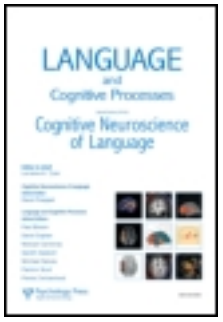
This article was downloaded by: [Max Planck Institut Fur Psycholinguistik]
On: 20 March 2013, At: 04:36
Publisher: Routledge
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK




Language and Cognitive Processes
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/plcp20

# The use of referential expressions in structuring discourse

Wietske Vonk [a], Lettica G.M.M. Hustinx [b] & Wim H.G. Simons [b]
[a] Department of Linguistics, Max Planck Institute for Psycholinguistics, Nijmegen and University of Nijmegen, Nijmegen, The Netherlands
[b] Department of Linguishes, University of Nijmegen, Nijmegen, The Netherlands
Version of record first published: 13 Dec 2007.

To cite this article: Wietske Vonk , Lettica G.M.M. Hustinx & Wim H.G. Simons (1992): The use of referential expressions in structuring discourse, Language and Cognitive Processes, 7:3-4, 301-333

To link to this article: http://dx.doi.org/10.1080/01690969208409389

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

LANGUAGE AND COGNITIVE PROCESSES, 1992, 7 (3/4) 301–333

# The Use of Referential Expressions in Structuring Discourse

Wietske Vonk

*Max Planck Institute for Psycholinguistics, Nijmegen and Department of Linguistics, University of Nijmegen, Nijmegen, The Netherlands*

Lettica G.M.M. Hustinx and Wim H.G. Simons

*Department of Linguishes, University of Nijmegen, Nijmegen, The Netherlands*

Referential expressions that refer to entities that occur in a text differ in lexical specificity. It is claimed that if these anaphoric expressions are more specific than necessary for their identificational function, they not only relate the current information to the intended referent, but also contribute to the expression of the thematic structure of the discourse and to the comprehension of the thematic structure. In two controlled production experiments, it is demonstrated that thematic shifts are produced when one has to make use of such an overspecified expression, and that overspecified referential expressions are produced when one has to formulate a thematic shift. In two comprehension experiments, using a probe recognition technique, it is shown that an overspecified referential expression decreases the availability of information contained in a sentence that precedes the overspecification. This finding is interpreted in terms of the thematic structuring function of referential expressions in the understanding of discourse.

## INTRODUCTION

Understanding text requires interpreting the incoming information and integrating it with the previous discourse. Not only has a syntactic and semantic analysis of the incoming sentence to be performed, but the information also needs to be related to the previous discourse. One of the ways in which sentences are connected is by reference to entities earlier in the text. Referential devices serve to map the information in the current

Requests for reprints should be addressed to W. Vonk, Max Planck Institute for Psycholinguistics, Wundtlaan 1, NL-6525 XD Nijmegen, The Netherlands.

The authors are grateful to Alan Garnham, Leo Noordman, Jane Oakhill and two anonymous reviewers who made valuable comments on an earlier draft of the manuscript.

© 1992 Lawrence Erlbaum Associates Ltd and V.S.P. Publications

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

sentence onto antecedents in the discourse representation. They identify the referent to which the current predicated information has to be attached.

In general, languages have several linguistic devices to refer to entities that occur in the discourse, e.g. zero anaphors, pronouns, names and definite noun phrases. These expressions differ in lexical specificity and, accordingly, in identificational explicitness.

The use of a particular device is generally assumed to depend on the effectiveness of the device to fulfil its identificational function. This will in turn depend on the degree of accessibility of the intended referent when the referential expression is encountered. For example, if there are two candidate antecedents in a particular text, both of which can be referred to by the same pronoun, a more specific referential expression, for example a noun anaphor, is needed to identify the antecedent. Another factor that affects the accessibility is the distance between antecedent and anaphor: the greater the distance, the greater the chance that a noun anaphor instead of a pronoun will be used. It may also be the case that an episode boundary decreases the accessibility of an antecedent, so that a more specific form is more appropriate for identification.

On the basis of the identificational function of referring expressions, one may assume that the less accessible a referent is in the context, the greater the lexical specificity of the referring expression has to be. This agrees with Givón's (1983a) scale of topic accessibility. Givón orders referential expressions according to accessibility of the topic (protagonist) as follows: zero anaphor, unstressed pronoun, stressed pronoun, definite noun and referential indefinite noun. This measures "how accessible the topic is to the hearer, given distance from prior mention in discourse, degree of referential interference from other referents, amount of semantic redundancy available in the proposition, and amount of thematic redundancy available in the discourse" (Givón, 1987, p. 178).

In general, there appears to be a tendency to use the least specific device as long as the intended referent is sufficiently available to be recovered. Frequently, a relatively unspecified device such as a pronoun is sufficient to recover the discourse entity and to add the predicated information to the entity referred to (Brown, 1983; Givón, 1983a). In the terminology of Chafe (1987): if an entity is already active, the concepts expressing that entity are verbalised in an attenuated manner, for instance by pronominalising it.

On the basis of the accessibility of the referents in the discourse, one can predict what kinds of device will be used in the text. Indeed, it is clear from the analysis of texts that an inverse relation can be found between the accessibility of the referent and the specificity of the description (Ariel, 1988; Brown, 1983). However, it is also clear that there are many exceptions to this. Nevertheless, at particular places in texts with only one

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

protagonist who is kept in focus and who is often referred to – thus enabling an unambiguous identification by a pronoun – noun anaphors are often used (Hinds, 1977). Clancy (1980) observed the use of noun phrases to indicate episode boundaries in cases where less explicit referential devices such as pronouns could have been used, and according to Fox (1984) nominal anaphors indicate a transition to a new narrative unit in the discourse. The occurrence of such noun anaphors cannot be explained by accessibility. Such referential expressions are overspecified with respect to the identificational function. But nevertheless they are not used randomly. The difference in form corresponds to a difference in function (Bolinger, 1977), but not at the referential level. They function with respect to the discourse structure.

That different forms of referential device have a discourse structuring function has been observed, for instance, by Linde (1979), in evoked apartment descriptions, for the use of the pronoun *it* and the more specified deictic expression *that*, *it* being used in references to a room "inside the focus of attention", and *that* preferably in references across so-called discourse nodes. Karmiloff-Smith (1985) observed in studies of child language that children prefer to use pronouns to refer to main protagonists, and definite noun phrases to refer to subsidiary protagonists. An indication that overspecifications contribute to the structure of discourse was obtained experimentally by Marslen-Wilson, Levy and Tyler (1982) and by Fletcher (1984). The choice of the referring expressions appeared informative with respect to the thematic structure of the discourse.

What is investigated in the present study is not the fact that an episode boundary makes an antecedent less accessible (cf. Anderson, Garrod, & Sanford, 1983). Rather, the claim is that a referring expression that is more specific than is needed for identification of the antecedent itself indicates a boundary.

The following example illustrates that an overspecification can be used at a theme shift (5) to make the text less awkward.

1. Sally Jones got up early this morning.
2. She wanted to clean the house.
3. Her parents were coming to visit her.
4. She was looking forward to seeing them.
5. She weighs 80 kilograms.
6. She had to lose weight on her doctor's advice.
7. So she planned to cook a nice but sober meal.

. . .
. . .

In our opinion, sentence (5) is awkward not just because the content does not fit the discourse as a whole: in (7) the connection becomes clear. However, the sentence cannot easily be connected to the preceding sent-

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

ences. Now note that, although the pronoun *she* in (5) does not cause any identificational problem, intuitively the more specific *Sally* makes the sentence sound better. Clearly (5) is about a different theme than the parents' visit and the preparations for it, and it seems necessary to indicate this. Note also that replacing the pronoun by the name in other sentences is also possible, especially in (4) and (7), but it does not make the sentence sound better. Sentence (4) is just slightly changed in its perspective, as if one may expect a continuation on the theme of looking forward. It seems that when a device is used that is more specific than is necessary for the recovery of the intended entity, it also has a discourse structuring function. It marks the beginning of a new theme concerning the same discourse referent.

This paper aims to demonstrate that language users exploit overspecification both in production and in comprehension of text. Two kinds of experiment will be discussed. The first set of experiments are controlled discourse production experiments. They aim to demonstrate that the use of overspecifications depends on theme shifts. The second set of experiments are comprehension studies. They test the claim that the differential use of anaphoric references not only correlates with thematic structure, but that overspecified anaphoric expressions contribute to building the mental representation of a text.

## EXPERIMENT 1: TEXT FRAGMENT COMPLETION

The aim of the experiment was to investigate the relation between the use of the different referring expressions and theme continuity in the production of text. This relation was experimentally investigated from two angles: Given the use of a particular referring expression, is the theme thought to be continuous or discontinuous, and given continuity or discontinuity of the theme, what kind of referring expression is used?

A controlled production task was used. The subjects were presented with two written sentences that introduced a protagonist of a (potential) story. There was only one protagonist in these sentences. Therefore, after these sentences, a pronoun would unambiguously refer to the protagonist. The subjects had to produce additional sentences to continue the story. At some point in each story, they were presented with a stimulus word (henceforth called the feeder word) that had to be used in the next sentence they were to produce (target sentence). In the conditions in which the controlling effect of the referring expression on theme continuity was investigated (anaphoric feeder conditions), this word was a pronoun or a noun, namely a proper name or a role description (e.g. *nurse*). In the conditions in which the controlling effect of theme continuity on the selection of a referring expression was investigated (thematic feeder condi-

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

tions), the feeder word was a noun, adjective or verb, the meaning of which was related (thematic continuation feeder) or unrelated (thematic shift feeder) to the theme described in the introductory sentences.

It was hypothesised that if the anaphoric feeder words were overspecified referring expressions (names or roles), this would lead to a greater number of theme shifts in the target sentence than pronouns would, while pronouns would give rise to the production of more theme continuations. It was further hypothesised that the thematic shift feeder words would more often result in the use of overspecified referring expressions, whereas the thematic continuation feeder words would produce more frequent use of pronouns.

## Method

*Subjects.* A total of 20 students of the N.L.O. Interstudie of Nijmegen participated in the experiment. They were all native speakers of Dutch, varied in age from 18 to 25 years, and were not paid for their participation.

*Materials.* The materials consisted of 16 text fragments, each containing two sentences. An example of the materials used is shown in Table 1. The first sentence introduced a character, indicated by a proper name (four male and four female first names) or a role description (four male and four female roles), in a simple action or everyday situation described in the present or past tense. The second sentence expanded this action or situation, and referred to the protagonist with a pronoun. Four kinds of feeder word were coupled to each text: in the anaphoric feeder conditions a pronoun, matching the gender of the protagonist, or a nominal anaphor,

TABLE 1
Example Text (Translated from Dutch) Used in Experiment 1

| | |
|---|---|
| | By five o'clock Leo came home from his work. |
| | He hung his coat on the hallstand. |
| | . . . . . |
| | . . . . . |
| (he) | . . . . . |
| (Leo) | |
| (tired) | |
| (play billiards) | |
| | . . . . . |
| | . . . . . |

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

identical to the referent in the first sentence; in the thematic feeder conditions a noun, adjective or verb in citation form, the meaning of which was either related or unrelated to the theme of the first two sentences. Whether a word was related or unrelated to the theme was determined by a panel of six native speakers of Dutch.

Each text fragment was typed on a separate page. The two introductory sentences occupied separate lines. They were followed by one or two dotted lines. Then the feeder word appeared before the next dotted line, and the page was concluded with one or two more dotted lines. A text had three to five dotted lines.

*Design.* Four different lists of texts were constructed. These lists contained the texts in a fixed order. Each of the four different feeder word conditions of a text was allotted to a different list. Each list contained the same number of texts in each feeder word condition. The same condition never occurred twice in succession. The subjects were assigned to lists randomly with the restriction that each list was assigned to an equal number of subjects. In this way, texts and subjects were counterbalanced.

*Procedure.* The subjects were instructed to read the two introductory sentences and to write a continuation to the text fragment. They had to write one sentence on each of the dotted lines following the text fragment. It was stressed that the description in the two introductory sentences applied to only one protagonist, and that continuations must concern that protagonist. The sentences produced, together with the two introductory sentences, did not need to form a complete story – they could form a text fragment. The subjects had to write as many sentences as there were dotted lines on a page. They were told that now and then a word would be presented in brackets before a dotted line and they were instructed to use that word literally (except for inflections) in their next sentence (the target sentence).

The subjects had to use an opaque piece of paper to cover each new page before beginning to read the text. They uncovered the page line by line, after reading or writing a sentence. This procedure ensured that the subjects did not know beforehand how many lines they had to write, and when they had to use what kind of feeder word.

The experiment was administered to the subjects individually. The subjects wrote their continuations in about 50 min.

## Results

Data from target sentences in which a second character was introduced before, or in, the target sentence were considered invalid as were data from target sentences in which the feeder word was not used.

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

The target sentences in the anaphoric feeder conditions were assessed on theme continuity by four judges. To eliminate possible effects of the different forms of anaphoric reference on the judgements, the names, role descriptions or pronouns referring to the protagonist in the target sentences were replaced by -X-. A produced target sentence was considered a theme continuation or a theme shift if at least three judges independently made the same assessment and if all four were in agreement after discussion. All other cases were considered invalid.

For the target sentences in the thematic feeder conditions, the number of pronoun productions (*she* and *he*) and the number of definite noun phrase productions (names and role descriptions) relating to the protagonist were counted. Data from sentences in which no reference to the protagonist was made were considered invalid data.

*Anaphoric Feeder Data.* The proportion of theme shifts was computed for each subject in each condition and for each item in each condition excluding invalid data. These proportions were analysed by two two-way analyses of variance, one with subjects as a random variable and one with items as a random variable. Anaphoric feeder (name/role *vs* pronoun feeder word) was a within factor in both analyses, subject groups was a between factor in the first analysis, as was item groups in the second analysis. Table 2 shows the number of theme continuation and theme shift sentences, together with the data considered invalid.

The anaphoric feeder conditions resulted in a differential production of theme continuity types: the mean proportion of theme shift sentences was greater in the name/role feeder condition than in the pronoun feeder condition [$F1(1,16) = 152.18$, $P < 0.001$; $F2(1,12) = 234.56$, $P < 0.001$; min $F'(1,27) = 92.30$, $P < 0.001$]. If the feeder word was a name/role, the mean proportion of theme shift sentences was greater than 50% (null hypothesis) [$t1(19) = 8.49$, $P < 0.001$; $t2(15) = 11.49$, $P < 0.001$]. Pronouns as feeder words yielded a mean proportion of theme shifting sentences less than 50% [$t1(19) = 9.18$, $P < 0.001$; $t2(15) = 6.06$, $P < 0.001$].

TABLE 2
Number of Thematic Shift and Thematic Continuation Sentences Produced in the Anaphoric Feeder Conditions (with Percentages of the Analysed Data in Parentheses)

| | *Theme Continuity* | | |
|---|---|---|---|
| *Feeder Word* | *Shift* | *Continuation* | *Invalid Data* |
| Name/role | 61 (88) | 8 (12) | 11 |
| Pronoun | 12 (17) | 57 (83) | 11 |

TABLE 3
Number of Name/Role Descriptions and Pronouns Produced in the Thematic Feeder Conditions in Experiment 1 (with Percentages of the Analysed Data in Parentheses)

| | *Anaphor Type* | | |
|---|---|---|---|
| *Feeder Word* | *Name/Role* | *Pronoun* | *Invalid Data* |
| Unrelated | 6 (12) | 46 (88) | 28 |
| Related | 1 (2) | 56 (98) | 23 |

*Thematic Feeder Data.* The proportion of name/role references to the protagonist was computed for each subject in each condition, and for each item in each condition excluding invalid data. These proportions were analysed in the same way as the anaphoric feeder data. Thematic feeder was the within factor in both analyses. The number of anaphoric reference types produced and the data considered invalid are presented in Table 3.

The thematic feeder conditions did not result in a pronounced differential use of anaphoric reference types [$F1(1,16) = 2.85$, $P = 0.11$; $F2(1,12) = 2.55$, $P = 0.13$]. Although the mean proportion of pronouns produced in the thematic continuation feeder condition was greater than 50% [$t1(19) = 28.61$, $P < 0.001$; $t2(15) = 30.02$, $P < 0.001$], the same pattern was found in the theme shift condition. It is clear that pronouns were predominantly produced, whatever the theme relatedness of the feeder word.

## Discussion

The results of the thematic feeder conditions provide no evidence of a preference for using overspecified anaphors in a condition that is meant to produce a theme shift by the obligatory use of a theme-unrelated feeder word. The question, therefore, arises as to whether the feeder words really led to more productions of a theme shift when they were unrelated than when they were related. Therefore, the target sentences of the thematic feeder word conditions were assessed for theme continuity by one judge in the same way as the target sentences of the anaphoric feeder conditions. The subjects produced more theme continuation sentences (71%) than theme shift sentences (29%). This indicates that notwithstanding the unrelatedness of the feeder word, the theme was continued. How the subjects accomplished this can be illustrated by an instance produced in the thematic shift feeder condition of the example text in Table 1 (feeder word: *play billiards*): "He switched on the TV and watched (people)

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

playing billiards". In Experiment 2, this issue is further investigated by eliciting theme shift in a more controlled way.

The results of the anaphoric feeder conditions are much clearer: they indicate that if one encounters an overspecification, one is led to produce a theme shift. Whether one interprets the pronoun feeder condition in a similar way, i.e. as showing that if one encounters a pronoun one prefers to continue the theme, depends on whether one concludes from the thematic feeder conditions that continuations of the theme are produced by default.

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

## EXPERIMENT 2: STRIP DESCRIPTION

In Experiment 1, the obligatory use of a content word unrelated to the theme of the story did not often result in a theme shift. A plausible explanation is that if one has to add some sentences to a couple of introductory sentences, one has a preference to continue the theme of the story. In Experiment 2, the subjects had to think up a story that would contain a theme shift or a theme continuation before writing it down. The stories were elicited by strip cartoons in which a particular drawing introduced a theme shift or continued the existing theme. The first sentence produced by the subject triggered by that picture was taken as the target sentence, because the first sentence of an episode is generally considered as the position in which a theme shift is realised (Kieras, 1978; Lorch, Lorch, & Matthews, 1985).

It was hypothesised that in the theme shift condition, the reference to the protagonist would be overspecified in the target sentence, whereas in the theme continuation condition pronouns would be used.

### Method

*Subjects*. A total of 30 subjects participated in the experiment. In this experiment and in the following experiments, the subjects were students in the different disciplines of the University of Nijmegen. They were all native speakers of Dutch, and varied in age from 18 to 25 years. They were paid for their participation.

The data for six subjects were not analysed because they failed to write the minimum number of sentences for each picture (two subjects) or produced stories in dialogue or "I" form (four subjects). The results are based on the data for the remaining 24 subjects.

*Materials*. There were six experimental strip cartoons and one practice strip, in all relevant respects similar to the experimental strips. An example of the materials used is shown in Fig. 1. Each cartoon consisted of three or four pictures, that together depicted a story. The first picture introduced

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

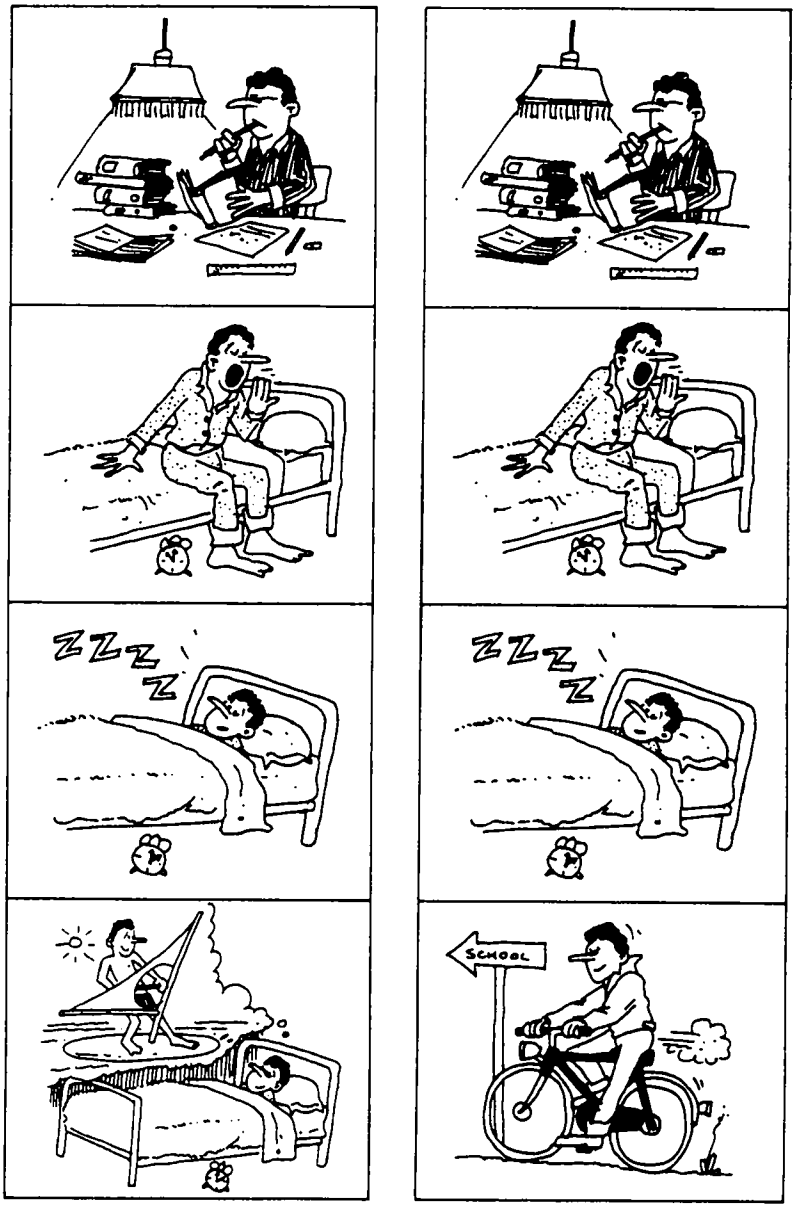



FIG. 1 Examples of strip cartoons used in Experiment 2: theme continuation version (left) and theme shift version (right). The theme shift is realised in the final picture.

one character in a situation or action. The other pictures also contained just this character. Each strip had two versions, differing only in the final picture. This picture either depicted an action of the protagonist that continued the action of the penultimate picture, or it had a theme shift: Its content did not directly link to the preceding action, but had an action at a different place and time from that of the penultimate picture. In both versions, however, the pictures of the strips belonged to one, more or less dull, story.

*Pre-test of the Materials.* The materials were pretested in a rating task. The subjects had to rate the thematic distance of each transition between pictures on a 5-point scale, ranging from 1 (no context change) to 5 (considerable context change). A total of 24 subjects participated in the pre-test. Two lists of strips were constructed. If a strip in one list was presented in the theme shift version, then the theme continuation version was put in the other list. In each list, half of the strips were of the theme shift version and half of the theme continuation version, the same condition never occurring more than twice in succession. Each list contained all strips. The lists were allocated to the subjects randomly, with the restriction that each list was assigned to an equal number of subjects. The order of the strips in the lists was quasi-random for each presentation of a list.

The mean ratings for each subject in each condition and the mean ratings for each strip in each condition were computed. The means from the penultimate and strip-final transitions were analysed together in two analyses of variance, one with subjects as a random variable and one with items as a random variable. Transition (penultimate *vs* final transition) and thematic continuity (shift *vs* continuation) were within factors. The mean ratings over subjects are shown in Table 4.

TABLE 4
Mean Theme Distance Rating of Penultimate and Final Transition in the Strip Cartoons as a Function of Thematic Continuity Type on a Scale of 1–5[a]

| | *Theme Continuity* | |
|---|---|---|
| *Transition* | *Shift* | *Continuation* |
| Penultimate | 1.61 | 1.93 |
| Final | 3.89 | 1.65 |

[a]1 = No content change; 5 = considerable content change.

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

There was a significant interaction between the factors theme continuity type and transition [$F1(1,23) = 141.54$, $P < 0.001$; $F2(1,5) = 57.91$, $P = 0.001$; min $F'(1,9) = 41.09$, $P < 0.001$]. The final transition between pictures was rated thematically more distant in the theme shift condition than in the theme continuation condition [$F1(1,46) = 299.15$, $P < 0.001$; $F2(1,10) = 116.72$, $P < 0.001$; min $F'(1,18) = 83.96$, $P < 0.001$], whereas the distance ratings did not differ significantly for the penultimate transition [$F1(1,46) = 4.46$, $P < 0.05$; $F2(1,10) = 1.74$, $P = 0.21$; min $F'(1,18) = 1.25$, $P = 0.28$]. (The analysis of the penultimate data can be interpreted as showing a tendency to a difference in thematic distance. Although the transition in both conditions is identical, such a tendency seems plausible. The subjects had to interpret the complete strip before rating the transitions. The final transition in the shift condition could easily have caused the penultimate transition to be perceived as more related and the final continuous transition could have caused the penultimate transition to be perceived slightly less related.)

The range of the mean ratings over subjects for the strip cartoons for the final transition in the theme continuation condition (1.42–2.17) did not overlap with the range in the theme shift condition (3.58–4.33). The mean ratings for the strips for the penultimate transition over both theme continuity conditions together ranged from 1.17 to 2.50. The results corroborate the effectiveness of the cartoonist's realization of thematic continuity type.

*Design.* Two lists of strips were constructed. The construction and presentation of the lists was the same as in the pre-test.

*Procedure.* The subjects were instructed to write a complete story on the basis of each strip cartoon. No separate description of the pictures was allowed, and the written story was to be comprehensible without the strips. The subjects had to look first at the complete strip and think of a story before starting to write. They had to write at least four sentences for each picture. After writing a story, the subjects had to indicate the (two or three) transitions between the pictures in their story by marking where the text accompanying each picture started. The sentence marked as the first sentence of the text for the final picture was considered the target sentence. In this way, the subjects themselves, not the investigator, determined which sentences were target sentences.

The experiment was administered to the subjects individually and took about 60 min.

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

## Results

A count was made of name/role expressions and pronouns (*she* and *he*) referring to the protagonist in the target sentences. Failures to refer to the protagonist, or introductions of a second character in either the target sentence or the preceding sentences, were considered invalid data. To determine whether the strip-final picture had indeed elicited the intended thematic continuity types, all of the target sentences were assessed on theme continuity. The data were analysed in two ways: with theme continuity of the strips (as defined by the cartoonist before the experiment was carried out) as the independent variable and with theme continuity of the subjects' productions (as determined by judges after the experiment had been carried out) as the independent variable. The same kind of assessment procedure was applied as in the anaphoric feeder conditions in Experiment 1, except for the number of judges involved. There were two judges in this experiment. For 9% of the target sentences, their assessments did not match. These sentences were counted as theme continuation or theme shift if the two judges agreed after discussion. The two remaining cases were considered invalid data.

*Thematic Continuity as Defined by Strip Condition.* The proportion of name/role references to the protagonist was computed for each subject in each condition, and for each item in each condition excluding invalid data. The proportions were analysed by two one-way analyses of variance, one with subjects as a random variable and one with items as a random variable. Thematic continuity as defined by strip condition (theme shift *vs* theme continuation) was a within factor in both analyses. The number of anaphoric reference types produced and the data considered invalid are presented in Table 5.

The theme continuity conditions as defined by strip conditions resulted in a marginally significant difference in production of anaphoric reference

TABLE 5
Number of Name/Role Descriptions and Pronouns Produced in the Thematic Continuity Conditions as Defined by Strip Cartoon Conditions in Experiment 2 (with Percentages of the Analysed Data in Parentheses)

| | *Anaphor Type* | | |
|---|---|---|---|
| *Strip Version* | *Name/Role* | *Pronoun* | *Invalid Data* |
| Theme shift | 31 (48) | 34 (52) | 7 |
| Theme continuation | 17 (25) | 50 (75) | 5 |

types [$F1(1,23) = 4.00$, $P = 0.05$; $F2(1,5) = 5.46$, $P = 0.07$; min $F'(1,19) = 2.31$, $P = 0.14$]. That this effect is only marginally reliable can be attributed to the theme shift condition. In this condition, as many pronouns were produced as nouns. In the theme continuation condition, however, the expectation was confirmed – the mean proportion of pronouns was greater than 50% [$t1(23) = 4.08$, $P < 0.001$; $t2(5) = 3.36$, $P < 0.02$].

*Thematic Continuity as Determined by Assessment of Subjects' Productions.* The subjects produced more target sentences with a theme continuation than with a theme shift [73 *vs* 57; invalid data = 14). The number of anaphoric reference types produced are presented in Table 6.

TABLE 6
Number of Name/Role Descriptions and Pronouns Produced in the Thematic Continuity Conditions as Determined by Assessment of Subjects' Productions in Experiment 2 (with Percentages of the Analysed Data in Parentheses)

| | *Anaphor Type* | | |
|---|---|---|---|
| *Produced Theme* | *Name/Role* | *Pronoun* | *Invalid Data* |
| Shift | 34 (60) | 23 (40) | 14 |
| Continuation | 14 (19) | 59 (81) | |

The proportion of name/role references to the protagonist in the target sentences was again computed for each subject in each condition, and for each item in each condition excluding invalid data. These proportions were analysed by a one-way analysis of variance with subjects as a random variable. Thematic continuity as determined by assessment of subjects' productions (theme shift *vs* theme continuation) was a within variable. There was a main effect of thematic continuity [$F1(1,23) = 19.76$, $P < 0.001$]. The mean proportion of pronouns produced in the theme continuation condition was again greater than 50% [$t1(23) = 4.86$, $P < 0.001$], but although the proportion of name/role references produced in the theme shift condition seems to have increased compared to the production in the same condition when defined by the cartoonist, the mean proportion of name/role references was not significantly different from 50% [$t1(23) = 1.14$, $P > 0.20$].

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

## Discussion

As in the thematic shift feeder condition in Experiment 1, the production of theme shift target sentences in stories describing the theme shift version of the strips in Experiment 2 was lower than was expected. This result may be due to the inadequacy of the experimental manipulation. Although this is rather plausible in Experiment 1, it is less plausible in Experiment 2. Transitions in theme shift versions were clearly judged more distant than transitions in theme continuation versions. Rather, the result must be attributed to the inclination of writers to produce a continuous story if they are asked to produce a continuation or a complete whole. This may be especially so if one has to write a short discourse.

Although in comparison to Experiment 1 the proportion of name/role references in the theme shift condition in Experiment 2 increased, it did not differ from 50%. Therefore, there may be other reasons why such a large proportion of pronouns was produced. In the theme shift condition of the strips, an action was depicted that occurred at a different time and place from that in the penultimate picture. References to time and place, therefore, are likely to occur in the stories. Preposed adverbial phrases and subordinate clauses of time and place are considered salient (Davison, 1984), tend to have a higher probability of carrying discontinuous themes (Givón, 1983b) and are commonly taken as markers of theme change (Brown & Yule, 1983; Van Dijk & Kintsch, 1983). Examples of these expressions produced in Experiment 2 are *The following day*, *After a good night's sleep*, *At the border* and *After the lessons were finished*.

The target sentences were once again grouped with theme continuity as determined by assessment of the subjects' productions as the independent variable, this time excluding target sentences containing a preposed expression of time and place from the anaphoric reference data. The numbers of the anaphoric reference types in the reduced data set are presented in Table 7, as is the number of sentences with preposed adverbial phrases and clauses of time and place.

The large number of expressions of time and place in theme shift sentences is striking. From the differences in the numbers presented in Tables 6 and 7 it can be seen that, in the theme shift condition, 21 of the time and place expressions were produced in sentences with a pronoun, and 12 in sentences with a name/role anaphor. If an expression of time and place is used in the sentence, the theme-shifting character of the sentence seems as well served with a pronoun as with an overspecified expression. It is also clear that, if expressions of time and place do not occur in a theme-shifting sentence, the pronoun is preferably not used, but instead the protagonist is referred to by an overspecified anaphoric expression.

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

TABLE 7
Number of Name/Role Descriptions and Pronouns in Sentences without Expressions of Time and Place and Number of Expressions of Time and Place Produced in the Thematic Continuity Conditions as Determined by Assessment of Subjects' Productions in Experiment 2 (with Percentages of the Analysed Data in Parentheses)

| | *Anaphor Type* | | | |
|---|---|---|---|---|
| *Produced Theme* | *Name/Role* | *Pronoun* | *Time/Place Expressions* | *Invalid Data* |
| Shift | 22 (92) | 2 (8) | 33 | 14 |
| Continuation | 14 (21) | 53 (79) | 6 | |

Summarising the results of the production experiments: It was experimentally established in Experiment 2 that overspecified anaphoric reference can be used when one has a theme shift in mind, and is usually used when there are no other linguistic means to indicate a theme shift. Experiment 1 showed that if one encounters an overspecified anaphoric reference, one is led to produce a theme shift. The question that should be addressed now is whether an overspecification can function as a signal of a thematic shift to the reader/listener. This will be investigated in Experiments 3 and 4.

## EXPERIMENT 3: PROBE RECOGNITION

If a thematic shift is expressed in a discourse, an anaphoric reference is used that is more specific than is necessary to unambiguously refer to the intended referent. From the point of view of comprehension, the question arises as to whether an overspecified referential expression can effect a theme shift or, put more neutrally, can function as a signal of a theme shift to the comprehender.

In comprehending a discourse, readers/listeners are incrementally constructing a representation of the information in the text. Sentence after sentence is integrated with the preceding discourse, especially with the information in the immediately preceding sentence. If, however, a shift in theme is intended by the writer/speaker, the information in the new sentence must not be attached to the information in the directly preceding sentences. A signal not to integrate would help the comprehender to set up a new partition in the text representation.

An overspecified anaphoric reference might function as such a signal. It would indicate that the reader should connect the current incoming infor-

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

mation not to the chain of information of the preceding sentences, but instead create a new chain of information (to be attached eventually to the antecedent that is referred to in the overspecified expression). As a result, a representation of the information in the sentence containing the overspecification is not connected – or at least it is more difficult to connect – to the representation of the immediately preceding sentence.

One way the system could accomplish this is that the focus of attention is shifted away from the running theme, and that the representation of the sentences in which the running theme was expressed becomes less available, with the outcome that the current incoming information is not attached to the information in the immediately preceding sentences. In an analogy to systematic storage: the overspecified expression signals that a new box should be opened in which a new chain of information can be stored. A referential expression that is not overspecified does not influence the storage process of the information in the current sentence, but, rather, it puts the incoming information in the current box: the information in the current sentence and the preceding sentences are stored in the same box.

To investigate whether an overspecified referential expression affects the incrementing text representation differently from a non-overspecified one, the availability of information in the sentence preceding the referential expression was measured. The information of interest is not the antecedent information in that sentence – availability of antecedent information would address the identificational function of the referring expressions – but information contained in the predicate of the sentence. The availability of this information was measured using a probe recognition technique (McKoon & Ratcliff, 1980). In this technique, a probe word is presented to the subject and the subject has to decide whether this word was contained in the text or not. The time to recognise the probe as a word from the preceding text is taken to reflect the availability of the information in that part of the text. A comparison of the recognition times under different conditions indicates the relative availability of the word under those conditions.

If the availability of the information in the preceding sentences is indeed diminished by encountering an overspecified referential device, then the recognition time of a word from the preceding text should be longer if such a referential device is used than if a non-overspecified device is used.

In research using the same probe recognition technique, Chang (1980) found that the antecedent itself was more available after reading an overspecified anaphor (the probe repeated the lexical material) than after encountering a pronoun, whereas Cloitre and Bever (1988), avoiding the lexical repetition of the probe word, concluded the opposite when testing an antecedent-related probe word (a modifying adjective). The availability effects were accounted for in terms of the identificational function of the

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

anaphor. In the present study, the availability of the antecedent is not the issue – this study addresses the availability of the content information. Therefore, the probe word was not directly related to the antecedent, although it was from the same clause as the antecedent. Dell, McKoon, and Ratcliff (1983) showed with the same technique that the activation of such a content word in the antecedent clause is increased by a later reference to the antecedent, but this activation appeared rather short-lived. If in the present study the referential expressions differentially affect the availability of a content word from the same clause as the antecedent, this effect cannot be easily explained in identificational terms.

A further question is *when* the overspecified device affects the availability of the preceding information. This might be only at the end of the incoming sentence (cf. sentence wrap-up: Just & Carpenter, 1980), or, more immediately after encountering the anaphor. Experiment 3 investigated the availability of information preceding the incoming sentence at the end of that sentence.

Because the occurrence of an anaphoric reference type is correlated with the type of theme continuity in the sentence, as was shown in Experiments 1 and 2, the linguistic material in natural text is likely to be different for the different anaphoric devices. For instance, information that expands the preceding sentence might keep lexically expressed concepts in that sentence more available than information from a new theme would do. Therefore, the content material itself can affect the availability of the preceding information, if this availability is measured after the sentence is read (cf. O'Brien, Duffy, & Myers, 1986). For that reason, overspecified anaphors were investigated in theme-continuing sentences as well as in theme-shifting sentences, as were non-overspecified anaphors. In the experiment, a name with an extension (e.g. *Hella Haasse, the celebrated writer*), was used as the overspecified anaphoric expression, whereas a pronoun (*she* or *he*) was used as the non-overspecified anaphor.

Since the overspecification is assumed to signal a new theme, the reasoning about the availability of words from preceding sentences should *a fortiori* apply to theme continuity – recognition time should be longer after a theme shift than after a theme continuation.

More specific predictions follow from what is hypothesised to be the text representation after reading the target sentence. If a target sentence contains an overspecified anaphor and a theme shift, the prediction is straightforward. The anaphor signals that the incoming information should not be attached to the current chain of information, but instead that a new chain should be opened, which will contain information that also by its content does not fit the current chain. Therefore, the probe word should be relatively unavailable. If the target sentence contains a pronoun and a theme continuation, the prediction is equally straightforward. The pro-

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

noun provides no reason not to attach the incoming information to the current chain, nor does the content of the sentence. Therefore, the probe word should be relatively available. The prediction of the occurrence of a pronoun in a theme-shifting sentence is more problematic. The pronoun provides no reason not to attach the information to the current chain, but the content itself does not fit in. It is assumed that after reading the theme-shifting sentence, the content is decisive in structuring the discourse. The probe word should be relatively unavailable once the information in the sentence has been processed. The most interesting condition is the use of an overspecified anaphor in a theme-continuing sentence. The overspecification indicates that the information should not be attached to the current chain, although the content could fit (indeed does fit) the current chain. It is claimed that after reading the incoming information, a new chain is opened, at least for the time being. The probe word should therefore be relatively unavailable. This claim is a strong one, because it states that under these circumstances guiding linguistic signals are relied upon more, at least initially, than semantic content (*vs.* e.g. Kieras, 1980).

This leads to the following predictions: In the theme continuation condition, the probe word should be less available if the overspecified anaphor is used than if the pronoun is used; if a pronoun is used, the probe word should be less available in the theme shift condition than in the theme continuation condition.

## Method

*Subjects.* A total of 36 subjects participated in the experiment. The data for four subjects were not analysed because they made more than 33% errors in the probe recognition task. The results are based on the remaining 32 subjects.

*Materials.* The experimental materials consisted of 24 expository texts about the profession and/or principles and opinions of a single protagonist. There were 21 filler texts and 10 practice texts. These texts were about one or more protagonists, or had no protagonist, and appeared in only one version. Each of the experimental and filler texts had a probe word associated with it. The filler texts had three purposes: first, to prevent subjects from realising what was manipulated in the experimental texts; secondly, to allow the presentation of probe words that had not occurred earlier in the text – the probes in the experimental texts all did occur in the text; and, thirdly, to allow probe words at different presentation positions in the text, so that the subjects could expect a probe at any position. In all other relevant respects, the filler texts were similar to the experimental texts.

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

TABLE 8
An Example of Text (Translated from Dutch), in Four Conditions, Used in Experiments 3 and 4

Professor Alan Johnson is a very busy man.
In addition to being the father of a large family, he is employed at the medical faculty of the University of Utrecht.
His current research subject is massage as therapy.
There are, he tells us, a large number of different massage techniques, and new techniques are added each year.
He mentions footsole-massage as one of the most important techniques.
Johnson was trained as a masseur in the past.
He still works regularly as a masseur.
In this way he keeps in touch with the *field* and (interesting) ideas for new research come up (again and again).

THEME SHIFT VERSION
*Johnson, a professor of medicine, is the father of seven children.*
*(He is the father of seven children.)*
Although his work demands a lot of him, he always finds time for them.
His children appreciate this greatly.

THEME CONTINUATION VERSION
*He considers this research important.*
*(Johnson, a professor of medicine, considers this research important.)*
Therefore he spends much time in his laboratory, where he supervises the research of many younger colleagues.

*Note*: Italic print indicates the target sentences and the probe word in the pre-target sentence. It is used for the benefit of the reader only; the subjects saw all materials in regular print.

The experimental texts were constructed in four conditions. Table 8 shows an English translation of one of the experimental texts (the original Dutch version is presented in Appendix 1). Each text had two main versions. In one version, a major theme shift occurred in a sentence at a certain point in the text, opening a new theme, and in the other version the sentence at that point was continuous with the preceding text (these sentences are henceforth called target sentences). In the target sentences, a reference was made to the protagonist. In both versions, this reference could be made by an unambiguously referring pronoun (viz. *he* or *she*) or by an overspecified anaphor, namely an extended name (e.g. *Johnson, a professor of medicine*). The extension consisted of information that was either explicitly stated earlier in the text, or was easily inferrable from the preceding text or from general knowledge (e.g. *the famous scientist*, as an extension to *Einstein*). The anaphor served as the grammatical subject of the sentence, and was positioned at its default position, the beginning of the sentence.

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

The probe word for the experimental texts was always contained in the sentence preceding the target sentence. This pre-target sentence consisted of two co-ordinated finite clauses connected with the conjunction *and*. The probe word was contained in the first clause of the pre-target sentence. The probed word and probe had exactly the same form. The probe word was not central to the current theme, nor directly related to the protagonist. It was a high-frequency noun ranging from 1500 to 5000 in the CELEX database (Burnage, 1990), a computerised lexical database based on 42 million tokens (CELEX Dutch Database, 1989, Release N2.7). It consisted of one to three syllables, most (15) being two syllables, and all but one probe occurred in citation form (the one exception being a plural). Some of the probe words for the filler texts occurred in these texts, and some did not. The positive filler probes were similar to the experimental probes, except for presentation position: they occurred at all positions in the text. The negative probe words and their presentation positions were similar to the positive filler probes. The negative probe words were related to the text as a whole, and could have been used in it, but they were not near synonyms of particular words from the text. Care was taken that a probe word or a compound or declination of it did not occur earlier, either in the current text or in any preceding text (except, of course, the positive probe word in the preceding sentence). There were 55% positive probes in total.

In the first clause of the pre-target sentence, the subject of the clause was a pronoun that referred unambiguously to the protagonist. In the second clause, no reference was made to the protagonist. The information in the target sentence in the theme continuation condition was related to the information in this second clause. The content and the lexical material in the second clause of the pre-target sentence was not related, and in particular not anaphorically related, to the first clause. It was chosen in such a way that the theme continuity of the target sentence in the theme continuation condition was supported by cohesive means (Halliday & Hasan, 1976; Hustinx, in prep.). As a result, the target sentence in the theme continuation condition was not lexically related to the *first* clause of the pre-target sentence, except, of course, for the anaphor.

The length of the target sentences in the theme shift and theme continuation conditions was about the same, the mean length of the sentences being one character longer in the theme shift condition than in the theme continuation condition. If the protagonist in the target sentence was referred to by a pronoun, the length of the second clause of the pre-target sentence was increased by about the same number of characters as the difference between the lengths of the extended name and the pronoun in the target sentence, the mean lengthening of the pre-target sentence being 3.5 characters more than the difference between the lengths of the refer-

ring expressions. Therefore, the mean length in characters between probed word and probe was not longer in the overspecified condition than in the pronoun condition. (Although the distance from probed word to probe was about the same, the length of the target sentence itself differed considerably between these conditions, making reading times incomparable.)

Up to the target sentence, the text versions were identical, except for the just-mentioned insertion of words in the pre-target sentence. The number of sentences in this part of the text varied from 6 to 10, that is from 10 to 16 lines of text. There was only one protagonist. This protagonist was introduced in the first sentence by his or her proper name. To keep the protagonist foregrounded, the protagonist was regularly referred to in the subsequent sentences, mainly by pronouns. One or two of the references were in sentences with a minor content shift, and were realised with an anaphor that repeated the name of the protagonist. This name anaphor was used to make it very unlikely that the proper name in the target sentence would be interpreted as a new character. There were at least two sentences between the last appearance of the name anaphor before the target sentence and the target sentence itself. The content shifts in the part of the text before the target sentence were intended to be minor, in order to allow a major theme shift in the target sentence.

The sentences after the target sentence were different for the two theme versions. Because in the theme shift version a major new theme was introduced, it was impossible to continue with the same sentence in the theme shift version as in the theme continuation version. The number of lines after the target sentence varied from 3 to 7.

To encourage reading for meaning, verification statements were added to 2 practice and 12 filler texts. Half of the statements were true and half false with respect to the information in the text.

*Pre-test of the Materials.* The materials were pre-tested in a rating task. The subjects had to rate the continuity of each sentence with its immediate preceding sentence on a 5-point scale, ranging from 1 (full content continuity) to 5 (content discontinuity). In the instructions, it was stressed that the subjects had to make their judgements on the basis of the content of the sentences. As in Experiment 1, the subjects had to use an opaque piece of paper, with which they uncovered the text sentence by sentence. This ensured that the information following the sentence that they had to rate could not influence their judgement. A total of 32 subjects participated in the pre-test. Four different lists were constructed. These lists contained the texts in two different fixed orders. Each of the four versions of a text was assigned to a different list. Each list contained the same number of texts in each condition, the same condition never occurring twice in succession. The subjects were assigned to lists randomly, with the restriction that each

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

list was assigned to an equal number of subjects. In this way, the texts and subjects were counterbalanced.

The mean rating score for each sentence was computed over subjects for all conditions for the sentences up to the pre-target sentence. For this sentence and the target sentence, the mean rating score was computed over subjects, separately for each condition. The rating scores were used to check for unintended content shifts and to check that the intended content shifts were perceived as major shifts. The results of the pre-test were used to make changes to the materials.

In any case, in the original versions, the range of the mean ratings for the target sentences in the theme continuation condition (1.19–2.69) did not overlap with the range of mean ratings in the theme shift condition (2.88–5.00).

The mean ratings for each subject in each condition and the mean ratings for each target sentence in each condition were computed. These means were analysed in two three-way analyses of variance, one with subjects as a random variable and one with items as a random variable. In the first analysis, theme continuity type and anaphoric reference type were within factors, and subject groups was a between factor. In the second analysis, anaphoric reference type was a within factor, and theme continuity type and item groups were between factors. The mean ratings over subjects are shown in Table 9.

Theme continuous target sentences were rated as having greater content continuity than theme shift target sentences [$F1(1,28) = 507.34$, $P < 0.001$; $F2(1,40) = 380.98$, $P < 0.001$; min $F'(1,67) = 217.59$, $P < 0.001$]. It was concluded that, overall, the subjects' perception of the constructed theme-continuous and theme-shifting target sentences was as intended.

TABLE 9
Mean Content Continuity Rating of Target Sentence in Experiment 3 as a Function of Anaphoric Reference Type and Theme Continuity Type on a Scale of 1–5[a]

| | *Theme Continuity* | |
|---|---|---|
| | *Shift* | *Continuation* |
| Extended name | 4.23 | 1.99 |
| Pronoun | 4.40 | 1.73 |

[a]1 = Complete content continuity; 5 = discontinuity.

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

The ratings of the target sentences in the extended name condition did not differ from those in the pronoun condition (both $F1$ and $F2 < 1$). The interaction between theme continuity and anaphoric reference was significant [$F1(1,28) = 6.89$, $P = 0.01$; $F2(1,40) = 7.12$, $P = 0.01$; min $F'(1,65) = 3.50$, $P = 0.06$]. Despite the instruction to assess continuity by content, it seems that an anaphoric reference type that is not used in accordance with theme continuity type causes the sentences to be perceived as more discontinuous.

The two sentences preceding the target sentences were rated as content connecting (means, in order of appearance in the text, 1.95 and 1.83). These means are comparable to the rating of the continuous target sentence.

*Design.* Four lists of the texts were constructed. The construction and presentation of the lists were the same as in the pre-test, except for the insertion of practice and filler texts. There were never more than three positive or negative probe words in succession.

*Procedure.* The experiment was run with the APLEX program (Kusters et al., 1989) on an Olivetti M24 personal computer. The text was presented using a moving window technique, the unit being a sentence. The sentences were masked by replacing each character by a hyphen, while the punctuation marks remained. Each sentence started on a new line. The presentation was subject-paced.

The presentation of each text was preceded by the words "New Text" for 2000 msec, followed by an asterisk, which disappeared when the subjects pushed a button. At that moment, the first sentence of a text appeared, the remainder of the text being masked. As soon as the subject pushed the button, the current sentence was masked and the next sentence appeared. The time between the onset of a sentence and the button press was registered. At some point in each text, the button press caused two asterisks, 14 character spaces apart, to appear on an otherwise empty screen, in the centre of the line where the subject expected the next sentence to appear. After 1000 msec, these asterisks disappeared and the probe word was presented where the empty space between the asterisks had been. The subjects had to indicate whether the word had occurred in the text by pressing one of two buttons. The time between the onset of the probe and the pressing of the button was registered as the probe recognition time. The probe word disappeared after the button press, and the remaining text was presented in the same way as the text before the probe, starting at the line where the probe word had been presented.

If the text was followed by a verification statement, the word "Verification" was shown for 1500 msec after the final sentence, whereafter the

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

verification statement appeared. The subjects had to press one of two buttons to indicate whether the statement was "true" or "false" with respect to the text. After the final sentence of the text or after the verification statement, the words "New Text" were presented again.

The subjects were instructed to read the text carefully so as to understand it and to be able to correctly verify the statements. It was explained to the subjects that they had to read at the pace they would normally read, but that they had to react as quickly as possible to the probe word, without making errors. To make it easier for the subjects to know when to use which button, they were instructed to be ready for one button when one asterisk appeared, and for two buttons when two asterisks appeared.

The experiment took about 60–75 min to complete.

## Results

The mean proportion of errors on the experimental probes was 12.5%, that on the filler probes 17.0%. The percentage of errors is shown in Table 10 as a function of theme continuity and anaphoric reference type. There was no significant difference in the number of errors between the anaphor type conditions [Wilcoxon $T(n = 23) = 86.5$, $P > 0.10$]. The only conditions that differed from each other were name in thematic shift and pronoun in thematic continuation [$T(n = 17) = 34$, $P < 0.05$].

Probe recognition times that exceeded the mean ± 2 standard deviations for that probe and that subject (1.4%) were substituted by estimates on the basis of subject and probe means. The recognition times were analysed by two three-way analyses, one with subjects and one with items as a random variable. Theme continuity type and anaphoric reference type were within factors in both analyses, subject groups was a between factor in the first analysis, as was item groups in the second. The mean probe recognition times are presented in Table 10.

TABLE 10
Mean Probe Recognition Times (msec) in Experiment 3 as a Function of Anaphoric Reference Type and Theme Continuity Type (with Percentage of Errors in Parentheses)

| | *Theme Continuity* | |
|---|---|---|
| | *Shift* | *Continuation* |
| Extended name | 1139 (15) | 1128 (14) |
| Pronoun | 1129 (13) | 1054 (8) |

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

Although the interaction of the two variables was not significant [$F1(1,28) = 1.34$, $P = 0.26$; $F2(1,20) = 2.20$, $P = 0.15$], the planned comparisons were as predicted. The recognition time for the probe word after a theme-continuous sentence was significantly longer in the extended name than in the pronoun condition [$t1(56) = 2.21$, $P < 0.025$ (one-tailed); $t2(40) = 2.24$, $P < 0.025$ (one-tailed)]. So, the probe word was less available after encountering an overspecified reference than after encountering a pronoun in continuous text. The recognition time for the probe after a theme-continuous sentence with a pronoun was shorter than after a theme-shift sentence with a pronoun [$t1(56) = 2.15$, $P < 0.025$ (one-tailed); $t2(40) = 2.09$, $P < 0.025$ (one-tailed)]. Therefore, in the pronoun condition, the probe word was less available in a theme-shift text than in a theme-continuous text. There was no difference between the pronoun and extended name conditions in the theme-shift text [$t1(56) = 0.27$; $t2(40) = 0.28$].

## Discussion

The results of the *a priori* comparisons show that an overspecified anaphoric reference causes a decrease in the availability of words from the preceding text. These words appear to be relatively unavailable after information that should have been integrated with the immediately preceding part of the text is encountered. Apparently, the influence of the overspecification was not overruled by the influence of the theme continuity of the content information. It is an open question as to whether the representation of the information is simply not connected, or whether the content is interpreted differently under the influence of the signal to expect a new theme. A further question, then, is whether the availability of the probe word is affected by the overspecified anaphor only indirectly, through a different interpretation of the content information, or also directly. As a first step towards answering this question, one could investigate when the overspecified anaphor affects the availability of the probed word. If this happens before the content information is encountered, we can rule out an explanation by which the overspecified anaphor exerts its influence only indirectly. In Experiment 4, the availability of preceding information is investigated not at the end of the target sentence as in Experiment 3, but after the anaphoric expression, early in the target sentence.

# EXPERIMENT 4

This experiment investigated whether an overspecified anaphor can directly affect the availability of previous information. Therefore, the probe word should be presented immediately after the referential express-

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

ion at the beginning of the sentence, before any content information in the target sentence is encountered.

In view of the rather long texts, it seemed inappropriate to present the text word by word, especially as the subjects were to read the texts self-paced. Therefore, the units generally contained more than one word. In such a presentation, a unit containing only a pronoun would have been marked. For that reason, the probe in the target sentence was not presented immediately after the referential expression, but after the next word, the finite verb. This presented the problem that content information could have been processed before the probe word was presented. In constructing the materials we tried to put little content information in the finite verb.

## Method

*Subjects.* A total of 19 subjects participated in the experiment. The data for three subjects were not analysed because they made more than 33% errors in the probe recognition task. The results are based on the remaining 16 subjects.

*Materials and Design.* In Experiment 4, the materials of Experiment 3 were used. The experiments were the same in all respects, except, of course, for the presentation position of the probe, which, in Experiment 4, was at the beginning of the sentence, after the anaphoric reference and finite verb (in Dutch, the main verb is separated from the auxiliary in the sentence constructions used in the target sentence).

The finite verbs of the target sentence consisted mainly of auxiliaries, copulas and light verbs, but they were not the same for the two anaphor conditions, and could potentially carry thematic information, either by their form or by their content. Therefore, as in Experiment 3, theme continuity was included as a factor in the design. The design was the same in both experiments.

*Procedure.* The procedure was the same as in Experiment 3, except for the units of presentation. These units were 2–6 words in length. The number of units between probed word and probe were the same in all conditions. In general, the words within a unit were all from the same constituent. Because the pronoun and finite verb in the target sentence, which were presented as a unit, do not form such a constituent, such combinations were also presented in a substantial number of other sentences. The verification statements were presented as one unit.

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

## Results and Discussion

The mean proportion of errors on the experimental probes was 9%, that on the filler probes 7%. The percentage of errors was not different between extended name condition (10%) and pronoun condition (9%).

The analyses of variance on the probe recognition data were carried out in the same way as in Experiment 3. There were 0.8% outliers. The main effect of thematic continuity was not statistically significant [$F1(1,12) = 2.58$, $P = 0.13$; $F2(1,20) = 3.77$, $P = 0.06$; min $F'(1,26) = 1.53$, $P = 0.22$]. The results show at most a tendency for the probe recognition time in the thematic shift condition (985 msec) to be slower than in the thematic continuation condition (952 msec). Because theme continuity and anaphoric reference did not interact [$F1(1,12) = 1.33$, $P = 0.27$; $F2 < 1$], the main effect of anaphoric reference is relevant. This main effect was significant [$F1(1,12) = 16.30$, $P < 0.01$; $F2(1,20) = 24.90$, $P < 0.001$; min $F'(1,26) = 9.85$, $P < 0.01$]. The mean probe recognition time was longer in the extended name condition (1007 msec) than in the pronoun condition (930 msec).

The results show that the overspecified expression makes the information in the preceding sentence less available early in the sentence. This is interpreted as support for the claim that the availability of previous content information depends on the choice of the anaphoric device. The overspecification of an anaphoric device itself functions as a signal to set up a new chain of information. The different finite verbs appeared not to affect the probe recognition time differentially. This may be because there was not much information in the verbs, or because the information in the two conditions did not differ in any relevant respect.

# GENERAL DISCUSSION

Observations about texts, and analyses in the literature, indicate that the specificity of anaphoric devices is related to the thematic structure of texts. The present experiments show that this relation can be interpreted in a functional way. Overspecified anaphoric expressions determine to a certain extent the construction of the representation. An overspecified anaphoric device as feeder word in the first production experiment triggered a thematic shift. This effect of overspecification is all the more striking as subjects appear to be inclined to continue themes instead of making thematic shifts in producing stories, a fact also demonstrated in Experiment 1: subjects given feeder words unrelated to the theme tend to make thematic continuations.

If people do make thematic shifts – as, for example, in the descriptions of the discontinuous strip cartoons in Experiment 2 – the anaphoric

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

expression does not function merely as an identificational device. Indeed, when in those situations other devices were used to mark thematic boundaries – such as preposed adverbial phrases of time and place – there was a preference to use pronouns. This demonstrates that pronouns were sufficient for identificational purposes and that overspecified anaphoric expressions and preposed place and time expressions can be used interchangeably as thematic structure indicators. These results show that overspecifications are functional in structuring the text thematically.

The comprehension experiments demonstrate that, corresponding to their role in language production, anaphoric expressions signal a thematic shift to the reader. The overspecification contributes to starting a new theme and closing the previous one. Two possible explanations of this effect were discussed. The first is that the overspecification changes the focus of attention to something new, making the previous theme less available. The second is that overspecification guides the understanding of the subsequent information, so that it is interpreted as a new theme, thereby making the information of the previous theme less available. The results of Experiment 4 indicate that at least the first explanation is correct. It is quite possible, however, that both explanations are correct. If one changes the pronoun *she* in line (4) of the story about Sally Jones in the Introduction into the overspecified proper name *Sally*, this intuitively seems to affect the perspective that is created by that sentence: one expects that the text will continue with a reason for her looking forward.

The fact that rather strongly overspecified referential expressions were used in Experiments 3 and 4 deserves some discussion. The overspecifications were extended names such as *Johnson, a professor of medicine*. Two alternative interpretations may be suggested for the results.

The first is that the probe recognition time difference between pronoun and extended name conditions in the theme continuation version of the text can be explained in terms of a difference in length of the descriptions. In order to exclude this factor, the length of the second clause of the pre-target sentence was increased in the pronoun condition in such a way that the length of the material between the probe word in the pre-target sentence and the sentence-final word of the target sentence – expressed as number of characters – was about the same in both conditions. But there is still the possibility that the time it takes to read the materials in the two conditions is different, and that this difference causes a difference in availability of the probe word, resulting in the difference in the recognition times. The time spent reading the materials between probed word and probe can be equated to the reading time of pre-target and target sentence together (minus an unknown constant). Analyses of variance of these summed reading times (Experiment 3) revealed that the reading times indeed differed. But the differences did not match the differences in the

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

probe recognition data; although the difference in reading time between the extended name and the pronoun conditions in the theme continuation condition matched the difference in probe recognition time between the same conditions, the differences in reading time and probe recognition time did not match for theme continuation and theme shift in the pronoun condition. Therefore, the results cannot be explained in terms of time elapsed between reading the probed word and testing the probe.

The second alternative interpretation concerns a difference between the pronoun and extended name conditions in terms of the number of propositions between the probed word and the probe. Although the extension in the extended name condition is old information, it is an extra proposition (e.g. *a professor of medicine*). One may wonder whether the insertions in the pre-target sentence in the pronoun conditions (e.g. *interesting*, and *again and again*) are propositions on a par with the proposition in the extension of the name. One may also argue that it makes a difference where the extension occurs – in the pronoun condition, the insertions occur before the sentence wrap-up (Just & Carpenter, 1980) of the pre-target sentence, whereas the proposition in the extended name condition occurs in the target sentence itself, just before the probe is presented. If these differences can be related to the probe recognition times, then the results could be explained in terms of (the number of) propositions between probed word and probe. However, although the number of propositions could explain the difference in probe recognition times between the pronoun and the extended name conditions in theme continuations, it is incompatible with the lack of such a difference in the theme shift condition (Experiment 3, Table 10). So the results are unlikely to be explained in terms of the number of propositions.

The effect of overspecified referential expressions should, of course, be confirmed by a comparison between other referential expressions, in particular between pronouns and names without extension. These data are not yet available. But data on preposed locative phrases obtained in an experiment similar to Experiment 3 (Hustinx, in prep.) do corroborate the idea that the number of propositions in the overspecification is not the decisive factor. In this experiment, the second clause of the pre-target sentence contained a reference to a place, as in *at the airport*, which was mentioned earlier in the text, for instance *Schiphol*. The target sentence started with a locative phrase, that was either an overspecification, for instance the locative *At Schiphol*, or was not an overspecification, for instance the prolocative *Here*. The interaction between the factors locative *vs* prolocative and theme continuity (shift *vs* continuation) was highly significant. The mean probe recognition time for the prolocative in the theme continuation condition differed significantly from the other three means (all *F*1s and *F*2s were significant), whereas these three means

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

were almost identical (all $F1$s and $F2$s $< 1$). This result confirms the conclusion of Experiments 3 and 4: Overspecified anaphoric expressions contribute to structuring the discourse in the comprehension process.

To sum up, there is a close relation between anaphoric expressions and discourse structure. Part of this relation can be explained in terms of what may be called the identificational function of anaphors. Thematic structure in the discourse may affect the identificational function of an anaphoric expression. A thematic change in the discourse affects the availability of the antecedents. The relation between anaphoric expressions and discourse structure also manifests itself in another way – an anaphoric expression may affect the thematic structuring of the discourse. Overspecified anaphoric expressions signal thematic boundaries. Entities from the previous discourse become less available, and subsequent information is less likely to be integrated with previous information. In this case, it is the anaphoric expression that affects the thematic structuring of the discourse. This is the thematic structuring function of the anaphoric expression.

The present study did not address the identificational function of anaphoric expressions, but their thematic structuring function. The results give some indication of how language users deal with thematic structure in discourse. Speakers or writers may decide to change the theme while continuing to talk about the same character. Once speakers or writers have made this decision, they may want to help the listener or reader to structure the incoming input by selecting a linguistic device to indicate the theme shift. This device may take several forms. A thematic shift can consist in a change of time or place and can be indicated by a locative or temporal expression. In such cases, a pronominal expression can adequately be used at a thematic boundary. In other cases, an overspecified expression may, on its own, signal a theme shift. It closes the current theme and refocuses the referent in a new theme. And, indeed, the language comprehender uses an overspecified referential expression to structure the discourse.

Manuscript received February 1992
Revised manuscript received March 1992

## 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, 35A*, 427–440.

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

Bolinger, D. (1977). *Meaning and form*. London: Longman.

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

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

Brown, G. & Yule, G. (1983). *Discourse analysis*. Cambridge: Cambridge University Press.

Burnage, G. (1990). *Celex: A guide for users*. Nijmegen: Centre for Lexical Information.

CELEX Dutch Database (1989). Release N2.7 Computer Software. Nijmegen: Centre for Lexical Information.

Chafe, W. (1987). Cognitive constraints on information flow. In R.S. Tomlin (Ed.), *Coherence and grounding in discourse, Vol. 11: Typological studies in language*. Amsterdam: Benjamins.

Chang, F.R. (1980). Active memory processes in visual sentence comprehension: Clause effects and pronominal reference. *Memory & Cognition, 8*, 58–64.

Clancy, P.M. (1980). Referential choice in English and Japanese narrative discourse. In W.L. Chafe (Ed.), *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production, Vol. 3: Advances in discourse processes*. Norwood, N.J.: Ablex.

Cloitre, M. & Bever, T.G. (1988). Linguistic anaphors, levels of representation, and discourse. *Language and Cognitive Processes, 3*, 293–322.

Davison, A. (1984). Syntactic markedness and the definition of sentence topic. *Language, 60*, 797–846.

Dell, G.S., McKoon, G., & Ratcliff, R. (1983). The activation of antecedent information in the processing of anaphoric reference in reading. *Journal of Verbal Learning and Verbal Behavior, 22*, 121–132.

Fletcher, C.R. (1984). Markedness and topic continuity in discourse processing. *Journal of Verbal Learning and Verbal Behavior, 23*, 487–493.

Fox, B.A. (1984). Anaphora in popular written English narratives. In R.S. Tomlin (Ed.), *Coherence and grounding in discourse, Vol. 11: Typological studies in language*. Amsterdam: Benjamins.

Givón, T. (1983a). Topic continuity in discourse: An introduction. In T. Givón (Ed.), *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam: Benjamins.

Givón, T. (1983b). Topic-continuity in discourse: The functional domain of switch reference. In J. Haiman & P. Munro (Eds), *Switch reference and universal grammar, Vol. 2: Typological studies in language*. Amsterdam: Benjamins.

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

Halliday, M.A.K. & Hasan, R. (1976). *Cohesion in English*. London: Longman.

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.

Just, M.A. & Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329–354.

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

Kieras, D.E. (1978). Good and bad structure in simple paragraphs: Effects on apparent theme, reading time, and recall. *Journal of Verbal Learning and Verbal Behavior, 17*, 13–28.

Kieras, D.E. (1980). Initial mention as a signal to thematic content in technical passages. *Memory & Cognition, 8*, 345–353.

Kusters, E.D.M., Tibosch, H.J.C.M., Noordman, L.G.M., Kerkman, J.P.M., & Evers, W.J. (1989). *APLEX: Apparatus for psychological experiments*. Groningen: Pro-Gamma.

Linde, C. (1979). Focus of attention and the choice of pronouns. In T. Givón (Ed.), *Syntax and semantics, Vol. 12: Discourse and syntax*. New York: Academic Press.

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013

Lorch, R.F., Lorch, E.P., & Matthews, P.D. (1985). On-line processing of the topic structure of a text. *Journal of Memory and Language*, *24*, 350–362.

Marslen-Wilson, W., Levy, E., & Tyler, L.K. (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*. Chichester: John Wiley.

McKoon, G. & Ratcliff, R. (1980). The comprehension processes and memory structures involved in anaphoric reference. *Journal of Verbal Learning and Verbal Behavior*, *19*, 668–682.

O'Brien, E.J., Duffy, S.A., & Myers, J.L. (1986). Anaphoric inference during reading. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *12*, 346–352.

Van Dijk, T.A. & Kintsch, W.A. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

## APPENDIX

*An example of Dutch text, in four conditions, used in Experiments 3 and 4.*

Professor Gerard Janssen is een druk bezet man.
Hij is, behalve vader van een groot gezin, werkzaam aan de medische faculteit van de Universiteit van Utrecht.
Zijn huidige onderzoeksobject is massage als therapie.
Er zijn, vertelt hij, bijzonder veel verschillende massagetechnieken, en er komen ieder jaar nieuwe technieken bij.
Hij noemt de voetzoolmassage als een van de belangrijkste technieken.
Janssen heeft in het verleden een opleiding tot masseur gehad.
Ook nu is hij nog regelmatig als masseur werkzaam.
Zo houdt hij contact met de *praktijk* en worden (telkens weer interessante) ideeën voor nieuw onderzoek gecreëerd.

THEME SHIFT VERSION

*Janssen, hoogleraar geneeskunde, is vader van zeven kinderen.*
(*Hij is vader van zeven kinderen.*)
Hoewel zijn werk veel van hem vraagt, vindt hij altijd tijd voor hen.
Zijn kinderen appreciëren dit zeer.

THEME CONTINUATION VERSION

*Hij vindt dit onderzoek belangrijk.*
(*Janssen, hoogleraar geneeskunde, vindt dit onderzoek belangrijk.*)
Daarom besteedt hij veel tijd in zijn laboratorium, waar hij het onderzoek van een groot aantal jongere collega's begeleidt.

Downloaded by [Max Planck Institut Fur Psycholinguistik] at 04:36 20 March 2013