DANDELION

Discourse Functions and Discourse Representation:
An Empirically and Linguistically Motivated,
Interdisciplinarily-Oriented Approach to Natural Language Texts

ESPRIT Basic Research Project 6665

Deliverable R1.4.1

Discourse functions of NP-anaphora

Linguistic and text-analytic form/function analysis of NP-anaphora

Lettica Hustinx, Fons Maes, Corine Schouten, Wietske Vonk

October 1993

CLS Tilburg/Nijmegen

University of Edinburgh — GMD-IPSI Darmstadt — Universität Saarbrücken

Universidad Complutense Madrid

Document ID	:	Deliverable R1.4.1: *Discourse functions of NP-anaphora*
Issue date	:	October 1993
Responsible	:	CLS Tilburg/Nijmegen
Version	:	Final
Status	:	Public
WP number	:	WP1: Empirical studies
Distribution	:	DANDELION Consortium and CEC
Keywords	:	Marked anaphoric reference, overspecified reference, accessibility, definite nominal anaphors, indefinite nominal anaphors, demonstrative nominal anaphors

©1993 The DANDELION Consortium

Discourse functions of NP-anaphora

Linguistic and text-analytic form/function analysis of NP-anaphora

Lettica Hustinx, Fons Maes, Corine Schouten, Wietske Vonk

CLS Tilburg/Nijmegen

Abstract:

This paper presents a functional analysis of nominal anaphors (NAs) in discourse. We survey the area of NP-anaphora starting from the traditional view on anaphoric expressions, according to which NP-anaphors form an accessibility hierarchy and NAs provide anaphoric access to discourse referents with low accessibility. This 'identificational view' is then challenged with empirical evidence showing various non-identificational discourse functions of definite, indefinite, and demonstrative NAs. NAs are shown to participate not only in the management of reference, but also in other semantic and pragmatic discourse functions at local and global levels. They express interactive relations between writer and reader, they contribute to the semantic representation of discourse, and they express structural thematic relationships in discourse.

Contents

1 The accessibility view of NP-anaphora **5**

1.1 Referential expressions and accessibility 5

1.2 Anaphoric hierarchies 5

1.2.1 The text-analytic requirement 7

1.2.2 The processing requirement 7

2 Exploiting and complementing the accessibility hierarchy for other discourse functions **8**

3 Definite nominal anaphora **11**

3.1 Unmarked occurrences of definite nominal anaphora 11

3.2 Overspecified occurrences of full definite nominal anaphora 12

4 Indefinite nominal anaphora **14**

4.1 Unmarked occurrences of indefinite expressions 14

4.1.1 Referential indefinites 15

4.1.2 Predicating indefinites 16

4.2 Marked use of indefinites: the referential function 17

4.3 Marked use of indefinites: the attributive function 20

5 Demonstrative nominal anaphora **23**

5.1 The interpretation perspective 23

5.2 The analytic basis of the demNA-proposals 24

5.3 The definite-demonstrative dichotomy in discourse 25

5.3.1 The MOD-proposal 25

5.3.2 Classifying the referent 26

5.3.3 Attributing information to the referent 27

5.3.4 Contextualising the referent 29

5.4 The dezeNA – dieNA variation in discourse 30

5.5 A two-way interpretation process for demNAs 33

6 Conclusion **33**

1 The accessibility view of NP-anaphora

1.1 Referential expressions and accessibility

Studies of NP-anaphora have traditionally been concerned only with the role of anaphoric expressions in the establishment and maintenance of reference in discourse, that is, with the identificational function of those expressions. The lexical form, particularly the degree of specificity, of an anaphoric expression is considered to reflect the accessibility of its referent at that point in the discourse. The lowest degrees of lexical specificity, realized by zero anaphora and pronouns, are appropriate for easily accessible, highly activated referents, while very specific expressions, such as descriptions and proper names, are reserved for referents with low accessibility.

The degree of accessibility of a referent can be influenced by a number of factors. The three most frequently observed and discussed factors are the number of competitors for the role of discourse referent, the distance between anaphor and referent, and the topical status of the referent in the text.

The most obvious factor affecting the accessibility of a referent is the presence or absence of *competing candidates*, to which the anaphor may refer. It has often been observed that when an anaphor occurs in an ambiguous context, i.e., when it can refer to two or more possible antecedents, a lexically more specific expression is preferred to the pronoun in order to avoid referential ambiguity (e.g., Ariel, 1988; Brown, 1983; Clancy, 1980; Givón, 1983).

The second important factor that affects accessibility is that of *anaphoric distance*. Only a relatively short distance between the anaphoric expression and the antecedent allows pronominal reference, whereas anaphora referring over longer distance contain more lexically explicit information (Ariel, 1988; Brown, 1983; Clancy, 1980; Givón, 1983, 1987).

The *topical status* or the role of the referent in the text may also influence the type of referential expression: Ambiguous pronouns are more easily assigned to subject antecedents than to object antecedents (Clancy, 1980; Frederiksen, 1981). New characters are mainly established by full noun phrases, and subsequently referentially maintained by pronouns (Ariel, 1988; Chafe, 1976; Clancy, 1980; Garrod & Sanford, 1982; Karmiloff-Smith, 1985; Marslen-Wilson, Levy & Tyler, 1982; Murphy, 1984). Francik (1985:29) shows that protagonists in narrative discourse are more often and more easily referred to by means of pronouns than non-protagonists, i.e., protagonists are more salient in a discourse than non-protagonists (see also Anderson, Garrod, & Sanford, 1983; Karmiloff-Smith, 1980:248). The establishment of a character with a proper name gives that character more prominence compared to the establishment with a role description (Sanford, Moar & Garrod, 1988). Furthermore, only globally important referents are introduced with a proper name (Givón, 1990).

1.2 Anaphoric hierarchies

Above, we described the inverse relationship between coding material and referent accessibility as a predictive principle explaining the distribution of anaphoric types. This relationship

can be seen as the organizing principle behind several rankings of different types of referential expressions according to their identificational value, i.e. referential hierarchies, that have been proposed in the literature.

These hierarchies differ in that their point of departure is either a ranking of degrees of accessibility of the referents or a ranking of expression types.

The first hierarchy type can be found in, for instance, Prince's (1981) *Given-new taxonomy* and in Gundel, Hedberg & Zacharski's (1989) *Givenness-hierarchy*. The starting point for these hierarchies is the assumed cognitive status of discourse referents. These statuses can be linked to types of referential expressions in a strict or in a loose fashion. The linking is strict when cognitive statuses are in a one-to-one correspondence related to types of referential expressions. This is the case in Gundel et al.'s hierarchy in which five cognitive statuses are determined in terms of necessary and sufficient conditions for the use of types of expressions. For example: The cognitive status *familiar* is a necessary condition for the use of all definite demonstratives and a sufficient condition for the demonstrative determiner *that*. Statuses and expression types are loosely linked in Prince's given-new taxonomy, in which the distinct entity-types in the taxonomy are merely exemplified with, but not tied up to types of referential expressions.

The second type of hierarchies starts from a set of expression types, ranked according to the assumed degree of accessibility of their underlying referent, such as the *anaphoric hierarchy* in Lakoff (1976:295), the *code-quantity scale of referential predictability* in Givón (1989:217) and the *accessibility marking scale* in Ariel (1990:73). These hierarchies differ in their fine-grainedness in that they range from four distinctions (Lakoff) to fifteen (Ariel), and they differ in the way in which the hierarchies are accounted for.

The most elaborate hierarchy is the anaphoric hierarchy in Ariel (1990). Her ranking of anaphora is as follows:

Low Accessibility
- full name + modifier
- full ('namy') name
- long definite description
- short definite description
- last name
- first name
- distal demonstrative + modifier
- proximal demonstrative + modifier
- distal demonstrative (+ NP)
- proximal demonstrative (+ NP)
- stressed pronoun + gesture
- stressed pronoun
- unstressed pronoun
- cliticized pronoun
- extremely high accessibility markers (gaps, including pro, PRO and wh-traces, reflexives, and agreement)

High Accessibility

When we look at this comprehensive list as an implementation of the accessibility view, we expect it to meet the following two requirements:

– Text analytic requirement: It should reflect the functional behavior of these types of NPs as can be analyzed in discourse.

– Processing requirement: It should provide language users with a heuristic device in producing and comprehending NP-anaphora in discourse. I.e. in interpreting or producing NPs, people should draw heuristic information from the position of each NP-type within the hierarchy.

1.2.1 The text-analytic requirement

According to Ariel, the NP-hierarchy reflects the fact that on average when going down in the hierarchy the distance in the text between these anaphoric types and their antecedents decreases. So, referential distance is taken to be a critical determiner of referent accessibility. Referential distance, however, is not a sufficient explanation for all occurrences. There are other contextual factors determining the degree of accessibility as well (as Ariel herself remarks) such as interference or topic persistence. These variables can (and should) be quantified as well in order to be able to give an explanation of all occurrences of NP-anaphora.

In addition, the position of an expression type in the hierarchy results from some kind of generalization over or reduction of the different uses of each particular expression type. The positions in Ariel's hierarchy are empirically based upon the overall average referential distance of all tokens of the same expression type in a text corpus. In Gundel's hierarchy, the position is determined by the most typical use of the expression, i.e. the use which distinguishes it from other expression types. This 'average' or 'typical' function expressed by the position in the hierarchy might mask important subdivisions or differences in use within one expression type. In principle, of course, this is not a problem for an analytic approach. On the contrary, it asks for a more fine-grained analysis.

As will be shown in the next sections, the accessibility as reflected in the anaphoric hierarchies does not completely capture the distribution of the different NP-anaphor types. NP types which are not in agreement with the accessibility hierarchy will be shown to have other discourse functions in addition to the identificational function. An explanation for these occurrences should be provided.

1.2.2 The processing requirement

A plausible way in which the hierarchy is used in *producing* and *interpreting* anaphoric expressions is that the accessibility marking degree of the present expression type and the accessibility of the discourse referents are matched. The degree of accessibility of the referent determines which NP-anaphor type is used. The less accessible a referent is in the context, the greater the lexical specificity of the NP-anaphor. The intuitive plausibility of this heuristic could be supported by differences in referential adequacy when pronominals are replaced by full definite NAs, and vice versa.

However, looking at Ariel's hierarchy, it is totally unclear what such accessibility considerations look like in the case of for instance a full name (second on the list), which is evaluated against a full name + modifier (first on the list) on the one hand and a long definite description (third on the list) on the other, or in the case of a demonstrative NA which is evaluated against a definite description. Therefore it is not plausible that the detailed list reflects noth-

ing but differences in referential accessibility. As a consequence, it cannot function as an adequate tool in a processing model.

In the next sections we will investigate the applicability of the following hypothesis. Readers interpret an expression in relation to the contextual conditions for the use of that expression concerning accessibility. If an expression is more (or less) specific than is required for identification in the present context, this mismatch between the specificity of the expression and the contextual conditions may suggest to the reader that the expression has another discourse function in addition to the identificational function. For example, the use of a name in a context in which the referent is highly accessible – in which case a pronoun would have sufficed for identificational purposes – may provide the reader with information about the thematic structure of the discourse.

2 Exploiting and complementing the accessibility hierarchy for other discourse functions

In this section we will investigate the assumption that a mismatch between the specificity of the anaphoric expression and the contextual conditions for identification has an additional function. If an anaphor is more specific than needed to unambiguously identify the intended referent, this expression is identificationally overspecified. The difference in form between such an overspecified expression and its less specified counterpart corresponds to a difference in function. The additional function depends on the standard identificational function of the NP-expression concerned.

The accessibility hierarchy is based on the assumption that the speaker/writer takes the assumed accessibility status of the referent in the discourse model of the addressee into account when choosing the referring expression. This assumption reflects pragmatic considerations of the speaker / writer, which have been sketched by Grice (1975) and further developed by Levinson (1983, 1987a, 1987b). Given the degree of accessibility of the referent, the speaker uses the least specific expression: the minimal expression that will do the identificational job. This agrees with the Cooperation principle, in particular with the maxim of Quantity, which consists of the following principles:

(i) Make your contribution as informative as is required for the current purpose of the exchange

(ii) Do not make your contribution more informative than is required

The latter is developed into the more specific I-principle, the maxim of Minimization, developed by Levinson (1987a, 1987b):

– Say as little as necessary, that is produce the minimal linguistic information sufficient to achieve your communicative ends.

and from the recipient's side:

– Assume that stereotypical relations obtain, unless this is inconsistent and the speaker has broken the maxim of Minimization by choosing a prolix expression.

When applied to the accessibility view on referring expressions, these principles explain the use of referring expressions based on accessibility assumptions. The production of these expressions is guided by the idea that they should be tuned to the accessibility status of the referent for the addressee. Similarly, in *comprehending* people assume that the referring expressions are produced according to the accessibility of their referents.

In addition, these conventions can explain the use of *marked* referring expressions, that is expressions that are not in agreement with the maxim of Minimization. In such circumstances, hearers / readers assume that the Cooperation Principle nevertheless still holds. Because the maxim of Minimization has been broken, the readers assign an additional, non-referential meaning to the anaphoric expression. They have to make an inference in order to understand the apparently additional meaning of the marked expression.

In this view of NAs we claim the following:

– when NAs are used in identificationally unproblematic discourse conditions, such that the NA is overspecified, the overspecification produced is used to perform non-identificational (pragmatic, semantic, topical/attentional) functions in discourse. The functional value of the overspecification is considered to depend on the standard (identificational) function of the expression type involved.

– the heuristics of NP-interpretation are not only based on a ranking of the expression type within a whole hierarchy but also on 'privileged substitution variants'. This means that in interpreting a nominal anaphor a reader in the first instance draws heuristic information from the fact that the expression type chosen is a substitution variant of one or more other expression types. For example, if a reader comes across a demonstrative nominal anaphor in discourse where a pronoun would be sufficient, the reader is provided with the heuristic information that the demonstrative NA is overspecified. As an overspecified expression, the demonstrative NA is further a substitution variant of the definite NA-version which provides the reader with heuristic information about the non-identificational function of the demonstrative NA, as we will show in section 5.

In the following sections we will concentrate on three types of nominal anaphora: full definite NAs and proper names, indefinite NAs, and demonstrative NAs. These types of NAs can all be used as an overspecification, violating the accessibility predictions in the same way. We will give a short illustration of each of these overspecifications.

As for the *proper name*, consider the following sentence sequence:

(1) *Mr. Jones* was walking down the street. *He* was carrying a heavy suitcase. *He* was on *his* way to the railway station. *He* was in a hurry.

If the last pronoun is replaced with a full definite noun phrase or a proper name, the interpretation of the last sentence changes in that it seems less continuous with the current theme:

(2) *Mr. Jones* was walking down the street. *He* was carrying a heavy suitcase. *He* was on *his* way to the railway station. *Mr. Jones* was in a hurry.

According to the accessibility scale a proper name identifies a less accessible referent than a pronoun. So, using a proper name instead of a pronoun to refer to a highly accessible referent, yields an overspecification. The default function of a proper name is to establish a referent or to recover a less accessible discourse referent. In the above example, however, the discourse referent has been established in the first sentence and is re-established in the last sentence by means of the overspecified proper name. As a consequence, the name anaphor induces a change in the thematic structure of the text. This will be taken up in section 3.

For the use of *indefinite NPs*, Ariel's accessibility hierarchy does not entail any predictions, as indefinites are not considered to be anaphoric expressions. Since indefinite NPs are normally used to introduce a referent into the discourse, they can be considered to express the lowest possible degree of accessibility, namely "inaccessibility". Therefore we propose to include indefinite expressions in the accessibility hierarchy, namely at the low accessibility pole of the scale. As will be shown in section 4, indefinite expressions can even be used to refer to a referent that has already been established in the discourse. Consider the following example:

(3) For several years, Mrs. H. T. Miller had lived alone in a pleasant apartment. [..] She was a widow. [..] She had no friends to speak of. [..]
Then she met Miriam. It was snowing that night. Mrs. Miller had finished drying the supper dishes and was thumbing through an afternoon paper when she saw an advertisement of a picture playing at a neighborhood theater. The title sounded good, so she struggled into her beaver coat [..] and left the apartment. [..]
A long line stretched in front of the box office; she took her place at the end. [..] The line seemed to be taking its own time and, looking around for some distraction, she suddenly became conscious of *a little girl standing under the edge of the marquee.*
Her hair was the longest and strangest Mrs. Miller had ever seen. [..]
(Truman Capote, 'Miriam', In *Contemporary American Short Stories*; cited in Ushie, 1986)

The reason for the use of the indefinite in this example exceeds the purely identificational function. According to its identificational function, this indefinite expression should introduce a referent into the discourse. However, since the referent has already been established, the indefinite expression is interpreted as reflecting an embedded domain in the discourse in which the referent in question is re-introduced.

As for the third type of nominal anaphor, the *demonstrative NAs*, consider the following examples:

(4) *Bush* has decided to raise taxes after all. It is the first time {*the / this*} *president* has dared to do so.

(5) *The sparrow* is very common in our low countries. {*The / this*} *bird* doesn't need special protection.

In these examples, the use of a demonstrative NA activates a non-default interpretation of the NA concerned. It effects a modification in the semantic representation of the referent: In the Bush-example other presidents are evoked and in the sparrow-example other birds.

The accessibility view can contribute to the explanation of this use of the demonstrative NA. The demonstrative NA constitutes an overspecification in comparison with a pronoun. The mismatch of the expression with the accessibility of the referent is used to convey additional information. The demonstrative aspect of the demonstrative NA triggers an additional meaning due to the markedness-function of the demonstrative. One effect of the markedness of demonstratives is that the reader amplifies the information about the referent by inferring the members of the class indicated by the nominal part of the demonstrative NA. This is the case in the examples above. However, there are other amplifications as well, as will be shown in section 5.

3 Definite nominal anaphora

3.1 Unmarked occurrences of definite nominal anaphora

A full lexical expression has one of two basic functions: either to establish a discourse referent that does not exist in the text representation, or to recover a referent that is present but less accessible in the text representation. The establishment of characters in a story, for instance, is most frequently realized with a proper name or role description, and reference is usually maintained by pronouns (e.g., Karmiloff-Smith, 1985; Marslen-Wilson, Levy & Tyler, 1982). Anaphora with more identificational information than pronouns are required to maintain reference in cases with competing candidates, long anaphoric distance, or low topical status of the referent in the text, as is illustrated in examples (1), (2) and (3). In example (1) two competing candidates for reference occur in the first sentence. The use of a pronoun in the second sentence would lead to two possible interpretations (although the real-world referent in this case would be the same for both). So, a name that unambiguously identifies the correct antecedent is used.

(6) *Mr. Flory*, de hoofdpersoon in deze roman, is in vele opzichten een zelfportret van *Blair*. *Flory* verafschuwt een systeem van onderdrukking waarvan hij deel uitmaakt.
Mr. Flory, the principal character in this novel, is in many respects a self-portrait of *Blair*. *Flory* detests a system of suppression, which he is part of.

When referring to a discourse referent which was last mentioned a number of sentences earlier, the referential distance has to be overcome with a nominal expression. In (2) the long anaphoric distance even requires an extension of the anaphor:

(7) *De bewoner ervan* beriep zich op een langduring koopcontract. De overeenkomst stond geregistreerd, maar het bestaan ervan kwam eerst in het geding nadat Einstein zijn toekomstige bezittingen was komen bezichtigen. ... [twee zinnen]...
De bewoner van het voor Einstein bestemde huis had nog andere pijlen op zijn boog.

The occupant appealed to a long term purchase contract. The agreement was registered, but its existence came only into play after Einstein had come to inspect his future property. ... [two sentences] ... *The occupant of the house intended for Einstein* had still more strings on his bow.

As for the influence of topical status on the form of the referential expression, Karmiloff-Smith (1985), for instance, noticed that children preferred to use a pronoun to refer to main protagonists, and full nouns to subsidiary protagonists. In the following example the *balloon man* requires a definite NP remention.

(8) *A little boy* is walking home. *He* sees *a balloon man. The balloon man* gives *him* a green balloon, so *he* happily goes off home with it. But the balloon suddenly flies out of *his* hand and so *he* starts to cry. (p. 72)

3.2 Overspecified occurrences of full definite nominal anaphora

When an anaphor is used that is more lexically specified than can be justified by any accessibility factor, as illustrated above, the anaphor is considered an overspecification.[1] In (9) the use of an overspecified anaphor is illustrated.

(9) *Hij* was gefascineerd door de mensheid en haar kinderrijk en door middel van gedichten, liederen, toneelstukken, proza, maximen en korte beschouwingen gaf *hij* uitdrukking aan *zijn* ideeën, vragen en worstelingen. *Goethe* was, naast kunstenaar, ook een verdienstelijk natuurwetenschapper.
He was fascinated by human nature and its children, and *he* expressed *his* ideas, questions, and struggling by means of poems, songs, plays, prose, maxims, and short essays. *Goethe*, besides being an artist, was also a meritorious physicist.

The use of the proper name in the last sentence is not necessary for identification of the referent and is consequently an overspecification. The proper name signals the theme shift in the last sentence. The signaling effect of the overspecification becomes apparent if the proper name is replaced with a pronoun: The sentence becomes awkward.

The phenomenon of overspecification has been observed in a number of studies, yet has not been interpreted in terms of overspecification.

In a discussion about the relationships between full noun phrases and pronouns in a discourse context, Hinds (1977) observed that in so called 'peak sentences' in paragraphs, in which the most important information of the paragraph is contained, mostly full noun phrases occur, whereas pronouns occur in non-peak sentences. The presence of these noun phrases was not determined by the accessibility of their antecedents.

For English narrative discourse, Clancy (1980) observed that lexically more specified forms were used at episode boundaries, even in the "unusual cases in which noun phrases were used

[1] Overspecification is defined in terms of the accessibility of the intended referent. It should be noted, however, that an overspecification is not defined in relation with the lexical form of its preceding referring expression (although in most instances the overspecification will consist of more lexical material than its preceding referring expression).

within very short intervals when there was no question of interference" (p. 71). She made the same observation at episode boundaries in Japanese narratives, where boundaries were marked by a switch from ellipsis to nominal reference for the same character.

The correspondence between the form of the referring expression and the structure of the text has also been investigated experimentally. In a production experiment of Marslen-Wilson, Levy and Tyler (1982) it was shown that the degree of lexical specificity of expressions referring to the main actors in a story depended on the level of information in the narrative structure. Apart from the use of lexically more specified anaphora where less specified reference would fail to correctly identify the intended referent, lexically more specified expressions (pronouns instead of zero anaphora, proper names instead of pronouns) were produced at episode and event boundaries. In addition, they were used to refer to actors that were not 'in focus'.

In a study by Fletcher (1984) it was demonstrated that a referential expression was produced as a function of topic continuity: Pronouns and zero anaphora were used in a continuous context and nominal anaphora when there was a theme shift.

In other studies, overspecification was acknowledged as a discourse structuring device. As Bates, Kintsch, Fletcher, and Giuliani (1980) suggested: "It may be that the decision to use an explicit form of reference involves the need to highlight, dramatize, mark topic shifts and important points, and generally "stage" utterances in a way that assures the listener's attention. Hence explicit forms tend to be used at high-interest points (even when they are not necessary for disambiguation)" (p. 47).

Fox (1984) argued that full NPs are used to demarcate new narrative units. Stark (1988) observed that paragraph boundaries appeared to be signaled by 'over-reference', i.e., using a full form when a pro-form would be sufficient.

In example (9), the theme shift in the last sentence is obvious: The theme of Goethe as an artist changes into the theme of Goethe as a scientist. The use of the overspecification is in agreement with the content of the sentence. An interesting issue is whether the use of an overspecified anaphor merely signals a theme shift in a discourse, or whether the anaphor itself induces such a shift. Then, the overspecification not only signals that the (highly accessible) discourse referent has to be re-established, but also invites the reader to consider the incoming information as belonging to a new chain of information.

This interpretation of the overspecified form entails an even stronger claim than described by Fox (1984), Stark (1988), Marslen-Wilson, Levy & Tyler (1982), and Fletcher (1984): Instead of interpreting the form of a referential expression as a function of the discourse structure, the form of the device itself is claimed to structure the discourse. An overspecified anaphor is regarded as a signal for the reader to set up a new chain of information. In this sense, an overspecified NA can be considered a boundary marker.

The claim that an overspecification can function as a boundary marker has already been corroborated in a number of experimental studies, in which the use of various forms of anaphora was investigated as a function of thematic (dis)continuity (Hustinx, in prep.; Vonk, Hustinx, & Simons, 1992). First, in an experiment in which subjects had to segment expository texts into paragraphs, overspecified NAs were found to occur predominantly in sentences that had

been indicated by the subjects as paragraph-initial. Second, in a production experiment, subjects had to write stories based on strip cartoons about one protagonist. In a critical picture the theme continued or shifted. When the theme continued, subjects produced more pronominal expressions in the first sentence describing the critical picture. In case of a theme shift in the strip cartoon (in which an action was depicted that occurred at a different time and place from that in the previous picture) boundary markers were used. Apart from preposed expressions of time and place, overspecified anaphoric expressions were used as boundary markers.

The stronger claim that an overspecified anaphor itself functions as a thematic structuring device has been investigated in a controlled production experiment and a series of comprehension experiments. In the production experiment it was demonstrated that an overspecification itself was *interpreted* by subjects as a boundary signal. In this experiment the subjects had to continue fragments of stories about one protagonist. They had to start the sentences with either a pronoun or an overspecified NA. It was demonstrated that when subjects had to continue the fragment with an overspecification, they were led to produce a theme shift, whereas they preferably produced theme continuations if they had to continue with a pronoun.

In the comprehension experiments it was investigated whether an overspecified NA affects the construction process of the text representation. The overspecified NA, functioning as a boundary marker, signals to keep the incoming information apart from the preceding information and not to integrate it with the preceding information. The boundary marker opens up a new chunk of information in which the incoming information can be stored. As a consequence the concepts in the sentence before the anaphoric reference should be less available when the anaphor was overspecified and better available when it was not. This was tested with a probe recognition task, using concepts other than the anaphors' referents. The availability of those concepts was shown to decrease after reading the overspecified anaphor.

These experimental studies corroborate the view that NA's that are more lexically specified than is required for identification in their present context, serve a thematic structuring function in addition to their identificational function.

4 Indefinite nominal anaphora

4.1 Unmarked occurrences of indefinite expressions

An indefinite expression can be used in several ways. It is generally acknowledged that an important function of an indefinite NP is to introduce a referent in the text that has not been mentioned so far. In terms of accessibility, an indefinite expression accesses a referent that has the lowest possible accessibility status. A second way to use an indefinite expression is to predicate something about a discourse referent, as in *John is a smart man*. In this example, the indefinite is used in a nominal predicate to attribute a property to the subject John. In contrast to the introductory indefinite, this indefinite expression is non-referential.

In the following subsections, the regular use of referential and predicating (attributive) indefinites will be described first. After that, it will be shown that indefinites can also be used to

refer to already identified entities, thereby violating the accessibility hierarchy. The functions of these marked indefinites will be shown to be derived from their regular functions of either introducing or predicating.[2]

4.1.1 Referential indefinites

Usually, referential indefinites are divided in three main classes, based on the kinds of entities they refer to. First, an indefinite expression may refer to a specific entity, as in the sentence *I saw a man walking in the street.* Second, it is possible to talk about a non-specific entity by means of an indefinite, as in *I was looking for a book about indefinites.* This sentence is actually ambiguous between a specific and a non-specific interpretation: either a particular book was being looked for, or someone was looking for any book about indefinites. This ambiguity can only be solved when the context in which the sentence occurs is taken into account (it depends on the verb, for instance). A third type of entity an indefinite expression can refer to is a generic concept, not denoting a specific or non-specific entity, but a class of entities as a whole. An example is *A horse has four legs.*

The function of these referential indefinites is to refer to concepts that are for some reason not uniquely identifiable to the recipient. A definite expression would not be appropriate in these cases, because it would imply that the referent was already known (see Hawkins, 1978; 1991). For instance, a specific indefinite expression, as in *I saw a man walking in the street*, cannot be replaced by a definite expression because a referent has not been mentioned in the discourse before and is therefore not identifiable. In the case of a non-specific indefinite, as in *I was looking for a book, but I couldn't find one*, the reason for the use of the indefinite is not in the first place that the referent has not been mentioned before, but that the referent is not uniquely identifiable because it is non-specific. Indefinites referring to generic concepts *can* be substituted by a definite expression: *The horse is a noble beast.* In addition, they can be substituted by an indefinite plural NP (*Horses are noble beasts*). In these cases reference is made to a class as a whole.

In sum, the difference between a definite NP and a referential indefinite NP is that an indefinite NP is used to refer to an entity that is not uniquely identifiable for the reader, whereas using a definite expression implies that there is an identifiable referent. Thus, using an indefinite expression implies that there is no uniquely identifiable referent because a definite expression would have to be used if that would have been the case. The non-identifiability of the referent may have various causes: a) the identity of the referent is known to the writer but has not yet been mentioned and is therefore unknown to the reader (specific referents); b) the referent is some not further specified member of a class, both for the writer and for

[2]Note that the distinction between referential and predicating (or 'attributive') expressions differs from Donnellan's (1966) distinction between referential and attributive definite descriptions. Donnellan meant to distinguish referential descriptions that pick out a particular individual, and attributive descriptions that denote anyone who satisfies the description in question. For instance, *the winner* may be used to pick out the relevant individual after the match, whereas it may be used to designate whoever will satisfy the description before the match (i.e. what is called non-specific reference in 4.1.1). In this section, *referential* is used to indicate the identificational function of a linguistic expression, and *attributive* or *predicational* is used to indicate the function of attributing properties to a referent (either by means of a (nominal) predicate or by means of a co-referential expression (see 4.1.2, 4.3)).

the reader (non-specific referents); or c) the referent is referred to as a prototypical member of its class and its specific identity is not relevant as such (generic referents).

4.1.2 Predicating indefinites

Predicating indefinites, e.g., *He is a doctor*, typically occur in nominal predicate position. They are usually analyzed as being non-referential. Referential predicating indefinites do exist (e.g., *The bank robber is a man called Bill Sikes*), but that issue will not be addressed here. The nominal predicate in *He is a doctor* is a classifying construction in which some property is assigned to the subject, namely, that he is a member of the set of doctors (Keizer, 1992).

As the following examples show, a predicating indefinite does not refer, that is, it does not introduce a referent. Instead, its function is to attribute certain properties to the referent. This predicating function is signaled by the copula *to be*.

(10) Willie B. is *a (*the) TV addict.* [...] (cited in Ushie, 1986)

11) Robert is – sinds enige tijd – *een (*de) man die met duistere blauwe ogen zit te roken en te zwijgen*, Erik kent hem al tientallen jaren. (M. de Moor (1992), Eerst grijs dan wit dan blauw, p. 15)
For some time now, Robert has been *a (*the) man who is sitting there with dark blue eyes, smoking and being silent*; Erik has known him for dozens of years.

The use of the copula is not necessary for an indefinite to be predicating, however. In examples (12) and (13) for instance, the predicating constructions are elliptic. In example (14) the predicating indefinite occurs in an (elliptic) apposition, and in examples (15) and (16) a verb is used that is similar to the verb *to be*, but not the same:

(12) Maar (Jo) Gijsen wordt met hoon en vermaak uitgeluid. *Een (*de) enge, zwart-rigide, gevoelloze man, tevens een Lulletje Wijwater met een inferioriteitscomplex, een man die vochtig spreekt, die spottend ook wel Joke genoemd werd, dommig, en zelfs niet in staat orde te houden als hij aan héél oude wijfjes lesgaf.* (VN, 30-1-1993:7)
(translated where possible)
But (Jo) Gijsen [a Dutch bishop] is given his farewell with scornful laughter. *A (*the) scaring, black-rigid, insensitive man, also a 'Lulletje Wijwater' with an inferiority complex, a man who talks moistly, who has mockingly been called Jody, foolish, not even capable of keeping order when he was teaching very old little women.*

(13) Hun zoon was inmiddels negen. *Een (*de) onbuigzame jongen die zijn geestesziekte tot in zijn vingertoppen, zijn tenen, zijn rare, alle kanten opgroeiende haar had weten door te voeren.* (De Moor, p. 55)
Their son was nine years old now. *A (*the) stubborn boy who had managed to extend his mental illness into his fingertips, his toes, his funny hair that grew in all directions.*

(14) Erik is volkomen van de kaart sinds uitgerekend hij, maandag, dat verschrikkelijke paar moest vinden: zijn jeugdvriend Robert als krankzinnige moordenaar en Magda,

*een (*de) vrouw die hem volgens mij nooit bijster veel heeft gezegd*, in de meest absurde staat waarin een mens maar kan verkeren, dood. (De Moor, p. 272)
Erik has been completely played out since he of all people, on Monday, had to find that horrible couple: his old friend Robert as an insane murderer and Magda, *a (*the) woman who I think never meant very much to him*, in the most absurd condition a human being can be in, dead.

(15) Er zijn ook acteurs en actrices wier aanwezigheid op het doek kennelijk zo volledig is dat bij het publiek vragen naar 'de mens achter de ster' niet meer opkomen. Audry Hepburn was zo iemand, ook al groeide ze ongekend snel uit tot *een (*de) ster*. (VN, 30-1-1993:54)
There are also actors and actresses whose presence on the screen apparently is so perfect that questions about 'the person behind the star' do not occur to the public. Audrey Hepburn was such a person, despite the fact that she became *a (*the) star* unprecedentedly fast.

(16) [over bisschop Gijsen] In al die afscheidsstukjes wordt hij afgeschilderd als *een (*de) impopulaire eenling, niet geliefd bij zijn collega-bisschoppen, niet geliefd bij pastoors en gelovigen*. (VN, 30-1-1993:7)
[about bishop Gijsen] In all those short farewell pieces he is depicted as *an (*the) impopular solitary individual, not much liked by his fellow-bishops, not much liked by priests and believers.*

In none of the above examples the indefinite article can be replaced by a definite article without changing the meaning of the sentence. When replaced the expressions become referential instead of predicating.

4.2 Marked use of indefinites: the referential function

By convention, it is inappropriate to use an indefinite expression to refer to a (specific) referent that has already been identified, because the use of an indefinite implies non-identifiability. Similarly, one cannot use a definite expression when the referent is unknown to the addressee or when the addressee can by no means identify the referent by inference from the situation or context.

These conventions can be violated: When the addressee encounters an expression that is not consistent with the conventional rules reflected in the accessibility hierarchy, he infers by the Cooperation Principle that there must have been a special reason for the speaker or writer not to use the appropriate expression. In this way the conventional definite / indefinite distinction can be exploited to express special meaning.

One example of this kind of exploitation is that modern novels often begin with a (definite) personal pronoun, which is an expression with relatively little information about the referent, in spite of the fact that the referent has not been mentioned:

(17) *Hij* staat elke dag om half zeven op. Een paar minuten voor die tijd opent hij zonder toedoen van enig machinaal geweld zijn ogen en komt tot bezinning. (De Moor, p. 9)
Every day *he* gets up at seven. A few minutes before, he opens his eyes without any mechanical violence and comes to his senses.

By presenting the referent as if it was a given entity, the writer creates the special effect for the reader to have the impression that he plunges right in the middle of an already existing (fictional) discourse world.

Likewise, indefinite expressions are sometimes used in contexts in which a definite expression would have been appropriate for identificational purposes. In the following example, which is a translation from a Dutch newspaper article, the indefinite is used twice to refer to a referent introduced in the first paragraph:

(18) (translation)
SERIES OF RAIDS: THREE YEARS AND 'TBR' DEMANDED
From our correspondent
ARNHEM – Public prosecutor R. Eigeman has demanded a court sentence of three years of unconditional imprisonment against a 31-year-old man from Beuningen. The man is suspected of a series of hold-ups at several banks and a post office in Nijmegen.
On May 18th of this year, *a raider* demanded money from an employee of a post-office in Nijmegen. He threatened her with a gun. This gun went off while the man tried to escape with the loot, which was almost 7700 Dutch guilders. According to the man from Beuningen, this was an accident. 'I didn't fire on purpose'.
Three months later, on August 5th, a branch-office of the Rabobank in Nijmegen was raided. Armed with a Riot-gun *a raider* approached a desk clerk and demanded money. This time the loot was over 7000 guilders. The suspect is said to have seized another 25.000 guilders by raids on other branch-offices.
The public prosecutor argued to the court of justice that it was not the first time the man was on trial for this sort of criminal acts. (De Gelderlander, 6-11-1992)
(*Note*: TBR = detention during Her Majesty's pleasure)

In these cases the intended referent is already known. Nevertheless, an indefinite expression is used. In this way it is implied that the referent was not yet identifiable at the moment that is being described in the text. The indefinite expression has the effect (or contributes to the effect) of creating a new subdomain in which the referent concerned has not been mentioned so far (see for the notion of domains for instance Fauconnier, 1985; Sanders, 1990; Sanders & Redeker, 1993, in press). So in fact, the indefinite expression does what it usually does, namely refer to some not identifiable referent, at least not identifiable in the subdomain. Although in this sense a new referent is introduced, the reader infers that the referent is the referent mentioned in the preceding domain.

Like most pragmatic inferences, this inference can be canceled: 'On August 5 a raider approached a desk clerk, but he (the raider introduced in the subdomain) appeared to be someone else (i.e. not the referent from the preceding domain)'. In example (18) it is actually only in the last sentence that the reader knows for sure that the two raiders introduced by indefinite expressions both are the man from Beuningen. The following example, taken from a Dutch newspaper, shows that such cancellations actually occur:

(19) NEWSPAPER-GIRL STABBED WITHOUT REASON
NIJMEGEN – Yesterday morning at a quarter past seven, a seventeen-year-old newspaper-girl was stabbed by an unknown man without reason. [many details about the incident

are described]. Immediately after the knifing-incident the police formed a special investigation team of nine officers. [...] So far this has not led to any concrete clues.
A short time after the crime the police was informed by some bus-passengers who said that they had seen *a man with a knife* in the bus. However, the description of this man, a 29-year-old citizen of Nijmegen, did not match the description of the perpetrator of the crime. (De Gelderlander, 21-11-1992)

The pragmatically implicated meaning of the re-introducing indefinite is particularly suitable in texts in which events are presented from different perspectives, an observation which has not gone unnoticed. By using an indefinite a perspectival domain can be signaled, in which for instance a character in the text perceives another character for the first time, whereas this other character had already been established in the preceding context but in a different perspectival domain (Hinds, 1977; Ushie, 1986; Sanders, 1990; Sanders & Redeker, 1993, in press). A clear example of a perspectival domain was given in section 2.

It seems possible to generalize the function of the re-introducing indefinites to signaling not only perspectival embedded domains, but also temporal or text structural domains, or domains with a combination of these characteristics. Consider the following examples:

(20) TWEEDE OVERVALLER SPARE RIB COUNTER GEPAKT BIJ INBRAAK
NIJMEGEN – Een 35-jarige Nijmegenaar, die gisternacht op heterdaad is betrapt tijdens een inbraak in een pizzeria aan de Sint Annastraat, blijkt veel meer op zijn kerfstok te hebben.
De man heeft bekend dat hij mededader was van de gewapende overval, die op 12 oktober plaatsvond bij een afhaalcentrum aan de Berg en Dalseweg.
De man werd al enige dagen gezocht door de politie op verdenking van betrokkenheid bij de overval, maar tot dusver zonder resultaat.
Gisternacht kreeg de politie plotseling een melding binnen van een eigenaar van een pizzeria, die *een inbreker* in de kraag had gevat. Het bleek te gaan om de 35-jarige Nijmegenaar die op het opsporingslijstje van de politie figureerde. (De Gelderlander, 3-11-1992)

SECOND RAIDER OF SPARE RIB COUNTER CAUGHT DURING BURGLARY
NIJMEGEN – A 35-year-old citizen of Nijmegen, who was caught in the act during a break-in in a pizzeria on St. Annastraat, appears to have much more on his slate.
The man has confessed that he was an accomplice of the armed raid on a take away center on Berg en Dalseweg, which took place on October 12.
The man had been sought by the police for several days, on suspicion of participation in the raid, but so far without success.
Yesterday night the police suddenly received a call from an owner of a pizzeria, who had caught *a burglar*. It turned out to be the 35-year-old citizen of Nijmegen who was on the tracing list of the police.

(21) GIRL SUBDUES ATTACKER
A brave young woman turned the tables on a robber and beat him with an iron pipe which she had wrested from him, then handed him over to the police in Osaka

Wednesday night.
At about 11:25 p.m. Wednesday, *a man* attacked Miss Mayumi Sanda, 23, of Oyodo-cho, Oyodo-ku, Osaka, on a street in the same ward. (cited in Hinds, 1977)

In (20) the coreferring indefinite expression signals a domain that is delimitated by cues that signal a temporal domain ('yesterday night'), as well as a perspectival domain (in 'the police received a call' *receive* indicates the perspective of the police). In this domain, the raider is reintroduced because at that moment in (story-)time the police did not know that the burglar was the one they had been looking for. Apart from that, it can also be argued that the clause containing the indefinite expresses the perspective of the owner of the pizzeria, who also met the burglar for the first time.

In example (21) the subdomain can be said to be text-structural in nature. The first sentence of this newspaper article is a summary of the whole article. This is typically used in news stories and usually referred to as a 'lead'. After the lead, the story starts again, elaborating upon the lead. In this new text-structural domain the main characters are re-introduced by means of indefinite expressions.

4.3 Marked use of indefinites: the attributive function

Indefinites that occur in nominal predicate position are not referential. Instead, they are used to attribute properties to a referent (see 4.1.2). This attributive function can, however, also be observed for indefinite expressions that refer to an already established referent. Because such a referent should be referred to by a definite expression, this attributive indefinite can be considered as another form of overspecification. Consider the following examples:

(22) In addition to the spiritual suffering of loneliness, of having to leave behind him 'the world which had made him what he was', Hammarskjöld had to endure [..] the plain physical suffering of constant nervous strain and overwork. If, as the reader goes through the entries between 1953 and 1957, he finds himself impatient [..] with their relentless earnestness and not infrequent repetitiousness, let him remember that most of them must have been written by *a (*the) man at the extreme limits of mental and physical exhaustion.* (cited in Ushie, 1986)

(23) In 1974 the American magazine Rolling Stone invited Jan Morris to write a series of travel articles. The fruits of that collaboration between *a (*the) romantic traditionalist Welsh author* and *a (*the) lovely and innovative American paper* appear in Destinations. (cited in Ushie, 1986)

(24) Daarop zette hij [= een spoorwegbeambte] zijn borst uit en bedolf ons onder het gewicht van zijn dienstmededeling. Dit was een moeder-met-kindcoupé. 'Raus!' vatte hij de situatie samen. 'Alle: Raus!' En terwijl de Slowaken, de zigeunerinnen, de baby en ik nog bezig waren het vertrek te ontruimen, kwam de moeder er al aan. *Een (*de) mooie jonge vrouw* haastte zich met nietsziende ogen, haar kind tegen zich aangedrukt, langs ons heen om zich ondanks de beproevingen van de reis aan haar moederschap te wijden. (De Moor, p. 237)

> After that he [= a railway officer] expanded his chest and buried us under the weight of his staff announcement. This was a mother-and-child compartment. 'Raus!' he summarized the situation. 'Alle: Raus!' And while the Slovaks, the gipsies, the baby and I were still clearing the compartment, the mother was already arriving. *A (*the) beautiful young woman* hurried past us without looking, her child pressed against herself, to devote herself to her motherhood despite the harassments of the journey.

The attributive nature of these indefinite expressions becomes evident when they are replaced with expressions without the attributive elements. In (22) for instance, the phrase 'let him remember that most of them must have been written by *a man*' is meaningless in the present context.

When the full lexical expressions are replaced with definite counterparts, which is possible in principle because of the accessibility of the referent, the definite expressions are still syntactically and semantically correct, but they imply a different meaning. Whereas the indefinite expressions are mainly attributive in nature, the definite expressions are mainly referring to the already known referent, thereby implying that the attributive information is also already known and identifiable. When the attributed information is not identifiable or inferable as being a unique property of the referent, it is not even possible to replace the indefinite expression with its definite counterpart, because a referent with those properties has not been mentioned before. This is demonstrated in the following example.

(25) [Miss Moss had been asked by the landlady to go out of her room if she couldn't pay her rent by eight o'clock tonight. Left alone in her room, Miss Moss saw herself in the glass and talked to the person in the glass.]
'You silly thing', scolded Miss Moss. 'Now what's the good of crying: you'll only make your nose red. No, you get dressed and go out and try your luck – that's what you've got to do.' She unhooked her vanity bag from the bedpost, rooted in it, turned it inside out. 'I'll have a nice cup of tea at an ABC to settle me before I go anywhere,' she decided. 'I've got one and thrippence – yes, just one and three.'
Ten minutes later, *a (*the) stout lady in blue serge, with a bunch of artificial 'parmas' at her bosom, a black hat covered with purple pansies, white gloves, boots with white uppers, and a vanity bag containing one and three*, sang in low contralto voice:
'Sweet-heart, remember when days are forlorn
It al-ways is dar-kest before the dawn.'
But then the person in the glass made a face at her, and Miss Moss went out. (K. Mansfield, 'Pictures'; cited in Ushie, 1986)

Ushie (1986) claimed that the indefinite in (25) signals a discontinuity in point of view. She argued that '(an) already identified individual is represented as if he was newly introduced and this is due to the discontinuity in point of view from which he is represented.' But it can hardly be maintained that it is the point of view of the writer that changes: The perspective from which the events are represented is the same before and after the use of the indefinite. Rather, the character in the story has undergone some change and this is indicated by means of the attributive indefinite expression. Similarly, in (26) and (27), which Ushie interpreted as signaling discontinuity in point of view and shift in point of view respectively, there is no

discontinuity or shift from one point of view to another. The already identified referent is not re-introduced as if it were introduced in a different subdomain. Instead, the indefinite expression is used to attribute information to the referent.

(26) Carl Bowler was a carpenter who loved to hunt and fish. [..]
One day, while out in the woods, Carl was nabbed by a local constable. The constable took him to the mayor, who also acted as the chief of police, tax collector, and justice of the peace. The constable claimed Carl had been trespassing, and the mayor agreed. Carl protested, 'Wait a minute! I was on public land the whole time. If this is a trial, I want a lawyer. And I want a jury, too.'
The mayor answered: 'This is a trial. And if you don't quit making such a disturbance, I'll find you in contempt of court. Thirty days in the county jail!'
*A (*the) shocked and furious sportsman* was hauled away. (M. Lipman, 'Point of law'; cited in Ushie, 1986)

(27) [Matthew Cuthbert and his sister Marilla decided to adopt a boy from an orphan asylum because they were old and wanted someone to help them on the farm. But for some mistake, instead of a boy, a girl was sent to them. Matthew couldn't tell her that she must go back to the asylum, seeing how glad she was to come to live with them. Besides he was charmed by Anne's personality. But Marilla was determined to send her back and told her that it was a mistake. Anne was grieved to hear it. Sometime after, Anne went upstairs to bed, [..] later, Marilla and Matthew began discussing the matter.]
'I'm not suffering for company,' said Marilla shortly. 'And I'm not going to keep her.'
'Well now, it's just as you say, of course, Marilla,' said Matthew rising and putting his pipe away. 'I'm going to bed.' To bed went Matthew. And to bed, when she had put the dishes away, went Marilla, frowning most resolutely. And upstairs, in the east gable, *a (*the) lonely, heart-hungry, friendless child* cried herself to sleep. (L.M. Montgomery, 'Anne of green gables'; cited in Ushie, 1986)

The fact that these indefinites are not perspective-signaling or re-introducing (see section 4.2), becomes even more evident when the cancelability of the pragmatic implication that is evoked by a re-introducing indefinite is considered. Whereas canceling is possible for the re-introducing indefinites, it is impossible for the attributive indefinites. In (27) for instance, a continuation like 'a lonely, heart-hungry, friendless child cried herself to sleep, but this child was not Anne,' is completely incoherent. In the case of a re-introducing indefinite the identity of the referent is deliberately left unspecified, whereas this is not at issue in the case of an attributive indefinite: The main function of this indefinite is to attribute information to an already established referent.

For examples (22) and (23) Ushie (1986) claimed that the indefinite expressions reflect a subjective interpretation of the referent on behalf of the writer. This is in accordance with the view that the indefinites attribute information to a referent, because this 'predication' can be done by the writer. However, if the other examples just described are considered to belong to the same category of attributive indefinites, these indefinites can no longer be claimed to reflect a subjective opinion of the writer about the referent. The indefinite in (25), for instance, is attributive, but the writer is not overtly present in the text. Rather,

the attributive indefinites should be considered to be vivid descriptions of a referent by an omniscient narrator or a first person narrator, which creates a literary style. Whether this information reflects a subjective opinion or not, depends on the point of view of the text fragment, being for instance the point of view of an overtly present narrator, an omniscient narrator or a character. The main function of the coreferential indefinites is, however, to convey attributive information about the referent.

To conclude, both types of marked, coreferential indefinites described in the last two sections are violations of the accessibility principle in order to convey additional, non-identificational meaning. The additional meaning is based on the original functions of the indefinite expressions, being introducing a referent into the discourse or predicating something about an already established referent. The marked indefinites described are used to introduce an already established referent into a subdomain of the text, or to attribute new, not-inferable properties to an already established referent.

5 Demonstrative nominal anaphora

5.1 The interpretation perspective

This section presents an interpretation proposal for demonstrative NAs in Dutch discourse, which is based on the functional idea of overspecification as described above. The section is largely based on Maes (1991, ch.3-4) and Maes & Noordman (forthc.).

The starting point of the interpretation of demonstratives as overspecified expressions in discourse is that a demonstrative nominal anaphor contributes in two different ways to the development of discourse. This twofold contribution is guided by the following interpretation heuristics:

- first, a demonstrative NA has to be interpreted as *not being a definite anaphor*
- second, a demonstrative NA has to be interpreted as *not being the other/concurrent demonstrative form.*

These heuristics are considered to guide a two-way interpretation mechanism of demonstrative NPs (defNPs), based on two simple substitution dichotomies:

- definite NA vs. demonstrative NA, abbreviated defNA vs. demNA and
- proximal demonstrative NA vs. distal demonstrative NA, abbreviated *deze*NA vs. *die*NA. *Deze*NA is the abbreviation for a proximal demonstrative NA, which in Dutch is realized with *deze* ('this'/'these') or with *dit* ('this'), *die*NA is the abbreviation for a distal demonstrative NA, in Dutch realized with *die* ('that'/'those') or with *dat* ('that').

Starting from a simple example like (28),

(28) *Bush*$_i$ has decided to raise taxes after all. It is the first time *this president*$_i$ has dared to do so.

the interpretation heuristics are claimed to be as follows:

(i) The functional variation between *this president* and *the president* offers the input for the semantic processing of demNAs (see section 5.3);
(ii) the functional variation between *this president* and *that president* is the information source for the pragmatic processing of demNAs (see section 5.4).

The interpretation heuristics reflect our interpretation perspective, in which demNAs and defNAs on the one hand and *deze*NAs and *die*NAs on the other are regarded as privileged substitution variants. This means that the analysis of a demNA should be guided by the functional difference with the corresponding defNA-version on the one hand, and the corresponding demNA-version on the other.

The two interpretation heuristics have to be considered to complement the basic heuristic of overspecified expressions in discourse described above. They are only activated when discourse conditions overrule the standard identificational function associated with nominal anaphora, i.e. when demNAs refer to highly accessible discourse referents as in (28).

5.2 The analytic basis of the demNA-proposals

The interpretation proposal for demNAs is based on the analysis of a large number of demNAs in a corpus of Dutch written texts. The corpus consists of different types of texts, such as informative and instructive leaflets, excerpts from political journals, popular scientific texts, book reviews, tv reviews, autobiographical columns and personal commentaries. The total number of demNPs in the corpus is 1259.

This corpus in the first place provides evidence for the fact that most of the demNA-occurrences cannot be accounted for simply by identificational considerations: A large proportion of demNAs represents a contextual type as is exemplified in (28), i.e demNAs used in one-topic contexts, referring to focal referents, that are not competing with other referents. For the basicality of this contextual type, the corpus analysis provides four pieces of evidences.

(i) In the majority of the demNA-occurrences in our corpus, the antecedent is in the same sentence (14.7%) or one sentence back (66.8%) and within the same text segment (682 out of 837 anaphoric demNPs, i.e. 81.5%).
(ii) As in (28), the majority of demNAs in the corpus occurs in *post-first-mention position* (63.7%).
(iii) The corpus contains hardly any demNAs that function as selecting devices amongst competing candidates in discourse; i.e. demNAs in discourse do not fit in with the standard deictic function of substantival demonstratives as in (29):

(29) Ik wil niet deze hier, maar die daar.
I do not want this one here, but that one there.

(iv) Most of the demNA-occurrences can be replaced by a defNA variant.

The corpus further makes clear that the difference between *deze*NAs and *die*NAs is not a difference in terms of referential or identificational explicitness:

(i) There is no significant difference between *deze*NAs and *die*NAs in terms of referential distance.
(ii) *deze*NAs and *die*NAs are not used contrastively, as in (29); there is no single sentence in the corpus in which the difference between *deze* and *die* is used to access contrastive competing referents.
(iii) There is no significant difference in the way in which *deze*NAs resp. *die*NAs are associated with focal discourse conditions, such as *post-first-mention position, subject position, first- constituent position* etc.
(iv) The difference between *die*NAs and *deze*NAs in the corpus is not to be interpreted as a difference of deictic force, i.e. low vs. high deixis (Kirsner, 1979). The replication of Kirsner's analysis in our corpus (based on twice as much demNAs from a wider range of text types) does not at all yield a clear difference in terms of deictic force.

These analytic results justify the assumption that identificational overspecifications are the pivotal and default type of demNAs in discourse.

5.3 The definite-demonstrative dichotomy in discourse

5.3.1 The MOD-proposal

The corpus analysis accounts for the fact that overspecification consitutes a pivotal use of demNAs in discourse. Our proposal for the function of anaphoric overspecification in these cases starts from the basic intuition that an overspecified demNA not just accesses a referent, but also modifies it, thus producing a kind of informational development in discourse. The modificational effect of demNAs is expressed in the following interpretation rule and interpretation strategy:

> INTERPRETATION RULE: THE MODIFICATION OF DEMONSTRATIVES (MOD)
>
> The markedness of demonstratives in anaphoric demonstrative NPs effects a *modification in the semantic representation* of the underlying discourse referent by activating a *non-default interpretation* of the NP involved.
>
> INTERPRETATION STRATEGY
>
> The reader relies on the formal-lexical relationship between antecedent and demNA as the initial trigger, which provides him with the following types of semantic modification, each of which implying the processing of a particular type of coherence relation vis à vis the definite-NP-version:
>
> | | | |
> |---|---|---|
> | superordinate demNAs | $\longrightarrow$ | classifying the referent |
> | demNAs with attr. modifiers | $\longrightarrow$ | attributing information to the referent |
> | partially repeated demNAs | $\longrightarrow$ | contextualizing the referent |

Assuming for the sake of simplicity that the definite counterpart of a demNA results in a simple identification of the antecedent, the non-default interpretation induced by the demonstrative version results in an interpretation which implies more than just an identificational relationship between anaphor and antecedent and which thus results in more complex coherence relations underlying the anaphor. The nature of the relation is provided or indicated by the formal relationship between antecedent and anaphor, yielding three prototypical formal demNA-types defining three semantic modification types.

In the analysis in this section we will only shortly mention the most important modification types of demNAs in discourse. In Maes (1991), the distinction between the types is further corroborated in that the contextual factors accounting for the basic modificational demNA-types are described in detail.

5.3.2 Classifying the referent

A first class of demNA-occurrences causes an activation of the class of elements which the referent is a member of. The prototype is constituted by anaphora, which are semantically or pragmatically superordinate to their antecedent, as is clear in (30) and (31):

(30) *Bush*$_i$ heeft besloten om de belastingen dan toch te verhogen. Het is de eerste keer dat {*de / deze*} *president*$_i$ een verhoging aandurft.
Bush has decided to raise taxes after all. It is the first time the / this president has dared to do so.

(31) *De mus*$_i$ komt in onze lage landen in grote getale voor. {*De / deze*} *vogel*$_i$ behoeft geen speciale bescherming.
The sparrow is very common in our low countries. The / this bird doesn't need special protection.

In terms of the MOD-proposal we can describe the function of the demNA in this example as follows. The default interpretation of the defNA (the president; the bird) is the simple anaphoric reading with the antecedent NP. The link between them is structurally as well as semantically or pragmatically simple. First of all an implicit coherence relation of the type *identify* can be attributed to the underlying referent representation, as is expressed in the a-versions. The demonstrative determiner overrules the simple *identify*-reading of the corresponding defNA. The semantic modification of the referent can be expressed by a *classify*-relation contained within the representation underlying the demNA, as is expressed in the b-versions:

(30a) *defNA-representation*
BUSH^n ⟵ PRESIDENT^n
(identify)
(Bush is the president in question)

(31a) *defNA-representation*
SPARROW^n ⟵ BIRD^n
(identify)
(the sparrow is the bird in question)

(30b) *demNA-representation*
BUSH^n ⟵ PRESIDENT^n
(classify)
(Bush is a member of the class of presidents (of the US))

(31b) *demNA-representation*
SPARROW^n ⟵ BIRD^n
(classify)
(the sparrow is a member of the set of birds)

The adequacy of this class interpretation is affected by a combination of the following conditions: (i) the naturalness of the antecedent as being a member of the anaphor-class and (ii) the plausibility with which the class-interpretation can be integrated within the relational and propositional network of discourse. Both adequacy conditions are met in the examples (30) and (31). The naturalness of the class-membership of Bush and sparrow prohibits the demNA of being unexpected or new. Moreover, the *classify*-relation induces interpretation possibilities which fit in with the coherence structure of discourse; in this case it induces an implicit contrastive interpretation, such as *other presidents might have increased taxes (more than once)* or *I'm not talking about other presidents* in (30) and *other birds might have needed protection* in (31).

5.3.3 Attributing information to the referent

The second prototype is constituted by demNAs containing extra attributive information vis à vis the antecedent, predominantly realized by attributive pre- or postmodifiers. The difference with the *classify*-type lies in the type of information with which the referent is modified. In the *classify*-type the referent is modified by its most privileged attribute, i.e. the natural class the referent belongs to, which is an invariable and basic characteristic of the referent. In the attribute-type, the referent is modified with all of the attributive information about the referent that is contextually relevant, as can be seen in the following examples:

(32) When *the washing machine*$_i$ first came on the market, {*this / ?the*} [our addition] *fantastic and truly revolutionary aid for the housewife*$_i$ was available only to very few. (Bosch 1983:150)

(33) De Londense Times maakt vanochtend stukken bekend uit *het overheidsrapport over de salmonella-besmetting van de Britse eieren*$_i$. {*Het / dit*} *vertrouwelijke rapport*$_i$ ...
This morning, the London Times publishes extracts from the government report on the salmonella infection in British eggs. {The / This} confidential report ...

There is a lot of evidence for an *attribute*-interpretation of these demNAs – as opposed to their definite variants (see Maes, 1991:104-107). According to our MOD-proposal, the difference between the definite and demonstrative versions can be expressed as a difference between an *identify*-relation for defNAs and an *attribute*-relation for demNAs. The demonstrative determiner in these cases is a signal to interpret the lexical content of the NP not in the referential default, but in the attributing non-default. For (33), for example, (33a) underlies the defNA-version and (33b) the demNA-version:

(33a) *defNA-representation*
REPORT^n ⟵ CONFIDENTIAL REPORT^n
(identify)
(the confidential report is the report in question.)

(33b) *demNA-representation*
REPORT^n ⟵ CONFIDENTIAL REPORT^n
(attribute)
(the report is confidential.)

The prediction of the contextual conditions under which an attributive anaphoric NP is by default identifying, and under what conditions an attributing non-default interpretation is or can be activated by means of a demNA, is guided by two principles outlined in Maes (1990). There, the adequacy of definite anaphora lexically differing from their antecedent is subjected to two pragmatic conditions, the *condition of centrality* and the *condition of relevance* respectively. The condition of centrality prohibits defNAs of containing central information, i.e. information which triggers coherence relations that are crucial in the realization of the central text intention. The condition of relevance enforces these defNAs to contain relevant information, i.e. information which triggers coherence relations that can be related in a relevant way to the basic text intention. That means that a reader must be able to ground a defNA on an implicit coherence relation or proposition which he conceives to be a relevant addition in the given context. In the following example, both conditions are illustrated. The a-version can be considered to be the adequate realization of the text intention described in (34). The defNA in the b-example infringes upon the condition of centrality, in (34c) it violates the condition of relevance.

(34) <*At a press conference Bill Clinton introduces his new press chief Donald Baker who, as is well-known, happens to be the grandson of Jesse James. The fact that he was going to be the new press chief, however, had been known for some days throughout the whole world and was certainly known to the press people. The press conference is to be looked upon as a kind of 'inauguration' of the new press chief*>

(34a) This is *Donald Baker*$_i$. *He*$_i$ is the new press chief. From now on *he* will speak to you every Wednesday afternoon. *The new press chief*$_i$ ....

(34b) This is *Donald Baker*$_i$. From now on *?the new press chief*$_i$ will speak to you every Wednesday afternoon.

(34c) This is *Donald Baker*$_i$. *?The grandson of Jesse James*$_i$ is the new press chief.

In view of the defNA-demNA variation these conditions can be amended as follows: By default, a defNA has a simple *identify*-interpretation when its lexical content fulfills the two conditions. DemNA-replacement yields a relaxation of one of these conditions. That means that the demonstrative causes one of the following two interpretative effects. In the first place it can assign a more central place in the text to the NP-information concerned because of the relaxation of the centrality condition. In these cases, the demNA emphasizes or increases the prominence of the information, as would be the case by introducing a demonstrative in (34b). In the second place, it can smuggle less relevant information into the text, because of the relaxation of the relevance condition (see also Bosch, 1983:150), an effect which can be illustrated (34c) in which a demonstrative would not be perfect but at any rate better than the definite version.

5.3.4 Contextualising the referent

The third modificational type of demNAs is constituted by the prototype of anaphora literally or partially repeating their antecedent, as in the following example:

(35) In de grote steden neemt *de luchtvervuiling*$_i$ toe. {*Die / deze / de*} *(lucht)vervuiling*$_i$ is het gevolg van wanbeleid.
In the big cities, air pollution is increasing. {This / that / the} (air) pollution is a result of bad policy.

Contextualising demNAs activate a non-default interpretation of the anaphoric NP: The demonstrative is an explicit signal or invitation to contextually accommodate the referent, i.e. the reader does not have to assign an interpretation based on the intrinsic semantic meaning of the following NP, but he has to activate a non-default-interpretation of that NP provided by the contextual information in the antecedent sentence.

This non-default interpretation causes a contextualising modification of the referent: The demonstrative explicitly induces the attachment of parts of the representation of the antecedent sentence 'under' the referent representation in the anaphor sentence, as being information modifying the referent-representation. The modifying of the demNA in (35) is given below:

(25a) *defNA-representation*
AIR POLLUTION^n ⟵ AIR POLLUTION^n
(identify)
(the air pollution is the air pollution in question)

(35b) *demNA-representation*
AIR POLLUTION^n ⟵ AIR POLLUTION IN THE BIG CITIES^n
(contextualise)
(the air pollution is the air pollution in the big cities just mentioned)

The representational information *in the big cities* in the antecedent sentence is explicitly linked to the referent *air pollution* in the anaphor sentence, thus modifying the representation of the referent. Congruent with the other demNA-types, this modification is looked upon as an implicit coherence relation within the representation underlying the demNA.

The basic empirical argument for the contextual interpretation of these demNAs is that their acceptability crucially depends on a difference between the 'standard' meaning of the lexical item in the demNA, as it can be derived from the mental lexicon, and the contextual interpretation of the anaphoric NP.

Most of the contextualising demNAs serve the installation of a referent as topic in the discourse. Yet, contextualisation and referent installation are distinct phenomena, which do not necessarily coincide. That is clear in the following examples:

(36) *PSV*$_i$ is de laatste weken in zeer slechte doen. {*PSV*$_i$ / *Dit PSV*$_i$ *(i.e., het PSV-in-slechte-doen*$_i$*)*} is voor K.V.Mechelen dan ook geen partij.
PSV has been playing very badly these last few weeks. As a result {PSV / this PSV (i.e., the PSV playing badly)} is no match for K.V.Mechelen.

(37) *Ronald Reagan*$_i$ heeft duidelijk last van geheugenstoornis. {*Reagan*$_i$ / *Deze Reagan*$_i$ *(i.e., de Reagan-met-geheugenstoornis*$_i$*)*} zou geen belangrijke beslissingen meer mogen nemen.
Ronald Reagan is clearly suffering from memory disturbance. Reagan / this Reagan (i.e., the 'Reagan suffering from memory disturbance)' should no longer be allowed to take important decisions.

(31) *De VVD*$_i$ is vorige week zwaar uit de bocht gegaan. Als lid vraag ik me op zo'n moment af of ik wel goed zit bij *die VVD*$_i$.
Last week, the VVD put their foot in it. As a member of this party I ask myself at moments like that whether that VVD really is my party.

In all these demNA-cases the contextualising MOD-interpretation is forced: The only reading which is applicable here is the one in which the demNA refers to the referent as it is presented in the preceding context, and not in its default interpretation. In the examples, the anaphor is a repetition of a proper noun, i.e., a referential type, which guarantees referential autonomy and unique identification, and which does not allow any demonstrative determiners (and no definite ones either). Still the demonstrative version is acceptable, at least in its contextualising interpretation. Moreover, the contextualising effect is not dependent on referent installation. The examples can easily be preceded by many sentences in which NPs access the same referent, without the contextualising effect disappearing.

5.4 The dezeNA – dieNA variation in discourse

In order to capture the nature of the inferences associated with the use of *this* or *that* in demNAs, we start from the basic intuition connected with the use of each demonstrative determiner that the accessed referent has to be *present in the discourse*, i.e., that it is related to the context or situation of discourse. In fact, this intuition hinges upon a minimal interpretation of the deictic quality associated with each demonstrative determiner, being that

deixis not necessarily implies *contrast*, *monstration* or *pointing* but only the presupposition of being present (see also Kleiber, 1983).

We will implement the presence intuition in the following way: The use of a demonstrative simply implies the association of the underlying referent with a particular referential domain which is present in the text. The *presence or P-Presupposition (PP)* of demonstratives is defined as follows:

> *Presence-Presupposition of Demonstratives*
> A demonstrative NP expresses the association of the underlying referent with a referential domain evoked in the discourse.

In order to be able to handle the different functions and uses of demonstrative NPs, and in line with the dichotomic nature of Dutch demonstratives (*deze/dit* vs. *die/dat*), we will distinguish two types of referential domains with which referents can be associated. Firstly, in each written text a *deictic reference domain* [DRD] can be discerned. This simply means that each text is embedded in its speech situation. DRD is constituted by the coordination points Bühler (1982:11) distinguishes in his *deictic field of human language*, i.e., *here*, *now* and *I*. The adaptation of these DRD-coordinates for written language yields four deictic reference points, i.e., *place*, *time*, *writer* and *discourse/text*. Each reference point can be made explicit in the discourse by means of referential-deictic expressions. Secondly, in texts referential domains can be established the coordination points of which are clearly different from DRD. The coordination points of these other referential domains [ORD] can be defined as *place*, *time* and *reader*.

The presence-presupposition, combined with the postulation of the two types of referential domains, produces the following characterization of the *deze*NA-*die*NA variation:

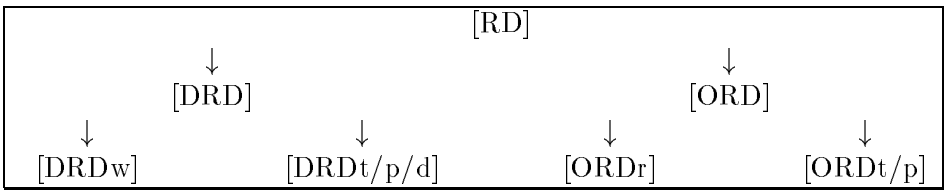
> *Deze-Presence-Presupposition*
> A *deze*NP expresses the association of the underlying referent with one or more coordination point(s) of the deictic reference domain (DRD) in the discourse involved: The underlying referent is assigned the pragmatic value [DRD].

> *Die-Presence-Presupposition*
> A *die*NP expresses the association of the underlying referent with one or more coordination point(s) of an other referential domain (ORD) in the discourse involved: The underlying referent is assigned the pragmatic value [ORD].

Given the distinct coordination points in DRD and ORD, mentioned above, the pragmatic values [DRD] and [ORD] are subdivided into distinct subvalues, which can be dichotomised as follows: On the one hand the association with *actor-coordinates* in the referential domains, i.e., writer and reader respectively, resulting in the pragmatic values [DRDw] and [ORDr]; on the other hand the association with other coordinates, i.e., place, time and discourse, resulting in the pragmatic values [DRDp], [DRDt], [DRDd], [ORDp], [ORDt]. As such, the different types of coordination points in a referential domain are a fruitful starting point in coming to grips with the problem of the relationship between 'literal', spatio-temporal uses of demonstratives and 'relational', actor-oriented uses of demonstratives.

In summary, the subdivision of the pragmatic values DRD and ORD on the basis of writer/reader and other coordination points results in the following figure:

(39) *Pragmatic referential-domain values*



The different values are illustrated in the following examples:

(40a) In de grote steden neemt *de luchtvervuiling*$_i$ toe. *Die luchtvervuiling (i.e., de luchtvervuiling die in de grote steden werkelijk toeslaat)*$_i$ [ORDp] is het gevolg van wanbeleid.
In the big cities air pollution is increasing. That air pollution (i.e., the air pollution which really hits the big cities) is a result of bad policy.

(40b) In de grote steden neemt *de luchtvervuiling*$_i$ toe. *Deze luchtvervuiling (i.e., de luchtvervuiling in de grote steden waarover deze tekst handelt)*$_i$ [DRDd]is het gevolg van wanbeleid.
In the big cities air pollution is increasing. That air pollution (i.e., the air pollution in the big cities which this text deals with) is a result of bad policy.

(41) *Deze sprankelende wereld van passie* [DRDw], zoals ik dat zou willen noemen, is eigen aan de schrijver.
This sparkling world of passion, as I would like to call it, is peculiar to the writer.

(42) *Die spanning van die twee werelden* [ORDr], begrijp je. (3:4155)
That tension of those two worlds, you understand.

As these examples demonstrate, the use of a demonstrative gives rise to pragmatic inferences in terms of the spatio-temporal interpretation of the referent or in terms of the relationship between referent, reader and writer. These inferences can be extended to pragmatic notions such as the appeal to the reader in resolving and conceptualizing the referent (such as in 42) or the writer's perspective in presenting the referent (such as in 41). These extensions are of course not caused only by the demonstrative choice. In the examples above the explicit mention of writer and reader, *begrijp je* 'you understand' and *zoals ik dat zou willen noemen* 'as I would like to call it' help the reader to make such inferences.

Apart from such contextual signals there are also more global discourse clues which cause such relational inferences. In an expository text, for example, the relationship between writer and reader is different from a conversational or narrative text. Consequently, it is more accurate to say that the use of a demonstrative, with its pragmatic value, is *congruent with* writer/reader-inferences resulting from global discourse conditions, than to say that these inferences are *inherent in* the semantics or pragmatics of demonstratives.

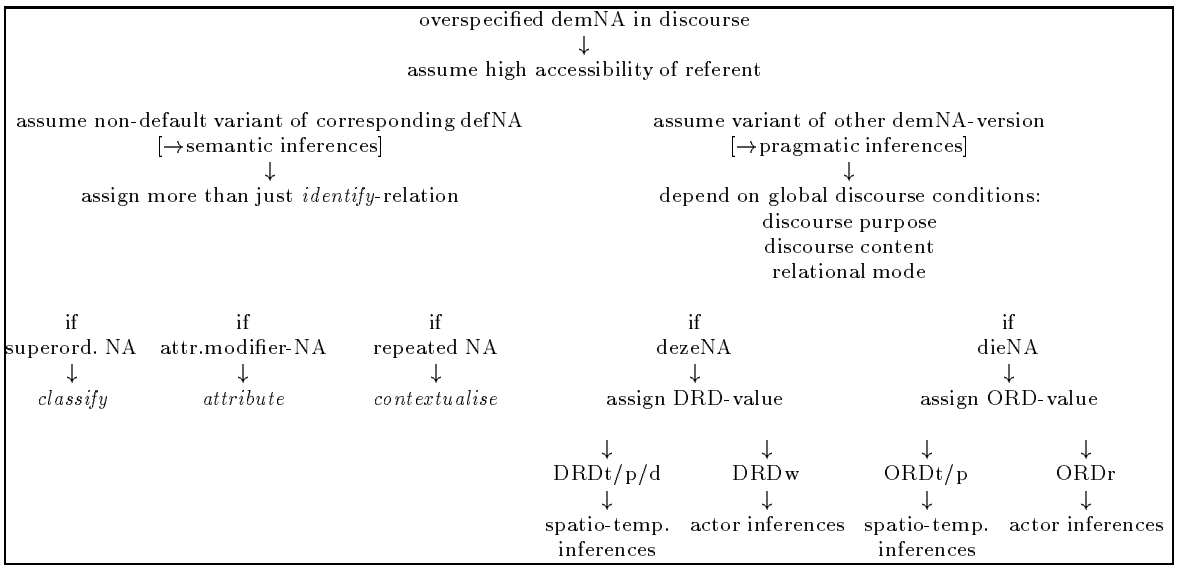
The global discourse conditions which seem to affect the *die/deze*-choice are first of all the *relational mode* between discourse participants, the *purpose* underlying discourse and the

content. These global conditions yield dichotomies such as formal – informal, conversational/narrative – expository. The relevance of these global conditions is reflected in the fact that the *deze/die*-ratio differs considerably over the different text types in our corpus. Further analysis of these discourse conditions and of the applicability of the pragmatic values in the corpus can be found in Maes (1991:189-202).

5.5 A two-way interpretation process for demNAs

The analysis given in the previous sections can be presented as a product oriented interpretation proposal for demNAs, which entails two separate components, based upon different contextual cues, and theoretically located on different levels of language description, i.e. semantics vs. pragmatics.

(43) *The processing of demNAs*



6 Conclusion

In this paper, we started from the standard view on discourse anaphora, in which the form, i.e., the degree of explicitness of referential expressions is inversely proportionate to the degree of accessibility of the underlying referent. We provided evidence for an amendment to this view, which is based on the use of so-called overspecified referential expressions. Overspecified expressions are expressions that contain more coding material than is necessary for enabling the addressee to unambiguously identify the intended referent. We demonstrated that all types of nominal referential expressions can be used as overspecifications: full definite NAs, indefinite NAs, and demonstrative NAs.

Full definite NAs that violate the accessibility hierarchy can be used to mark a theme shift; indefinite NAs may be used to signal a subdomain or to attribute new information to a referent, and demNAs may be used to establish implicit coherence relations, being classification,

attribution, and contextualization.

The implications of these text-analytic observations for the processing of overspecified NAs are now being investigated (Dandelion task 1.4.2). The interpretation and comprehension of demonstrative NAs that are assumed to evoke a class-membership property of the referent are tested in a number of experiments. As for the indefinite NAs, their function of signaling a subdomain of a character in the text will be investigated, as well as their attributive function.

References

Anderson, A., Garrod, S.C., & Sanford, A.J. (1983). The accessibility of pronominal antecedents as a function of episode shifts in narrative text. *Quarterly Journal of Experimental Psychology, 35*, 427-440.

Ariel, M. (1988). Referring and accessibility. *Journal of Linguistics, 24*, 65-87.

Ariel, M. (1990). *Accessing Noun-Phrase Antecedents*. London/New York: Routledge.

Bates, E., Kintsch, W.A., Fletcher, C.R., & Giuliani, V. (1980). Recognition memory for surface forms in dialogue: Explicit vs. anaphoric reference. In J. Kreiman & A.E. Ojeda (Eds.), *Papers from the parasession on pronouns and anaphora* (41-48). Chicago: CLS.

Bosch, P. (1983). *Agreement and anaphora. A study of the role of pronouns in syntax and discourse*. London: Academic Press.

Brown, C. (1983). Topic continuity in written English narrative. In T. Givón (Ed.), *Topic continuity in discourse: A quantitative cross-language study* (313-341). Amsterdam: Benjamins.

Bühler, K. (1982). The deictic field of language and deictic words. In R.J. Jarvella & W. Klein (Eds.), *Speech, place, and action. Studies in deixis and related topics* (9-30). Chichester: Wiley.

Chafe, W.L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C.N. Li (Ed.), *Subject and topic*. New York: Academic Press.

Clancy, P. (1980). Referential choice in English and Japanese narrative discourse. In W. Chafe (Ed.), *The pear stories: cognitive, cultural and linguistic aspects of narrative production* (127-202). Norwood, NJ: Ablex Publishing Cie.

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

Fauconnier, G. (1985). *Mental spaces: Aspects of meaning construction in natural language*. Cambridge/London: MIT Press.

Fletcher, C.R. (1984). Markedness and topic continuity in discourse processing. *Journal of Verbal Learning and Verbal Behavior, 22*, 121-132.

Fox, B.A. (1984). *Discourse structure and anaphora in written and conversational English*. (Ph.D. thesis). University of California, Los Angeles.

Francik, E.P. (1985). *Referential choice and focus of attention in narratives*. (Ph.D. thesis). Stanford University.

Frederiksen, J.R. (1981). Understanding anaphora: rules used by readers in assigning pronominal referents. *Discourse Processes, 4*, 323-347.

Garrod, S.C., & Sanford, A.J. (1982). The mental representation of discourse in a focussed memory system: Implications for the interpretation of anaphoric noun phrases. *Journal of Semantics, 1*, 21-41.

Givón, T. (1983). Topic continuity in discourse: an introduction. In T. Givón (Ed.), *Topic continuity in discourse: a quantitative language study* (1-41). Amsterdam/Philadelphia: Benjamins.

Givón, T. (1987). Beyond foreground and background. In R.S. Tomlin (Ed.), *Coherence and grounding in discourse: Typological studies in language, Vol. 11*, (175-188). Amsterdam: Benjamins.

Givón, T. (1989). *Mind, code and context: essays in pragmatics.* Hillsdale, NJ: Erlbaum.

Givón, T. (1990). *Syntax, a functional-typological introduction, Vol. 2.* Amsterdam: Benjamins.

Grice, H.P. (1975). Logic and conversation. In P. Cole, & J.L. Morgan (Eds.), *Speech acts.* (Syntax and Semantics, 3, 41-58). New York: Academic Press.

Gundel, J., Hedberg, N., & Zacharski, R. (1989). Givenness, implicature and demonstrative expressions in English discourse. In B. Music, R. Graczyk, & C. Wiltshire (Eds.), *Papers from the 25th Annual Regional Meeting of the Chicago Linguistic Society. Part two: parasession on language in context* (89-103).

Hawkins, J.A. (1978). *Definiteness and indefiniteness. A study in reference and grammaticality prediction.* London: Croom Helm.

Hawkins, J.A. (1991). On (in)definite articles: implicatures and (un)grammaticality prediction. *Journal of Linguistics, 27*, 405-442.

Hinds, J. (1977). Paragraph structure and pronominalization. *Papers in linguistics, 10*, 77-99.

Hustinx, L.G.M.M. (in prep.). [*Textual coherence, use and comprehension of theme-shift markers*]. Doctoral dissertation, University of Nijmegen.

Karmiloff-Smith, A. (1980). Psychological processes underlying pronominalization and non-pronominalization in children's connected discourse. In J. Kreiman & A.E. Ojeda (Eds.), *Papers from the parasession on pronouns and anaphora* (231-250). Chicago: CLS.

Karmiloff-Smith, A. (1985). Language and cognitive processes from a developmental perspective. *Language and Cognitive Processes, 1*, 61-85.

Keizer, M.E. (1992). *Reference, predication and (in)definiteness in Functional Grammar: A functional approach to English copular sentences.* Unpublished doctoral dissertation, Vrije Universiteit, Amsterdam.

Kirsner, R.S. (1979). Deixis in discourse: an exploratory quantitative study of the modern Dutch demonstrative adjectives. In T. Givón (Ed.), *Discourse and Syntax.* (Syntax and Semantics, 12, 355-375). New York: Academic Press.

Kleiber, G. (1983). Les dimonstratifs (di)montrent-ils? Sur le sense rifirentiel des adjectifs et pronoms dimonstratifs. *Le français moderne, 51*, 99-117.

Lakoff, G. (1976). Pronouns and reference. In J.D. McCawley (Ed.), *Notes from the linguistic*

underground. (Syntax and semantics, 7, 275-335). New York: Academic Press.

Levinson, S.C. (1983). *Pragmatics.* Cambridge: Cambridge University Press.

Levinson, S.C. (1987a). Pragmatics and the grammar of anaphora: a partial pragmatic reduction of binding and control phenomena. *Journal of Linguistics, 23*, 379-434.

Levinson, S.C. (1987b). Minimization and conversational inference. In J. Verschueren & M. Bertucelli-Papi (Eds.), *The pragmatic perspective* (455-472). Amsterdam: Benjamins

Maes, A.A. (1990). The interpretation and representation of coreferential lexical NPs in expository texts. *Journal of Semantics, 7*, 143-174.

Maes, A.A. (1991). *Nominal anaphors and the coherence of discourse.* Unpublished doctoral dissertation, Tilburg University.

Maes, A. A., & L.G.M. Noordman (forthc.) Demonstrative nominal anaphors: a case of non-identificational markedness.

Marslen-Wilson, W., Levy, E., & Tyler, L. (1982). Producing interpretable discourse: the establishment and maintenance of reference. In R.J. Jarvella & W. Klein (Eds.), *Speech, place and action. Studies in deixis and related topics* (339-378). Chichester: Wiley.

Murphy, G.L. (1984). Establishing and accessing referents in discourse. *Memory and Cognition, 12*, 489-497.

Prince, E.F. (1981). Towards a taxonomy of given-new information. In P. Cole (Ed.), *Radical pragmatics* (223-255). New York: Academic Press.

Sanders, J. (1990). Expliciet of niet? Referentie-bepalende factoren bij personen in nieuwsberichten. *ttt, 9*, 159-180.

Sanders, J., & Redeker, G. (1993). Linguistic perspective in short news stories. *Poetics, 22*, 69-87.

Sanders, J., & Redeker, G. (to appear). Perspective and the representation of speech and thought in narrative texts. In G. Fauconnier & E. Sweetser (Eds.), *Mental spaces, grammar, and discourse.* Chicago: Chicago University Press.

Sanford, A.J., & Garrod, S.C. (1981). *Understanding written language. Explorations of comprehension beyond the sentence.* Chichester: Wiley.

Sanford, A.J., Moar, K., & Garrod, S.C. (1988). Proper names as controllers of discourse focus. *Language and Speech, 31*, 43-56.

Stark, H.A. (1988). What do paragraph markings do? *Discourse Processes, 11*, 275-303.

Ushie, Y. (1986). 'Corepresentation' – A textual function of the indefinite expression. *Text, 6*, 427-446.

Vonk, W., Hustinx, L.G.M.M., & Simons, W.H.G. (1992). The use of referential expressions in structuring discourse. *Language and Cognitive Processes, 7*, 301-333.