friend Fritz.
 d. A man I met that day turned out to be my old friend Fritz.

By the end of both sentences it is perfectly clear whom the initial NP refers to: my old friend Fritz. But this establishes neither the functionality of the definite NP in (7c) nor does it turn the indefinite NP in (7d) into a functional concept. The difference between (7c) and (7d) is just that in (7c) the speaker must assume that the initial NP can identify a referent, given the knowledge then available, whereas the speaker of (7d) assumes the contrary.

This is not to say though that the human processor could not use knowledge that is provided by the remainder of the sentence for the understanding of a preceding definite NP. Such delayed processing however would usually lead to distinct markedness of the discourse and should be regarded as exceptional. The corresponding "delayed processing" strategy then would look for a situation with respect to which the sentence (clause, paragraph) as a whole could be true and attempt to find the referent in that situation. But such a strategy would rather be typical for the processing of an indefinite NP and must, when used for the comprehension of definite NPs, be regarded as a repair strategy, i.e. as an attempt on the part of reader or listener to make up for an apparently mistaken assumption on the part of the writer or speaker.

Of course, people are also able to solve riddles, like finding out which famous historical person a particular text is about. But this kind of pastime should be clearly distinguished from linguistic processing of the garden variety.

When we now think of modelling the understanding of definite NPs, we may, on the one hand, subscribe to the view just outlined, but still, in our computer implementation, use what we referred to above as the two-stage approach, which postpones reference resolution until after the sentence in which the definite NP occurs has been analysed syntactically and the compositional semantics for the sentence has been constructed. The computational model we are proposing, on the other hand, attempts to resolve definite NPs immediately when they are encountered during the parse process. This interpretation strategy is motivated by considerations of computational efficiency as well as of psychological realism.

To begin with, there are fairly clear-cut psychological data to show that humans process definite NPs immediately. Komisarjevsky Tyler and Marslen-Wilson (1982) and Marslen-Wilson and Komisarjevsky Tyler (1987) have shown in experiments using on-line processing techniques that listeners quite generally make use of semantic and pragmatic information as soon as it becomes available. And – even more directly related to our issue – there are experiments by Crain and Steedman (1985) that show how the syntactic analysis is affected by what referents are contextually available for the definite NPs that are being processed.

One of Crain and Steedman's experiments uses materials as in (8):

- (8) *definite:*
 a. The teachers taught by the Berlitz method passed the test.
 b. The children taught by the Berlitz method passed the test.
indefinite:
 c. Teachers taught by the Berlitz method passed the test.
 d. Children taught by the Berlitz method passed the test.

Subjects were presented these sentences under appropriate experimental conditions with the request to judge their grammaticality. Sentences (8b) and (8d) were judged to be grammatical significantly more often than (8a) and (8c), which shows a clear influence of plausibility of the corresponding reading – which of course should never influence grammaticality judgments if semantic processing would start only after the completion of the syntactic analysis.

Further, sentences (8c) and (8d) were judged to be grammatical significantly more often than (8a) and (8b), which would indicate that the relative clause analysis that is required in these sentences is more readily accessible for indefinites than definites. With the definite NPs readers would more readily immediately attempt a referential interpretation (by accommodation, which is no problem, because we have no preceding text). Since there is no reason why this interpretation should not succeed, it may count as semantically confirmed and the first syntactically possible option for a definite NP (*The teachers/children* rather than *The teachers/children taught by the Berlitz method*) is assumed. What follows can then only be the main verb of the sentence. But this analysis turns the sentences into garden path sentences (cf. Bever 1970): the reader has reached the end of what according to his analysis is the complete sentence after *method* and is unable to extend this analysis to integrate *passed the test*; hence he can only judge the string ungrammatical.

If we may generalize from these experimental results, it would seem that listeners or readers guess at the referent of a definite NP as soon as it occurs, and correct their guesses only if the remainder of the discourse (in particular, the remainder of the current sentence or clause) is inconsistent with their earlier choice. In the case just discussed a correction would mean a fresh start: backtracking up to the incorrect decision that the NP was completed after the noun (*teachers/children*) and enlarging the NP by what now comes to be a relative clause rather than the VP of the sentence.

Now, guessing is probably not a matter of what linguists call "competence". But what if on-line linguistic comprehension is impossible without it? Psychological evidence indicates that our short-term working memory simply cannot hold all syntactically and semantically possible alternatives until the end of the sentence has been reached in order to then start upon a complex selection procedure, while, at the same time, the next sentence in the discourse must already be taken in to be processed. But then competence alone is insufficient for understanding linguistic utterances on-line.

Indeed, no sensible linguist has ever claimed the opposite. The idea of sticking a piece of competence into a performance model as if it were a processing module was no more than a first guess in computational linguistics.

tics, and entirely unpretentious as regards cognitive reality. What seemed to recommend this approach of first working out what is in theory possible and then letting semantics and pragmatics constrain the result is only the tidiness of modular programming. But the particular notion of modularity that would support this approach is only one of several and has no special advantage over other, more realistic, notions of modularity that serve computational practice and cognitive theory equally well.

Modularity offers us, in the case at hand, a clean relation between syntactic, semantic, and, if you wish, pragmatic theory on the one hand and the programmes in the mind or machine that implement these theories on the other. For "realist" syntacticians it is an important matter that their theories are not "about nothing", but about a mental capacity that does syntax and nothing else. But syntactic realists can have their way, and we may indeed programme a syntax module that is cleanly separated from other modules. All that our considerations require us to admit is a certain amount of interaction between modules (cf. Crain and Steedman 1985).

A second argument in favour of local definite NP resolution derives from considerations of computational efficiency. As already hinted, syntactic constraints alone generally cannot guarantee a unique syntactic analysis for a given sentence (cf. Hirst 1987 for a survey). There are many types of structural syntactic ambiguity, which, in cooperation with each other and with ambiguities of syntactic category assignment or morphological ambiguity, often produce a large number of theoretically possible analyses for a sentence. In parsing, where we cannot be sure that each attempted analysis that works for a certain substring of the sentence that is being analysed can eventually be extended to yield an analysis for the complete sentence (as is obvious also for the human processor e.g. in the garden path sentences above), the number of (local) ambiguities is even greater.

The earlier a syntactic processor thus comes to know that a particular analysis that it is pursuing will (with great likelihood or even with full certainty) not be the right one, the less time it will waste on constructing parses that are eventually not required.

Let us take as an example one particularly notorious ambiguity, which arises when, syntactically, there is more than one place to attach a prepositional phrase, as in

(9) The painter hit the writer with the brush on the head in the studio.

For most run-of-the-mill contexts the object NP in (9) would be *the writer* or perhaps *the writer with the brush*. Occasionally though, *the writer with the brush on the head* may be the right object NP, or even *the writer with the brush on the head in the studio*. It is clear that it is pragmatic factors that determine which syntactic analysis is, in the given context, the correct one. And one of these factors is whether or not the appropriate referents can be found for the various definite NPs that the syntactic analyses would require.

Now supposing that reference is only considered after the entire sentence

has been syntactically analysed, then, obviously, all syntactically possible analyses of the sentence must be figured out in detail and must be scored until a decision can be made. This requires a considerable amount of time and resources. In the case of (9) we have at least nine different possibilities of attaching the various PPs, all of which yield syntactically possible analyses of the complete sentence. But in most reasonably well-behaved contexts only one of these analyses will actually be at issue. All the rest are produced only to be thrown away – but before that happens they occupy even more time and resources, because semantic and pragmatic processes will have to deal with all of them in order to find out which one is best suited for the current context.

These observations suggest (cf. already Winograd 1972) that an immediate resolution of definite NPs would be a good idea not only for reasons of cognitive realism, but also for computational efficiency. If we try to resolve definite NPs referentially as soon as they occur, we can prune unpromising parses at a very early stage, which – assuming that the resolution process isn't too costly – would imply a considerable gain in efficiency (cf. also Mellish 1985).

Sure enough, early reference resolution may go wrong and discard just the analysis that is eventually needed, together with a whole set of incorrect analyses. We saw this happen to the subjects in Crain and Steedman's experiment. On the whole, however, from the point of view of global efficiency, it does not matter much when occasionally wasting time on doing something wrong and having to make a new start is balanced out by regular success. Machine processing of natural language, just like human processing, must strike a balance between efficiency and reliability.

3. *Where do referents come from?*

If a definite NP or, for that matter, any other expression, is to refer at all, its referent must be represented somewhere in some form or other. Since we may assume that there is not just one single form of mental representation, we must ask what sort of representation a definite NP can access.

Let us start with what is probably the most straightforward case. Just like an unmarked anaphoric pronoun a full definite NP may select its referent from a set of entities in memory that are attentionally prominent, or "in explicit focus" (cf. Bosch 1988), as we shall say:

- (10) a. Harry's long awaited book on Edward Wyman is out.
b. (*The book/It*) makes for rather laborious reading, because it has so many footnotes.

Here the definite NP *the book* seems to function exactly like the anaphoric *it*. The intended referent can be recovered easily because it was introduced by the subject term of (10a), and is therefore the most salient entity when (10b) is up for processing.

In (10) an entity is brought into focus with the help of linguistic means. Other means would have done equally well however. If the speaker sees the

hearer pick up Harry's book he could volunteer (10b) to precisely the same effect. In sum: Just like an anaphoric pronoun a definite NP can pick up any referent that is at the focus of attention. But while that is basically all an anaphoric pronoun can do, definite NPs can also access referents that are not focussed, as we shall see presently. Moreover, the degree of attentional prominence that a definite full NP requires of its referent is less than that needed for pronominal reference.

- (11) *Ferdinand I*, Emperor of the Holy Roman Empire, was served by another famous mechanic, Hans Bullmann of Nuremberg, who acquired great fame for his skill since he had learnt without instruction, and though unable to read or write had studied machinery and taught himself its secrets. He made several automata for [the Emperor / ?him] in the form of men and women.

Although the context makes it clear enough whom the speaker intends to refer to, a pronoun doesn't suffice here; it requires a full definite NP to get the job done. An intuitively plausible explanation for this observation is that, after the first sentence is processed, the Emperor is not sufficiently salient any more, i.e. is no longer in explicit focus. Therefore an unmarked anaphoric pronoun cannot ordinarily link up to this referent, whereas the definite NP can.

If a definite NP cannot find a referent in explicit focus, it may try to find another, unfocussed, object elsewhere in the discourse model. If we assume that entities receive a representation in the discourse model only when they are focal at a particular point in the discourse, then the discourse model consists of entities that either currently are in focus or were in focus at some earlier point in the discourse.

If a definite NP cannot find either a focussed or unfocussed referent in the discourse model, we have reached the end of what a narrowly understood anaphora-based account of definiteness would allow us to expect. But we have not yet reached the limits of definiteness. Discourse models are based on what is said in discourse and on what is presupposed. They cannot sensibly be assumed to contain representations of anything and everything that is relevant to a discourse. Discourse models are partial representations of what is actually relevant in discourse. This partiality explains why it is both necessary and possible to accommodate entities in the discourse model that were not explicitly introduced before.

As we already said, accommodation is constrained not only by the need to avoid contradiction with what is explicitly represented but, more generally, by the need to preserve coherence. One coherent extension of a discourse model that is always available, unless local considerations make different provisions, is by addition of elements from the global context to which the discourse context belongs. The global context that forms the implicitly understood background on which a discourse takes place is only partially represented in the discourse model, in that it is modelled only to the extent that is made explicit in the discourse. Since local considerations don't matter at the beginning of a discourse, Hawkins' (1978) "wider situation uses" of the

definite article are always permissible in discourse-initial position, and their reference is a matter of what is the most plausible global context extension of what is understood to be the current context. In this town then, Manfred Rommel is *the Mayor*; in this country Richard von Weizsäcker is *the President*; in this world John Paul II is *the Pope*, the sun is *the sun*, and so on.

The determination of the global context is a matter of the local context. And the local context is determined by preceding discourse or, by default, by where you think you are, or by what you think you are busy with, and by what you know about the wider context in which you think you are busy. Thus for a definite NP in search of a referent, the global context is determined by what parts of background knowledge the NP can access, starting out from the current discourse model. This, unfortunately, is about all we can say about global context at the moment.

If no referent for a definite NP can be found either in the current discourse model or in a global context that can count as its extension, another variety of accommodation may be tried that does not extend the discourse model globally but locally.

The entities to be accommodated then are not "given", as it were, with the discourse model as a whole, but with particular entities in the discourse model (on the basis though, again, of what we know about those entities by default and in general, i.e. on the basis of what is explicitly represented in background knowledge about entities of that kind). Cf. the following examples:

- (12) a. When I came home and entered *the kitchen*, I noticed *the fridge* was open.
b. (Situation: two people standing in front of a huge canvass covered with paint, one saying to the other.)
I wonder if *the artist* died of a natural cause?

Assuming the hearer is not already familiar with the speaker's kitchen, he will probably accommodate the kitchen as a part of the speaker's home, and, subsequently, the fridge as belonging to that kitchen. So what he must do is introduce new objects into his discourse representation to represent the kitchen and the fridge. Analogously for the artist in (12b), even though the painting on which the reference of *the artist* depends is not introduced linguistically but via the physical situation.

If one could always simply introduce new referents for definite NPs to refer to, every occurrence of a definite NP would be acceptable and interpretable, which is not the case (cf. van der Sandt's (1988:152) criticism of Lewis' notion of accommodation). Clearly accommodation must involve more than simply positing referents; what is needed in addition, as we have seen, is that these referents are linked up to something else that is already identified. The hearer should infer that the newly introduced fridge in (12a) belongs to the kitchen, the kitchen to the speaker's home, and that the artist referred to in (12b) is the one whose painting is contextually prominent.

Finding such cohesive links between referents presents an interesting problem in itself, which imposes non-trivial requirements on the representa-

tion of generic and specific knowledge and on inference procedures (cf. e.g. Charniak 1986 and Norvig 1987 for attempts to tackle such problems). But it is not a problem that we can delve into here. We are merely concerned with the type of referent that is involved in the process of interpreting NPs. In this connection we note that, as a rule, when a referent is to be successfully accommodated, it must either link up to a focussed referent in the discourse model or it must itself be an entity that forms part of the global context. We shall henceforth speak of *focal* and *global accommodation* respectively. Consider the following discourse (the first sentence of which is copied from example 11):

- (13) *Ferdinand I*, Emperor of the Holy Roman Empire, was served by another famous mechanic, *Hans Bullmann of Nuremberg*, who acquired great fame for his skill since he had learnt without instruction, and though unable to read or write had studied machinery and taught himself its secrets. But before that, he worked for three years as clerk (?at the town hall / ?there).

The definite NP *the town hall* cannot be focally accommodated via a link to the city of Nuremberg, because Nuremberg was mentioned too long ago and is not in focus any more. That neither Nuremberg nor any other place is in focus at the end of the discourse is demonstrated by the fact that also the (attempted) anaphoric *there* cannot be naturally interpreted. Also global accommodation of the town hall is not possible, since there is really no plausible way of embedding (13) into a global context that would contain a town hall. This leads us to the following constraint on the accommodation of referents for definite NPs:

- (ii) *The range of objects to which a referent can be linked via focal accommodation is the same as the range of objects from which an anaphoric pronoun can select its referent, i.e. the discourse referents that are currently in explicit focus.*
The range of objects to which a referent belongs that is introduced via global accommodation is the global context of the discourse.

Note that there is no way of linking a referent that is to be accommodated to an entity in the global context. Global accommodation is only by direct reference to the entity itself that is to be accommodated.

In summary: unmarked referentially occurring anaphoric pronouns (apart from some odd exceptions with no current interest, cf. Bosch 1988:211) can refer to objects in explicit focus only. Full definite NPs on the other hand can refer to:

- objects in explicit focus
- unfocussed objects in the discourse model
- objects that are part of the global context (global accommodation)
- objects linked via focal accommodation to objects in explicit focus

A condition on all these forms of interpretation of definite NPs is that the lexical information, together with the information provided by the various representations mentioned, must yield a referent-defining function. This is the core of what definiteness means. Note, by the way, that this approach also

covers the case of what Löbner regards as semantic definites, which we can construe via focal or, more usually, global accommodation.

From the preceding considerations we can straightforwardly derive a computational model for definite NP resolution. We saw in the previous section that definite NPs can get their referents from several representational structures. These representations are ordered in the sense that, if a suitable referent can be provided by more than one of these representations, there is a default preference that takes care of the selection. In particular, we propose a computational model for resolving definite NPs which examines, consecutively

- objects in explicit focus
- objects that can be accommodated in relation to explicit focus
- unfocussed objects in the discourse model
- objects in the global context

and which reports success as soon as the referent is found.

We would not want to claim that the sequentiality of the search process in this model is necessarily psychologically real. Parallel search is at least an equally good candidate for a psychologically realistic model. But then of course we may occasionally get more than one referent candidate and there must be a way of choosing among them. How this is to be done is a difficult question, but a heuristic that would probably be involved in such cases is that, if there are candidates for reference suggested by, say, the focus and the global context, the referent that is the most *pertinent* to the current purposes of the discourse should be preferred.

Now it seems plausible to assume that objects that are in focus are, *ceteris paribus*, more pertinent than objects that depend on focal objects via focal accommodation; that focally accommodated objects are, again *ceteris paribus*, more pertinent than unfocussed objects in the discourse representation, and so on. In short, we suggest that although a sequential model of the kind we propose may not be psychologically real, its net effects upon comprehension and acceptability are still by and large correct and hence linguistically real.

The procedure of search and selection just outlined should take place, as we have argued, in interaction with the parsing process. Avoiding technical details we can best describe this interaction with the help of an example. Suppose sentence (14a) is already processed and has resulted in a discourse model as in (15a).

(14) a. A writer was having a row with a painter.

(15) a.

x y
writer (x)
painter (y)
x was having a row with y

Now sentence (14b) is to be processed as a continuation of (14a) and hence with respect to the discourse model in (15a). The discourse model is here given

in the format of a Discourse Representation Structure (for short: DRS, cf. Kamp 1981).¹

(14) b. The writer hit the painter with the brush on the head.

The parser would, in a left-right order, propose that the string *the writer* could be analysed as a definite NP. Whenever the parser proposes such an analysis for a particular string, it would, before it goes any further, call up another module, which we have called the *NP-resolver*. The *NP-resolver* would now try to give an interpretation to the definite NP, starting its search in explicit focus. Since in the case at hand all conditions are met by the writer represented in the DRS, the parser's first proposal may count as semantically confirmed. The string is appropriately indexed for its interpretation with respect to the discourse referent, and the parser can proceed. The procedure will be entirely analogous when the parser comes across the following definite NP, *the painter*. A consequence of the semantic confirmation of the NP *the painter* is that further syntactic possibilities, such as *the painter with the brush* or *the painter with the brush on the head* are not even considered by the parser.

Now this should of course be different when (14b) occurs as a continuation of (14a) rather than (14a). (15a') is what the DRS would look like after the interpretation of (14a).

(14) a'. A writer was having a row with two painters. One of them kept poking at him with a brush.

(15) a'.

x	Y	y	z
writer (x)			
painters (Y)			
Y = 2			
x was having a row with Y			
y ∈ Y			
brush (z)			
y kept poking at x with z			

Here the parser's proposal to consider *the painter* as a definite NP is rejected by the *NP-resolver* because there are two painters in the DRS. The parser therefore continues and comes up next with the proposal of considering the string *the painter with the brush* as a definite NP. Although this requires a small amount of additional inferencing, the *NP-resolver* should be able to figure out that *the painter with the brush* can reasonably be interpreted by reference to the painter that kept poking at the writer with a brush, mentioned in the second sentence in (14a'). The remaining PP *on the head* can thus only be parsed as attached to the main verb, *hit*, and the contextually excluded analysis of *the painter with the brush on the head* is never even considered.

¹ In this paper we avoid any differentiation of salience within explicit focus, even though such a differentiation, as already hinted at in the discussion of example (1), is eventually necessary, certainly for the treatment of pronominal anaphora. A plausible computational framework for focus differentiation is found in Alishawi (1987) and in Asher and Wada (1988).

In general then, we would get the following procedure for the interaction between parser and *NP-resolver* (leaving out the precise nature of the relevant tests and heuristics):

1. The parser finds a string that is syntactically a definite NP
2. The parser calls the *NP-resolver*
3. The *NP-resolver* attempts to find a suitable referent for the proposed NP by considering consecutively the following options:
 - a reference to an object in explicit focus
 - b reference to an object that can be focally accommodated
 - c reference to an unfocussed object in the discourse model
 - d reference to an object in the global context
- 4a. If the search succeeds at either a, b, c, or d no further options are considered,
 - control is returned to the parser
 - the same syntactic analysis is pursued further on the basis of the assumption that the current string is correctly parsed as a definite NP
 - the NP is indexed for its reference
- 4b. If the search fails after d the proposed NP is considered semantically uninterpretable,
 - control is returned to the parser
 - the same syntactic analysis is not pursued any further and the analysis of the current string as a definite NP is rejected
 - the parser attempts alternative syntactic analyses

4. Some Remaining Problems.

Having presented the basics of our approach to definites and having argued for the correctness of its basic assumptions, we now want to list a number of problems that arise.

First of all, the approach does not automatically work for definite generic NPs. We hope that this class of cases can eventually be taken into account under the umbrella of "global context", but just saying this is of course not even a sketch of a solution.

Second, there is a problem in cases where the relevant referent for a definite NP is introduced earlier in the same sentence. In such cases, the referent is not yet represented in the DRS when the definite NP is encountered by the parser and the search will fail merely for this artefact reason. Since DRS construction, in current Discourse Representation Theory at least, cannot take place clausewise or word by word, but only at sentence boundaries, there is no immediate principled solution to this problem.

Third, the generalizations made above do not yet cover definite NPs with a demonstrative article (*this/that man*). In order to treat those cases, a more sophisticated focus structure is required than we are using in this paper. There are also special problems involved with the role of the lexical content of those NPs.

Fourth, as we have not in detail considered the inferencing that is carried out in the course of the reference search by the NP-resolver, we have also missed out the problems linked to the role of the lexical content of definite NPs in cases where the NP occurs coreferentially with another NP and their lexical contents are different (the relation may be anything in between straightforward synonymy and re-conceptualization or metonymy or metaphor).

These are surely not all the remaining problems, but they are those that we are aware of and know we still have to solve, though not in this paper.

There is one further interesting problem for which we would like to offer at least the sketch of a solution though. Since we have adopted a functional view of definiteness, it may seem that the following examples present a serious problem:

- (16) a. Helma took her poodle to *the vet*.
 b. As our hero approached the castle, he saw his beloved standing at *the window*.
 c. Theo broke *his leg* when he was mowing the lawn.

Sentences (16a)-(16c) can be interpreted perfectly well without the hearer having to assume that there are, in whatever sense, unique objects for the italicized definite NPs to refer to. There need not be one particular vet, contextually singled out or otherwise, for (16a) to be interpretable; (16b) does not imply that the castle has just a single window; and (16c) is compatible with the fact that Theo has more than one leg.

These observations are problematic for virtually all theories of definiteness, because they seem to show that, in the general case, uniqueness of reference is not a prerequisite for definiteness: it simply does not matter which leg Theo broke for (16c) to be interpretable. Analogously, it wouldn't usually matter to which vet Helma took her poodle, or at which particular window the lady in (16b) was standing.

The definite NPs in (16a)-(16c) are reminiscent of what Donnellan (1966) calls "attributive reference". Donnellan's notion of attributive reference is not quite appropriate to these cases however.

- (17) *The murderer of Smith is insane.*

Donnellan observes that the speaker need not have a unique referent in mind for (17) to be uttered felicitously, and that the definite NP in (17), when used by a speaker who does not know who murdered Smith, can be paraphrased as "whoever is Smith's murderer". Donnellan seems to suggest (but does not actually say) that Russell's original analysis should be restricted to attributively used definite NPs, and that the definite article can only be used because the phrase *the murderer of Smith* happens to refer uniquely. If that is what attributive reference is about (but it is by no means clear that Donnellan claims it is), then clearly the definite NPs in (16a)-(16c) do not refer attributively, simply because there may be more than one vet, window, and leg for the respective definite NPs to refer to. Hence referential uniqueness cannot be at stake here.

What then is at stake in these cases? We want to propose that it is uniqueness of roles, rather than of actual referents, that renders the definite NPs in (16a)-(16c) appropriate. In our view each of (16a)-(16c) evokes a situation type that is to some degree stereotyped, and it is by designating a unique role in such a stereotype that a definite NP can come about, even if it does not refer uniquely – that is, even if, or regardless whether, more than one object could fill the designated role.

Let us take (16c) as our first example. The situation of having a broken leg is a familiar one, and important to us because it interferes with our normal functioning in some way or the other. It is these consequences of having broken a leg that normally will be interesting to us – not the fact that it is one leg in particular. Moreover, our functioning is affected in different ways when we break an arm or a leg, and hence we are likely to have specialized stereotypes for having a broken arm and having a broken leg.

Let us assume then that there is a stereotype that allows people to reason about the implication of having broken a leg. Thus they would predict that normally the person who happens to be attached to the broken leg can't walk very well, has the leg in plaster, can't wear ordinary shoes, and so on. What is important is that there appears to be a cognitively useful way of describing what it means to have a broken leg that is *indifferent to which leg it is* that is broken. That is, there appears to be a broken leg stereotype with a single role for a leg. It is by designating this role that the definite NP in (16c) comes to "refer" uniquely.

It is perhaps not implausible that "going to the vet" evokes a stereotype in much the same way as "breaking one's leg" and that a similar explanation as the one just given applies to *the vet* as it occurs in (16a). That "standing at the window" may evoke a specialized concept too can be seen when we ask ourselves what we expect a person who stands at the window to do. Probably we would assume that he faces the window, not to look at it but *through* it. Accordingly, if someone concentrates his attention on the window itself, he is not standing at the window in the intended, stereotyped, sense. Compare (16b) with:

- (18) ?As our hero approached the castle, he saw his beloved standing at *the window*. She was cleaning *it*.

The discourse in (18) is awkward because the first sentence lets us expect that the lady in question is standing inside, looking through the window. The second sentence forces us to retract these inferences, making it clear that the preposition *by* should have been used instead of *at*.

In order to appreciate the relevant notion of role and interchangeability or arbitrariness of the actual filler of the role, consider the following story. Theo and Walter are milkmen who share a round, i.e. they work on alternating days. Since they are also both enthusiastic anglers, they always sit by the river with their rod when the other is working. Now imagine Mr. X saying to Mrs. X as she comes home from work wondering why there is no milk for the tea:

(19) The milkman didn't turn up today.

This utterance is perfectly normal and intelligible. But if Mr. X comes home from angling and utters (19), meaning that Theo or Walter was not out angling today, the utterance would be rather odd. With the notion of role designation that we have introduced, these observations can be explained if we assume that in the former context the definite NP *the milkman* is appropriate because Mr. X's utterance evokes a stereotype with a unique role for a milkman, whereas in the latter situation such a stereotype does not apply: after all for delivering or not delivering their milk, milkmen are pretty much interchangeable. But who says milkmen are equally interchangeable as anglers?

We have seen that the definite article may function in two rather different ways, which in English are not reflected in surface form. When we consider how one would translate sentences (16a) and (16b) into German, however, we see that languages sometimes do distinguish between both functions. For some prepositions and articles (in fact whenever phonology permits), German offers the choice between either cliticizing the article to the preposition or not (cf. Haberland 1985). E.g. for *zu* ("to") + *dem* ("the", third person singular masculine or neuter), German has the contracted form *zum*, and for *an* ("at") + *dem* there is *am*. Hence (16a) and (16b) can be translated as (20a) or (20b), and (21a) or (21b), respectively.

(20) a. Helma brachte ihren Pudel *zum* Tierarzt.

b. Helma brachte ihren Pudel *zu dem* Tierarzt.

(21) a. Als unser Held sich der Burg näherte, sah er seine Geliebte *am* Fenster stehen.
b. Als unser Held sich der Burg näherte, sah er seine Geliebte *an dem* Fenster stehen.

Semantically, the choice between these translations is not arbitrary. Normally (16a) would translate as (20a), and (16b) as (21a). (20b) would only be used if an anaphoric reference to a particular vet was intended. That is, for the interpretation of (20b) the identity of the vet is relevant, whereas for (20a) it is not. Analogously, (21b) is really only acceptable if some specific window is contextually given and singled out for attention. (21a), on the other hand, only requires generic information about castles and windows. In other words, contraction of a preposition and an article is possible if a role is at issue, but not if specific reference is involved.

In Fering, a North Frisian dialect, we find two definite articles, classified as the A-article and the D-article respectively (Ebert 1971, Keenan and Ebert 1973) and some German dialects of the Rhineland draw essentially the same distinction (Hartmann 1982). The contraction of preposition and article in Standard German follows the same pattern as does the distribution of the A- and D-articles in the dialects mentioned: in general, contraction in Standard German is possible only in situations where in the Fering or Rhineland dialects the A-article would be found (Haberland 1985). The D-article, on the other hand, requires that a referent be identified.

It is not clear whether these data can be fully explained with the help of the distinction we have drawn between reference to a role and reference to

particular individuals. The generic use of the definite A-article, for instance, is still quite open (as is the entire phenomenon of genericity). But it is clear that the distinction we have proposed² provides at least a partial explanation of the Standard German, Rhineland, and Fering data, and therefore receives some support from them.

Address of the authors:

Wissenschaftliches Zentrum der IBM Deutschland
Institut für wissenschaftsbasierte Systeme
Postfach 80 08 80
D-7000 Stuttgart 80
W.Germany.

Acknowledgements.

This paper reports on research carried on in the LILOG Project at the Scientific Center of IBM Germany at Stuttgart. The LILOG Project engages in research on Linguistic and LOGical Methods in automatic text understanding.

The authors wish to acknowledge the contribution of Matthias E. Kölln (now at the University of Hamburg) to initial stages of the work on NP resolution.

² A specific variety of roles in the current sense is discussed in Bosch (1983:141ff).

- Alshawi, H. (1987), *Memory and Context for Language Interpretation*, Cambridge University Press, Cambridge.
- Asher, N. and H. Wada (1988), "A computational account of syntactic, semantic and discourse principles for anaphora resolution", *Journal of Semantics* 6: 309-344.
- Bever, T.G. (1970), "The cognitive basis for linguistic structures", in J.R. Hayes (ed.), *Cognition and the Development of Language*, Wiley, New York: 279-353.
- Bosch, P. (1983), *Agreement and Anaphora*, Academic Press, London/New York.
- Bosch, P. (1988), "Representing and accessing focussed referents", *Language and Cognitive Processes* 3: 207-231.
- Charniak, E. (1986) "A neat theory of marker parsing", *Proceedings AAAI* 86: 584-588.
- Crain, S. and M. Steedman (1985), "On not being led up the garden path: the use of context by the psychological syntax processor", in D.R. Dowty et al. (eds.), *Natural Language Parsing*, Cambridge University Press., Cambridge: 320-358.
- Donnellan, K.S. (1966), "Reference and definite descriptions", *Philosophical Review* 75: 281-304.
- Ebert, K.H. (1971), *Referenz, Sprechsituation und die bestimmten Artikel in einem nordfriesischen Dialekt*, Nordfriisk Institut, Bredstedt.
- Haberland, H. (1985), "Zum Problem der Verschmelzung von Präposition und bestimmtem Artikel im Deutschen", *Osnabrücker Beiträge zur Sprachtheorie* 30: 82-106.
- Hartmann, D. (1982), "Deixis and anaphora in German dialects: the semantics and pragmatics of two definite articles in dialectal varieties", in J. Weissenborn and W. Klein (eds.), *Here and There. Cross-Linguistic Studies on Deixis and Demonstration*, Benjamins, Amsterdam: 187-207.
- Hawkins, J.A. (1978), *Definiteness and Indefiniteness*, Croom Helm, London.
- Hirst, G. (1987), *Semantic Interpretation and the Resolution of Ambiguity*, Cambridge University Press, Cambridge.
- Kamp, H. (1981), "A theory of truth and semantic representation", in J. Groenendijk et al. (eds.), *Formal Methods in the Study of Language*, Mathematical Centre Tracts, Amsterdam: 277-322.
- Keenan, E.L. and K.H. Ebert (1973), "A note on marking transparency and opacity", *Linguistic Inquiry* 4: 421-424.
- Komisarjevsky Tyler, L. and W. Marslen-Wilson (1982), "Processing Utterances in Discourse Contexts: On-line Resolution of Anaphors", *Journal of Semantics* 1: 297-314.
- Lewis, D. (1979), "Scorekeeping in a language game", in R. Bäuerle et al. (eds.), *Semantics from Different Points of View*, Springer, Berlin/New York: 172-187.
- Löbner, S. (1985), "Definites", *Journal of Semantics* 4: 279-326.
- Marslen-Wilson, W. and L. Komisarjevsky Tyler (1987), "Against modularity", in J.L. Garfield (ed.), *Modularity in Knowledge Representation and Natural-Language Understanding*, MIT Press, Cambridge, Mass: 37-62.

- Mellish, C.S. (1985), *Computer Interpretation of Natural Language Descriptions*, Ellis Horwood, Chichester.
- Norvig, P. (1987), *A Unified Theory of Inference for Text Understanding*, Report No. UCB/CSD 87/339, Computer Science Division, University of California at Berkeley.
- Sandt, R.A. van der (1988), *Context and Presupposition*, Croom Helm, London.
- Winograd, T. (1972), *Understanding Natural Language*, Academic Press, New York.
- Woods, W.A. et al. (1972), "The Lunar Sciences Natural Language Information System", Final Report, Report No 2378. BBN.