Chapter 4

The effect of the second NP: Processing embedded \(wh\)-clauses

1 Introduction

The general aim of this study is to investigate to what extent word order preferences can be influenced by the discourse-related properties of the NPs in the clause. In the previous chapter we have seen that the strength of the subject-object preference is influenced by the nature of the first NP. In this chapter, the relative contribution of the second NP is dealt with. The discourse-related properties of the second NP may be such that a subject reading for this NP, and hence, an object-subject reading for the entire clause is favored. This bias thus competes with the syntactic bias for the subject-object order.

In a majority of the experiments reported in the literature, the second NP has been a definite NP. In the experiments to be presented below, order preferences are investigated for embedded \(wh\)-clauses containing a pronoun. Pronouns and definite NPs differ in their discourse-related properties and may therefore have a different effect on word order preferences.

In this introductory section, first, some reasons will be given why embedded \(wh\)-clauses are the most suitable for the present purposes. Next, the different properties of definite NPs and pronouns, and their potential effects on order ambiguities will be discussed briefly. The Syntactic and the Discourse Hypotheses are repeated and their predictions specified. Finally, an overview of the experiments is given.

1.1 Embedded \(wh\)-clauses in Dutch

In the previous chapter we have seen that the subject-object preference is fairly strong when the first NP is a definite NP. The preference for this order may be so strong that the discourse-related properties of the second NP hardly have any effect (cf. the Lamers (1996) data). To avoid such ceiling effects, the effect of the second NP will be investigated using \(wh\)-clauses only.

Furthermore, only embedded \(wh\)-clauses will be used. An example is given in (1). Just like main clause questions, embedded \(wh\)-clauses can be ambiguous between a subject-object and object-subject reading: the \(wh\)-phrase \textit{welke dichter} in (1) can either be the subject or the object of the clause, as paraphrased in (2a) and (b), respectively.

\begin{enumerate}
\item \(wh\)-clauses in Dutch
\end{enumerate}
Ik vroeg me af welke dichter de boer had gegroet.
I wondered which poet the farmer had greeted

a. I wondered which poet had greeted the farmer.
b. I wondered which poet the farmer had greeted.

Embedded clauses are more suitable than main clauses to investigate the effect of the second NP on order preferences. First, in contrast to main clauses, embedded clauses can be syntactically disambiguated after the second NP. In embedded clauses in Dutch, the finite verb appears after the second NP, that is, in the clause final or penultimate position. Hence, if the clause is disambiguated by number information at the verb, the properties of the second NP may influence word order preferences before the point of disambiguation. In main clauses, on the other hand, the finite verb always precedes the second NP. At best, main clauses can be disambiguated at the second NP by case information. This, however, limits the choice of second NPs to pronouns, the only NPs that are case-marked in Dutch.

The second reason for using embedded clauses is that the length of the ambiguous region can easily be manipulated. This allows investigation of how word order preferences develop over time. Embedded clauses can be disambiguated at the second NP by case information, or later at the finite verb by number information. The ambiguous region can be arbitrarily lengthened by inserting adverbial expressions or other material between the second NP and the disambiguating verb. Main clauses lack this variability: as mentioned before they are either disambiguated before or at the second NP. Embedded clauses are thus more suitable than main clauses to investigate the impact of the second NP on order preferences.

1.2 The nature of the second NP

Let us now consider how order preferences in embedded wh-clauses may be affected by the nature of the second NP. As has been shown in Chapter 2, Section 4.1.2, embedded wh-clauses are structurally similar to main clause questions and relative clauses under certain linguistic assumptions. This means that for embedded wh-clauses, too, a subject-object order is favored on structural grounds. This syntactic bias for a subject-object order may be influenced by the properties of the NP following the wh-phrase. The properties of this second NP may be such that a subject reading for this second NP, and hence, an object-subject reading for the clause as a whole is favored. This may be the case when the second NP is a pronoun. As discussed in Chapter 2, Section 5.3, pronouns refer to highly salient entities in the discourse. Elements referring to salient, given entities are often encoded as the subject of the clause. This has been shown by some corpus studies (Nieuwborg, 1968; Prince, 1981, 1982): pronouns more frequently occur as a subject than as a non-subject. A pronoun thus has a strong bias for being encoded as the subject of the clause. When a pronoun follows the wh-phrase, it may favor
The effect of the second NP

an object-subject reading of the clause. This may weaken or even override the general bias for a subject-object interpretation.

The syntactic bias will not be counteracted, at least not to a large extent, if the second NP is a full definite NP. As discussed in Chapter 2, definite NPs differ from pronouns in that they need not refer to given, salient elements, but can be used to (re)introduce other elements into the discourse. Especially when they are used in isolated contexts, they will be interpreted as introducing new information into the discourse. Compared to pronouns, then, there is no strong tendency to encode these elements as the subject of the clause. As a consequence, the syntactic subject-object bias will not be counteracted when the second NP is a definite NP.

Ambiguity resolution of structurally ambiguous clauses may therefore be influenced by the nature of the second NP. If the second NP is a pronoun, the discourse-related subject bias for this pronoun competes with the structural bias for the subject-object. The net effect may be a relatively weak subject-object preference for the entire clause, or even a preference for the reverse order if the pronoun bias is strong enough. If, on the other hand, the second NP is a definite NP, the syntactic bias is not opposed and the subject-object order may still be preferred.

1.3 Predictions

Whether the nature of the second NP actually affects order preferences, and if so, when, depends on which type of information is given precedence in ambiguity resolution. The two hypotheses formulated in Chapter 2 represent two extreme views on this issue. According to the Syntactic Hypothesis, ambiguity resolution is mainly driven by structural information. Discourse-related information such as the object-subject bias introduced by a pronoun in second position hardly has any effect. If it does, it will only become apparent at a later stage. Embedded wh-clauses are therefore expected to show a subject-object preference. The nature of the second NP will not affect this preference, at least not immediately.

The Discourse Hypothesis, on the other hand, predicts an immediate influence of the type of the second NP. According to the Discourse Hypothesis, discourse-related biases are stronger than syntactic biases. If the second NP is a pronoun, an object-subject reading is immediately preferred; if, in contrast, the second NP is a full definite NP, the syntactic bias is not opposed and a subject-object reading is preferred.

The two hypotheses thus differ concerning which order is initially preferred when the second NP is a pronoun: a subject-object order (Syntactic Hypothesis) or an object-subject order (Discourse Hypothesis).
1.4 An overview of the experiments

In the following sections, two off-line and four on-line experiments are reported investigating order preferences in Dutch embedded \textit{wh}-clauses.

The first experiment in this Chapter, Experiment 2, is an off-line completion study. The aim of this experiment is to confirm that embedded questions indeed show a structural bias for a subject interpretation of the \textit{wh}-phrase when no information concerning a second NP is provided.

Next, order preferences are investigated in \textit{wh}-clauses containing a case-marked pronoun as the second NP. First, off-line preferences are investigated using a forced decision and completion task (Experiment 3). The results show that the object-subject bias introduced by the pronoun can indeed override the syntactic bias for a subject-object interpretation. Next, two on-line studies are conducted to see whether and when the discourse-related object-subject bias introduced by the pronoun affects order preferences when the clauses are processed word-by-word, and the clause is syntactically disambiguated by case marking at the pronoun. The first experiment, Experiment 4, is an explanatory study using a self-paced grammaticality judgment task. The second, Experiment 5, is a self-paced reading study serving as a replication of the judgment experiment. In both experiments, an object-subject preference is indeed attested on-line.

The aim of Experiment 6 (self-paced reading) is to see whether the same results can be obtained when the \textit{wh}-clauses are disambiguated by number information at the verb rather than by a case-marked pronoun. In addition, the length of the ambiguous region is manipulated to obtain a better view of how the order preferences develop over time.

Finally, Experiment 7 explicitly tests order preferences as a function of the type of the second NP. In this self-paced reading study \textit{wh}-clauses containing a pronoun are contrasted with ones containing a non-pronominal definite NP as second NP.

2 A subject preference for embedded \textit{wh}-phrases

2.1 Introduction

Before the impact of a second NP on order preferences can really be investigated, it must first be shown that there really is a structural preference for a subject-first order in embedded \textit{wh}-constructions. Dutch embedded \textit{wh}-clauses have never been tested in an experimental paradigm. Although previous research has shown a subject-object preference for main and relative clauses, it cannot be assumed a priori that also embedded \textit{wh}-clauses in Dutch show this preference. Therefore first a small completion study was carried out to see whether there indeed is a preference to interpret the \textit{wh}-phrase as the subject of the embedded clause when no second NP has been provided yet.
2.2 Experiment 2

2.2.1 Subjects and materials

A completion study was carried out among 100 native speakers of Dutch, mainly students at the University of Groningen. They were given seven sentence fragments and asked to complete them with the first thing that came to their minds. The crucial fragments were *Ik vroeg me af welke student...* ‘I wondered which student...’ and *Ik vroeg me af wie...* ‘I wondered who...’. The remaining five fragments also introduced embedded clauses, either declaratives, or questions, starting with a *wh*-phrase other than *wie* or *welke*. Ten different orders of presentation were used. Care was taken that the *welke* fragment preceded the *wie* fragment in half of the versions; that the two experimental fragments were always separated by at least two fillers; and that a critical fragment was never presented as the first item on the list.

2.2.2 Analysis and results

The number of subject, object and ambiguous completions for the *welke student* and *wie* fragments are given below. First, Table 4.1 displays the number of subject completions when all types of completion were taken into consideration.

<table>
<thead>
<tr>
<th>function of the <em>wh</em>-phrase</th>
<th>subject</th>
<th>object</th>
<th>ambiguous</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>welke student</td>
<td>96</td>
<td>3</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>wie</td>
<td>98</td>
<td>1</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>total</td>
<td>194 (97%)</td>
<td>4 (2%)</td>
<td>2 (1%)</td>
<td>200</td>
</tr>
</tbody>
</table>
Chapter 4

In an overwhelming majority of the cases (97% on average), the wh-phrase was taken as the subject of the clause.\(^1\) The type of wh-phrase, wie or welke student did not make a difference.

About half of the completions did not contain an object NP, however. In the experiments below, the wh-clauses always involve an object NP. It may therefore be informative to also consider the number of subject completions for transitive and ditransitive completions only. These numbers are given in Table 4.2. Excluded are intransitives such as ‘to leave’, copula constructions such as ‘is the tallest’, and expressions like *zin hebben in* (‘to feel like’) or ‘t in zijn hoofd halen om.. (‘to get it in one’s head to..’).

\[
\begin{array}{|c|c|c|c|}
\hline
\text{function of the wh-phrase} & \text{subject} & \text{object} & \text{ambiguous} & \text{total} \\
\hline
\text{welke student} & 62 (94\%) & 3 (5\%) & 1 (2\%) & 66 \\
\text{wie} & 35 (95\%) & 1 (3\%) & 1 (3\%) & 37 \\
\hline
\text{total} & 97 (94\%) & 4 (4\%) & 2 (2\%) & 103 \\
\hline
\end{array}
\]

Note: due to rounding, percentages may sum up to more than 100%.

The presence of a second NP did not affect the relative number of subject completions very much: the wh-phrase was the subject of the clause in 94\% of the cases. Although the number of transitive and ditransitives completions was smaller for the wie fragment (37) than for the welke student fragment (66), this did not affect the order preference, either: in the (di)transitive completions, the wie phrase was the subject in 94\% of the cases; the welke student phrase in 95\%.

### 2.3 Discussion

The completion data suggest that people have a clear preference to take a wh-phrase as the subject of the clause when no information concerning other NPs or

\(^1\) One case in the welke student condition, and 27 in the wie involved a copula construction containing a second NP, such as welke student / wie dat is ‘which student / who that is’. These are treated as subject-initial completions in Table 4.1 and omitted in Table 4.2.
The effect of the second NP

The verb is available. This is in accordance with the general subject-object preference reported for other clause types in Dutch (cf. Experiment 1, and the literature cited in Chapter 2). Furthermore, the data pattern with the subject-object preference for embedded wh-clauses in German reported by Meng (in prep.). Additional evidence for a general subject preference for both main clause and embedded wh-phrases comes from frequency data reported in Chapter 5.

The data thus confirm the view that there is a structural bias for a subject-initial interpretation of embedded wh-clauses. Some reservations are in order, though, as only two wh-phrases were tested. However, the wh-phrases are comparable to the ones used in the experiments below (all denote persons). It is therefore reasonable to assume that the structural preference for a subject-initial reading extends to the embedded wh-clauses used in the following experiments.

3 Order preferences with case-marked pronouns

3.1 Introduction

Experiment 2 showed a preference to interpret a wh-phrase as the subject of the clause in absence of any other information. But what happens if the wh-phrase is followed by a pronoun, as in (3)?

(3) a. Ik vroeg me af welke mannen hem hebben gezien. [SO]
    I wondered which men him-ACC have seen

b. Ik vroeg me af welke mannen hij heeft gezien. [OS]
    I wondered which men he-NOM has seen

In these sentences, the function of the wh-phrase welke mannen is disambiguated by case information at the pronoun. In (a), the pronoun is accusative. The wh-phrase therefore is the subject of the clause; in (b), the pronoun is nominative, and hence, the subject.

On the basis of the structural bias, subject-object clauses (3a) should be preferred to object-subject ones (3b). On the other hand, given its discourse properties, a pronoun is preferably encoded as the subject of the clause. From a discourse point of view, clauses such as (3b) in which the pronoun is the subject should be favored over clauses like (3a) in which it is the object.

Which bias, the syntactic or the discourse, is strongest? According to the Syntactic Hypothesis the syntactic bias is the strongest. This predicts that in a forced choice task, subject-object clauses such as (3a) are preferred to object-subject clauses such as (b). In on-line tasks, subject-object clauses will elicit shorter response times compared to object-subject clauses. According to the Discourse Hypothesis, on the other hand, the pronoun bias is the strongest. Object-subject wh-clauses in which the pronoun is the subject, such as (3b) will be preferred to clauses in which it is the object, such as (3a), and will be responded to faster.
Chapter 4

These predictions were tested in the following three experiments. Experiment 3 is an off-line experiment; Experiments 4 and 5 investigated on-line preferences.

3.2 Experiment 3

3.2.1 Introduction

Experiment 3 investigated which interpretation of an embedded wh-question, subject-object or object-subject, is preferred when the second NP is a pronoun. This was done by using a forced-choice completion task. Subjects saw sentence fragments like the ones in (4):

(4) Zij vroeg wie hem / hij in het café...

She asked who him-ACC / he-NOM in the café

Subjects were requested to choose between the accusative or nominative pronoun and complete the sentence. The rationale of the task is that when people are presented with two possible structures, they will favor the one that is easiest to process. A choice for the accusative hem (‘him’) indicates a preference for the subject-object interpretation of the clause; a choice for the nominative hij (‘he’) a preference for the object-subject interpretation.

If discourse information affects word order preferences more strongly than syntactic information, the object-subject order is easier to process than the subject-object order. The nominative form of the pronoun will then be chosen more often than the accusative form. If, on the other hand, ambiguity resolution is mainly determined by syntactic strategies, the wh-phrase is preferably interpreted as the subject of the clause. As a consequence, the accusative form of the pronoun will be chosen more often than the nominative.

Order preferences were investigated for embedded wh-clauses starting with wie ‘who’ or welke-N ‘which N’. In addition, the experiment served as a pilot study to test the influence of the verb introducing the wh-questions. Two matrix verbs were compared: vragen ‘to ask’, and zich afvragen, ‘to wonder’. The verb vragen allows an implicit direct object (‘to ask someone something’, where ‘someone’ is not expressed); zich afvragen ‘to wonder’ does not. So, with vragen, the pronoun in the embedded clause can refer to the implicit direct object of the matrix clause. The pronoun referent may thus be more ‘given’ than in the zich afvragen conditions. As there is a correlation between givenness and subjecthood (cf. Chapter 2), this may affect word order preferences of the embedded wh-clause.
3.2.2 Methods

3.2.2.1 Materials

The format of the experimental items
Thirty-two sets of sentence frames were constructed of the format given in Table 4.3.

<table>
<thead>
<tr>
<th>wh-phrase</th>
<th>matrix verb</th>
<th>example</th>
</tr>
</thead>
<tbody>
<tr>
<td>wie</td>
<td>vroeg</td>
<td>a. Zij vroeg wie HEM / HIJ in het café... she asked who him/he in the café</td>
</tr>
<tr>
<td></td>
<td>vroeg z. af</td>
<td>b. Zij vroeg zich af wie HEM / HIJ in het café... she wondered who him/he in the café</td>
</tr>
<tr>
<td>welke-N</td>
<td>vroeg</td>
<td>c. Zij vroeg welke man HEM / HIJ in het café... she asked which man him/he in the café</td>
</tr>
<tr>
<td></td>
<td>vroeg z. af</td>
<td>d. Zij vroeg zich af welke man HEM / HIJ in het café... she wondered which man him/he in the café</td>
</tr>
</tbody>
</table>

The full set of experimental items is given in Appendix 2. The four experimental conditions were derived in the following way. In conditions (a) and (b), the embedded question started with wie (‘who’). In conditions (c) and (d), the wh-phrase consisted of welke (‘which’) followed by a noun denoting a gender neutral or masculine entity. Sixteen different nouns were used, eight singular, eight plural.

In conditions (a) and (c), the main verb was vroeg or vroegen (simple past singular and plural forms of ‘to ask’) depending on the number of the main clause subject. In conditions (b) and (d), the matrix verb was the appropriate form of the reflexive verb zich afvragen (‘to wonder’).

The choice was always between hij (‘he’) and hem (‘him’), the nominative and accusative forms of the masculine singular pronoun. To prevent that the subject’s choice would be influenced by biases induced by the materials, the following precautions were taken. First, the embedded clause pronoun could never refer to the subject of the main clause: the main clause subject was either ik ‘I’, singular zij (‘she’), plural zij (‘they’) or wij (‘we’). This was to exclude any potential biases of such a coreference relation on the order preference. Second, the embedded clause fragments never contained a verb. In this way, word order would not be biased by the semantics of a particular verb. Third, sentence frames ended either in a locative or a temporal expression. This was to avoid
repetitiveness of the same frames in the wie conditions (a) and (b). Finally, the order of the choice item (hij / hem versus hem / hij) was counterbalanced as described below. A preference for one particular form could therefore not be due to its linear position in the materials.

**Item groups and subject lists**
A total of eight subject lists were created in the following way. First, the 32 sentence sets were divided into four item groups of eight sentence sets each. In each item group, the order of the choice item was hij / hem in half of the cases, and hem / hij in the other half. Next, four subject lists were derived from these item groups using a Latin Square design. Sentences from one item group always appeared in the same condition on a list. Each list contained an equal number of sentences in each condition; no list included more than one member from each set; and no list included more than one occurrence of each of the 16 welke nouns. Finally, from the four lists thus constructed, four additional lists were derived by reversing the presentation order of choice items: hij / hem was replaced by hem / hij and vice versa. In this way a total of eight subject lists was obtained.

**Fillers**
In addition to the experimental sentence frames, 48 fillers were constructed to yield a total of 80 fragments per list. Also in these filler items, the choice concerned two different pronouns within an embedded clause. Fillers were of various types. In sixteen filler sentences, the choice was between a nominative and an accusative pronoun (other than hij or hem) but only one form was grammatical. These fillers were included in order to check whether subjects were attentively completing the task. In all cases, the accusative was the grammatical option. It was not possible to construct comparable items in which only the nominative was grammatical. In order to balance the number of (correct) accusative and nominative answers, 16 additional fillers were included in which the choice was between two, equally grammatical, nominative pronouns. In the 16 remaining fillers, the choice was between the equally grammatical zij (‘she’) and hij (‘he’) or hem (‘him’) and haar (‘her’) to refer to noun phrases denoting professions such as dentist, doctor or plumber.

**The order of presentation**
Fillers and experimental items were presented in a pseudorandom order (see the blocking procedure in Experiment 1; Chapter 3, Section 2.1.1.6). The order was different for each subject.

3.2.2.2 Subjects
Sixty-four students of the University of Groningen participated. All were native speakers of Dutch. Subjects were randomly assigned to one of the eight subject lists such that each list was completed by eight subjects.
3.2.2.3 Procedure
Subjects received a booklet containing the instructions and the 80 sentence frames, with eight frames on one page. Subjects were requested to encircle the form of the pronoun which they thought was the most appropriate, and to complete the sentence. They were asked not to look back to or revise already completed items. Repetition of words used to complete the sentence was allowed. Subjects took around 25 minutes to complete the test.

3.2.2.4 Analysis
For each subject and each item, the number of *hij* responses was counted. In one case, no choice was indicated, but number agreement of the completed sentence unambiguously implied the choice of one alternative. This alternative was scored as the subject’s answer. In one other case, the number of the completed finite verb form was incompatible with the choice indicated. This item was scored as a missing datum. In three other cases (each in a different subject) no choice was indicated and verb agreement did not provide disambiguating information. These cases were treated as missing, as well.

Since the data are binary and measurements are repeated within subjects, a multilevel logistic regression was carried out using the VARCL program (Longford, 1993). The number of *hij*-responses (i.e. object-subject completions) was taken as the dependent variable; type of *wh*-phrase (*wie* or *welke*-noun), type of verb (*vroeg* or *vroeg zich af*) and the interaction of *wh*-phrase type by type of verb were explanatory factors at the level of the responses. *P*-values of the effects were obtained by dividing the estimates of the coefficients by the standard error for each effect, yielding values with a standardized normal distribution. To see whether there was a preference for either *hij* or *hem*, a model was fitted including the constant (general mean) and individual deviance only. Individual deviance was modeled by estimating the deviation of the grand mean for each subject (item). When reporting the probability of a response below, I will only report the normal probabilities derived from the logit values.

3.2.3 Results
Subjects were attentively completing their forms. This is shown by the small number of errors made on those filler trials in which only one of the two alternatives, namely the accusative form, was grammatical: in total, 0.4% errors were made on these trials.

The mean percentage of occurrence of *hij* is displayed in Table 4.4. Across the board, *hij* is preferred to *hem* in 61% of the cases, with a 95% confidence interval (means plus or minus twice the Standard Error) of 57%-65%. As the chance level of 50% falls well outside this interval, *hij* is chosen significantly more often than *hem*. 
Chapter 4

The nature of the *wh*-phrase (*wie* v. *welke*-noun) did not affect the choice preference: mean percentages of *hij* were 61% in both *wie* and *welke*-noun conditions [two-tailed tests: \( z = -1.53, p > .1 \) by subjects; \( z = -1.52, p > .1 \) by items]. The main clause verb (*vroeg* v. *vroeg zich af*) did not affect the responses either [ \( z = -1.12, p > .2 \) by subjects; \( z = -1.11, p > .2 \) by items].

The preference for the nominative *hij* was slightly stronger in embedded questions introduced by *welke*-noun than by *wie* if the matrix verb was *vroeg*. This pattern was reversed if the embedding verb was *vroeg zich af*. The interaction of type of verb and type of *wh*-phrase failed to reach significance [ \( z = 1.87; p = .06 \) by subjects; \( z = 1.86; p = .06 \) by items].

3.2.4 Discussion

In sum, off-line completions show the following results:

- Nominative pronouns are chosen significantly more often than accusative pronouns. This means that the embedded *wh*-clause is preferably assigned an object-subject reading;

- The type of *wh*-phrase (‘who’ or ‘which-N’) did not affect the choice;

- The matrix verb (‘to ask’ or ‘to wonder’) did not show any reliable effect.

The object-subject preference runs counter to what is predicted on the basis of syntactic information only. According to the Syntactic Hypothesis, the *wh*-phrase should have been favored as the subject of the clause. However, the pronoun is preferred as the subject, instead. This suggests that the discourse information introduced by the pronoun is strong enough to override the syntactic bias. These results are in accordance with the Discourse Hypothesis.
The off-line results thus suggest that *wh*-clauses containing pronoun are preferably assigned an object-subject order. This, however, does not mean that the object-subject interpretation is favored immediately when the *wh*-clause is processed word-by-word. Results from off-line studies only reflect the ultimate outcomes of on-line processes. The following experiments were therefore aimed at investigating whether the object-subject preference could also be seen on-line; and, if so, when this preference would become apparent.

### 3.3 Experiment 4

#### 3.3.1 Introduction

In order to obtain a first approximation of the time course of on-line processing preferences, a continuous grammaticality judgment task was conducted. In this task, sentences are presented word-by-word. The subject decides at each word position whether the sentence can still be continued in a grammatical way. The rationale is that response times will be longer and more judgment errors will be made when the current word cannot readily be combined with the preceding context. Word-by-word decision times and accuracy data thus reflect where processing difficulties occur, and to what extent.

In general, judgment tasks yield more robust results than tasks in which only reading is required. Since subjects have to make a judgment at each word, they are forced to really pay attention to each word they are reading. This means that effects of processing difficulties tend to appear on the first word that can cause them (cf. Holmes, Stowe and Cupples, 1989). A continuous grammaticality judgment task was considered most appropriate as a first on-line experiment, as it was not obvious whether the present experimental manipulation would be strong enough to elicit any processing difficulties. If this task failed to show any effects of processing difficulty, it is unlikely that more subtle reading tasks would.

Decision times and judgment accuracy were compared for subject-object and object-subject *wh*-clauses containing an accusative and nominative pronoun, respectively. Predictions were the following.

First, if syntactic information takes precedence, as the Syntactic Hypothesis predicts, a subject-object reading will initially be preferred. This means that people have more difficulty integrating the nominative than the accusative pronoun: the *wh*-phrase is initially assigned a subject interpretation; a subsequent nominative pronoun contradicts this analysis. Response times will therefore be longer and more errors will be made on the nominative pronoun (object-subject condition) than on the accusative pronoun (subject-object condition).

According to the Discourse Hypothesis, on the other hand, the object-subject bias introduced by the pronoun takes precedence over the syntactic bias for a subject-object reading. People have more difficulties in integrating the accusative than the nominative pronoun. This means that responses times will be
longer and more errors will be made on clauses containing an accusative pronoun (subject-object conditions) than on clauses with a nominative pronoun (object-subject conditions). The Discourse Hypothesis predicts that this effect starts immediately at the disambiguating pronoun. In contrast to the predictions of the Syntactic Hypothesis, no difficulties with object-subject clauses are expected at all.

3.3.2 Methods

3.3.2.1 Materials

The experimental conditions

The experimental conditions and the format of the experimental sentences are illustrated in Table 4.5.

Thirty-two experimental sentence sets were created. Sixteen sets contained a plural wh-phrase, and a singular pronoun; in the sixteen remaining sets, the wh-phrase was singular, and the pronoun plural.

Sentences were disambiguated by a case-marked pronoun. In the subject-object conditions cf. (1a,b) and (2a,b), the embedded clause pronoun is the accusative hen ‘them’ or hem ‘him’. In the object-subject conditions (1c,d) and (2c,d), the embedded pronoun is nominative (wij ‘we’ or hij ‘he’). As in Experiment 3, coreference of the pronoun with the matrix clause subject was excluded: wij (‘we’) or ik (‘I’) served as the matrix clause subject in the plural wh-conditions, and jullie (plural ‘you’) or jij (singular ‘you’) as subject in the singular wh-conditions. The matrix clause verb was either a (simple past) form of weten ‘to know’ or vragen ‘to ask’.

To check whether subjects successfully parsed the sentences on the basis of case information, the grammaticality of the embedded clauses was manipulated at the auxiliary. In the grammatical conditions (1a,c) and (2a,c), the auxiliary had (singular) or hadden (plural), agrees in number with the subject. In the ungrammatical conditions (1b,d) and (2b,d), the auxiliary and the embedded clause subject do not agree in number. Whether the sentence is still grammatical at the auxiliary can therefore only be decided upon correctly if the object and subject roles of the embedded clause have been assigned properly. A preference for a subject or object interpretation of the wh-phrase may therefore be reflected in the reaction times at the auxiliary and the error patterns at this word position. If a subject-object reading of the wh-clause is very strongly preferred, the disambiguating case information at the pronoun may either be neglected, or it may

---

2 It was not possible to use two plural pronouns with the same person features. Wij and hen are the only plural pronominal forms in Dutch that are unambiguous with respect to case and number.
The effect of the second NP

TABLE 4.5
Experimental conditions in Experiment 4.

1. Singular wh-phrase

subject-object

a. grammatical
   Jij vroeg welke man hen bij het schuurtje achter het huis had gezien.
   you asked which man them near the shed behind the house had-\text{SG} seen

b. ungrammatical
   Jij vroeg welke man hen bij het schuurtje achter het huis hadden
   you asked which man them near the shed behind the house had-\text{PL}.

object-subject

c. grammatical
   Jij vroeg welke man wij bij het schuurtje achter het huis hadden gezien.
   you asked which man we near the shed behind the house had-\text{PL} seen

d. ungrammatical
   Jij vroeg welke man wij bij het schuurtje achter het huis had
   you asked which man we near the shed behind the house had-\text{SG}

2. Plural wh-phrase

subject-object

a. grammatical
   Ik vroeg welke vrienden hem in het dorp aan de kust hadden opgezocht.
   I asked which friends him in the village near the beach had-\text{PL} visited

b. ungrammatical
   Ik vroeg welke vrienden hem in het dorp aan de kust had
   I asked which friends him in the village near the beach had-\text{SG}

object-subject

c. grammatical
   Ik vroeg welke vrienden hij in het dorp aan de kust had opgezocht.
   I asked which friends he in the village near the beach had-\text{SG} visited

d. ungrammatical
   Ik vroeg welke vrienden hij in het dorp aan de kust hadden
   I asked which friends he in the village near the beach had-\text{PL}

* Ungrammatical sentences were not completed beyond the point of ungrammaticality.
take a while before the initial subject-object analysis is revised into an object-subject structure in the object-subject conditions. Decision times will then be longer and more errors will be made at the auxiliary in the object-subject conditions compared to the subject-object conditions. Because of this grammaticality manipulation, only welke-N phrases were used as the first NP of the embedded clause. The other type of wh-phrase used in Experiments 2 and 3, wie (‘who’), is ambiguous with respect to number. The grammaticality manipulation is therefore not possible with this sort of wh-phrase.

Finally, in order to obtain a better view of the time course of the word order preference, the disambiguating pronoun was separated from the auxiliary by a complex prepositional phrase of six words in length. The prepositional phrases were expressions of time and place, and were not selected as arguments by the sentence final verb. To avoid interference effects of number (Bock and Miller, 1991; Bock and Cutting, 1993), the two nouns inside this phrase were always singular and denoted non-animate entities.

Plausibility rating
A paper-and-pencil rating task was carried out to determine the degree of plausibility and reversibility of the embedded proposition of 40 potential experimental sentence sets. Only grammatical, single clause, declarative versions of the experimental sentences were used. This was to avoid possibly confounding effects due to the temporal ambiguity of the wh-phrase, among other things.

To remain as close to (in)definiteness of the wh-N as possible, the wh-phrase was replaced by the NP één van de N (‘one of the N’) if the wh-phrase was singular and by een paar N (‘some N’) if it was plural. For instance, the subject-object version of set (1) in Table 4.5 became: Eén van de mannen had hen bij het schuurtje achter het huis gezien. (‘One of the men had seen them near the shed behind the house.’). The reversed version, in which the pronoun was the subject, became: Wij hadden één van de mannen bij het schuurtje achter het huis gezien. (‘We had seen one of the men near the shed behind the house’).

Two experimental lists were created using a Latin Square design, such that each list contained an 20 sentences in which the pronoun was the object, and twenty in which the pronoun was the subject; and no list contained more than one version of each sentence. Thirty fillers were pseudorandomly interspersed among the experimental items yielding a total of 70 sentences. In 16 fillers, the phrase een paar N or één van de N was used. All fillers expressed more or less implausible situations (e.g. De toerist had één van de reisleiders de Martinitoren laten zien ‘The tourist showed the Martini Tower to one of the tourist guides’; Eén van de baby’s had mij een schone luier omgedaan. ‘One of the babies had changed my nappy’). Twelve different orders of presentation were used. Forty subjects, all students at the University of Groningen and native speakers of Dutch, were requested to rate each sentence on general plausibility on a scale from 1 (‘very strange’) to 7 (‘excellent’). The plausibility score of a sentence version was determined by computing the mean score for this sentence version.
On the basis of the scores obtained, the 32 most reversible and plausible sentence sets were selected to be used in the on-line experiment. The complete set of experimental materials is provided in Appendix 3. Sentences containing a singular full NP and a plural pronoun (corresponding to the singular *wh*-phrase conditions in the on-line experiment) had a mean plausibility score of 6.18 (SD .32) for the version in which the pronoun was the object and 6.01 (SD .41) for the reversed version; sentences containing a plural full NP and a singular pronoun scored 6.03 (SD .43) for the version in which the pronoun was the object, and 6.17 (SD .26) for the reversed version. There was no significant difference between the object pronoun and the reversed versions, or between the singular and plural full NP conditions \( F \)'s < 1; n.s.. Note that the plural full NP sentences were somewhat more plausible in the version in which the pronoun was the subject (corresponding to the object-subject condition in the on-line experiment) than in the version in which the pronoun was the object (corresponding to the subject-object condition), whereas the reverse was true for the sentences containing a singular full NP. This interaction almost reached significance \( F(1,30) = 3.93; p = .057 \). I will return to the implications of this difference in the results section.

**Assignment to groups and lists**

Sentence sets were divided into four item groups of eight sentence sets each. Four sets in each item group contained a singular *wh*-phrase; four a plural one. Four subject lists were created using a Latin Square design. Within a list, items from one item group appeared in the same condition. Thus each list contained an equal number of sentences in each condition, and no list contained more than one member from each sentence set.

To avoid having all somewhat implausible sentences or all sentences with less frequent words in one condition, care was taken that the four item groups differed neither in plausibility and reversibility of the predicates, nor in the number of characters and logarithmic (word form) frequency of the open class words used (source: CELEX, Nijmegen). Furthermore, each type of matrix subject and matrix verb occurred equally often within and between item groups.

**Fillers**

The 32 experimental items in each list were pseudorandomly interspersed among 76 filler items, yielding a total of 108 trials. Thirty-four fillers were grammatical.

---

3 In addition, conditions were matched for the mean z-scores of the ratings. In this way, inter-subject variability in the use of the rating scale was controlled for. First, for each subject, the plausibility z-scores for each item were calculated on the basis of that subject’s rating of all 40 potential experimental items. Next, mean z-scores were obtained for each item and condition by averaging the z-scores across subjects. Conditions remained as closely matched as when raw plausibility ratings were used.
42 were ungrammatical. The ungrammaticality was caused either by a lack of subject-verb agreement, lack of gender agreement between a deteminer and a noun, or incorrect word order. In 18 fillers, the ungrammaticality became apparent at the sentence end; in the remaining fillers, the point of ungrammaticality was varied.

Order of presentation
The order of experimental and filler items was pseudorandomized by creating blocks of experimental sentences and fillers (cf. Experiment 1). The order of the blocks, and the order of the materials within each block was automatically scrambled for each subject. The first four sentences on each list (two grammatical and two ungrammatical fillers) were the same for each subject, however.

3.3.2.2 Subjects
Forty-eight students of the University of Groningen participated (age 18-28; 14 male; 5 left-handed). All subjects were native speakers of Dutch and had normal or corrected to normal vision. None of them had participated in Experiment 3 or in the materials rating discussed above. Subjects were paid ƒ12.50 for participation in this experiment and a memory test conducted at least one day before the online task. A bonus of ƒ2.50 was given if the subject made less than ten errors on the continuous grammaticality judgment task. This was mainly to encourage subjects to avoid false positives by automatically pressing the ‘Grammatical’-button.

Subjects were assigned to one of the four lists, such that each list was seen by twelve subjects, and the number of males and left-handers was reasonably balanced across the four lists.

3.3.2.3 Procedure
Subjects were tested individually. They were seated in front of a computer screen and a button box. Sentences were displayed one word at a time in the center of the screen. Subjects initiated the presentation of the sentence by pressing a bar on the button box. Each sentence started with four hashes presented at the center of the screen. The first word of the sentence was obtained by pressing one of two response buttons on the box. Subjects were requested to indicate at each word whether the sentence continued to be grammatically correct by pressing one of

---

4 Results from this memory span task allowed investigation of the effect of working memory on word order preferences. Data from Kaan and Stowe (1995), Mecklinger et al. (1995) and Friederici et al. (1996) suggest that word order ambiguities are processed differently depending on memory capacity. However, no differences were attested between low and high span subjects in this and the following two experiments. The factor span was therefore dropped from the analyses.
The effect of the second NP

two buttons on the button box: the right button for ‘grammatical’, the left button for ‘ungrammatical’. As soon as the response was given, either the word was replaced by the next word, or a feedback message was given. Ungrammatical sentences were only displayed up to and including the point of ungrammaticality. At the end of each grammatical sentence or ungrammatical sentence fragment, subjects received a message indicating whether their last response had been correct (prima! ‘excellent!’) or not (fout antwoord ‘wrong answer’). If at other word positions the subject erroneously indicated that the sentence was incorrect, no feedback message was given, but the sentence continued with the next word, providing implicit feedback. Punctuation and use of lower and upper case letters was normal.

Subjects first read a block of ten practice trials to familiarize with the procedure and the type of errors that occurred in the sentences. Six of the practice trials were ungrammatical (agreement errors, gender errors). Only if the subject really understood the task and could properly operate the button box was the actual experiment started.

Subjects were instructed to read and respond carefully and make as few errors as possible. As already stated, this was encouragement by putting up a bonus. Subjects took about 25 minutes to complete the task.

3.3.3 Analysis and results

Decision times and error rates were separately analyzed. One experimental item was excluded from further analyses. In this sentence the noun portiek (‘doorway’) followed the masculine/feminine determiner de. However, some of the subjects reported that according to their intuitions, portiek is neuter and hence requires het as a determiner. This item thus showed a fair number of errors and long decision times on portiek and following words.

3.3.3.1 Decision times

For word positions 1 (first determiner) through 11 (the noun before the auxiliary), reaction times slower than 3 seconds or faster than 90 ms were regarded as missing data. For word positions 12 and 13 (the auxiliary and the sentence final word), where decision times were fairly long, the upper cutoff was set to 9 seconds. For the reaction time analyses, false alarms (that is, erroneous negative responses) were treated as missing as well. Responses to all word positions following outliers or false alarms were excluded from the analysis as well. Next, for each subject, mean and standard deviation for non-missing responses were computed for each word position. Reading times longer (shorter) than the subject’s mean plus (minus) 2.5 times the standard deviation for that word position were replaced by the mean plus (minus) 2.5 times the standard deviation for that word position. These trimming and cutoff procedures affected 6% of the data.

Two analyses were conducted: one on all non-missing responses; a second excluding those sentences which were responded to incorrectly at the
auxiliary or, if available, at the sentence final verb. The results of the second analysis should be interpreted with caution: the errors on the auxiliary and verb were not distributed evenly across the conditions. This means that the variance is not equal across conditions, which increases the chance of a Type I error. In the main text below, \( p \) and \( F \)-values will be reported for the analysis on the complete data set only. I will indicate the differences with the analysis on the correct trials only.5

Analysis was done on the subject \((F1)\) as well as item means \((F2)\). SPSS/PC+ MANOVAs were carried out separately for each word position. In the analysis by subjects the within-subject factors were: order (subject-object, object-subject), grammaticality (grammatical, ungrammatical) and number of the \(wh\)-phrase (singular, plural). Subject list (4 levels) was included as a between-subjects factor. In the analysis by items, order and grammaticality were the within-item factors, and number of the \(wh\)-phrase and item group (4 levels) the between-item factors. For the sentence final word, only data for the grammatical conditions were available, as in the ungrammatical cases this word was not displayed. The factor grammaticality was therefore omitted in the analyses for this word position.

In Table 4.6 and Figure 4.1, mean decision times for word position 3 (\(welke\)) through 11 (the noun immediately preceding the auxiliary) are displayed. For clarity of presentation, the factors number and grammaticality are collapsed. Table 4.7 is an overview of the significant effects for all word positions. Means for the penultimate and sentence-final position are given in Table 4.8.

Conditions did not differ until the disambiguating pronoun. At the pronoun, the object-subject sentences were reacted to 77 ms faster than the subject-object sentences, resulting in a significant effect of order \([F1(1,44) = 16.68, \ p < .001; \ F2(1,23) = 35.63, \ p < .001]\). This suggests that sentences with a nominative pronoun and, hence, an object reading of the \(wh\)-phrase, are preferred to sentences with an accusative pronoun, that is, a subject reading of the \(wh\)-phrase.

Although the difference between the object-subject order and subject-object order was largest at the pronoun, a significant difference could also be seen at several following word positions [Word position 7: \(F1(1,44) = 5.74, \ p < .025; \ F2(1,23) = 7.65, \ p < .025\). Word position 8: \(F1(1,44) = 9.45, \ p < .01; \ F2(1,23) = 5.26, \ p < .05\). Word position 9: \(F1(1,44) = 11.01, \ p < .01; \ F2(1,23) = 7.14, \ p < .025\)].

---

5 For one subject all data at the auxiliary were either missing or incorrect in one of the ungrammatical conditions. Hence, in the analysis of the restricted set, this subject was dropped.
At the auxiliary, number information indicated whether the sentence was correct. Mean decision times and error rates for the (un)grammatical auxiliary and the following word are displayed in Table 4.8. Again, number is collapsed for reasons of exposition. Ungrammatical sentences were reacted to 570 ms more slowly than grammatical sentences \([F1(1,42) = 51.72, p < .001; F2(1,23) = 53.20, p < .001]\). This confirms the well-known fact that negative responses take longer than positive responses. Furthermore, when all word positions are taken into consideration, ungrammatical continuations were less frequent than grammatical ones. This difference in frequency might also have contributed to the longer decision times on the ungrammatical auxiliary.

The auxiliary and the sentence final verb showed no main effect of order: subject-object and object-subject clauses were responded to equally fast. This suggests that before the auxiliary, the correct order of subject and object has been established on the basis of case information provided by the pronoun. However, grammaticality decision times for this position cannot be considered separate from the error rates to be dealt with below.

At the auxiliary and verb, no interaction of order and number of the \textit{wh}-phrase was found, either. Recall that the \textit{wh}-singular and the \textit{wh}-plural conditions differed with respect to reversibility (cf. plausibility ratings, Section 3.3.2.1). Apparently, this difference did not affect the decision times.
The analysis also included the factors subject list and item group. At several word positions, these factors interacted with one or more of the factors order, number and grammaticality. The effects of subject list can be attributed to differences in the length, frequency or plausibility of the item groups, especially of the function words, which were not deliberately matched across the items groups. Item group by condition interactions were mostly due to subjects in a particular list reacting more slowly than subjects in other lists.

3.3.3.2 Error data
The corrected percentage of errors for each word position is shown in parenthesis in Tables 4.6 and 4.8. For each word position, the number of errors was divided by the number of non-missing data for that word position and condition. For example, if a subject made one error at the first word position in a particular condition, his or her error rate for that position and condition is 1 out of 8 (the number of errors that could possibly have been made in that condition) yielding an error rate of 12.5%. If on the next position this subject made another error in the same condition, the error rate becomes 1 out of 7 (14.3%) for that position and condition.
The effect of the second NP

As one can see, very few errors (all less than 1%) are made before the auxiliary.

### Table 4.7

Significant effects for decision times by subjects (F1) and items (F2) (Experiment 4).

<table>
<thead>
<tr>
<th>effect</th>
<th>word position</th>
<th>significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>order</td>
<td>5 pronoun</td>
<td>*** ***</td>
</tr>
<tr>
<td></td>
<td>6 bij ‘near’</td>
<td>.1</td>
</tr>
<tr>
<td></td>
<td>7 het ‘the’</td>
<td>* *</td>
</tr>
<tr>
<td></td>
<td>8 schuurtje ‘shed’</td>
<td>** +</td>
</tr>
<tr>
<td></td>
<td>9 achter ‘behind’</td>
<td>** *</td>
</tr>
<tr>
<td>grammaticality</td>
<td>12 had(den) ‘had’</td>
<td>*** ***</td>
</tr>
</tbody>
</table>

Order: subject-object vs object-subject order; grammaticality: grammatical versus ungrammatical. *** p < .001; ** p < .01; * p < .05. * This effect was somewhat stronger when the analysis was restricted to correct trials only [F1: p < .1; F2: p < .025]. ** This effect was weaker for correct trials only [F1: p < .01; F2: p = .230].

### Table 4.8

Mean decision times (corrected percentage of errors) for the (un)grammatical auxiliary and sentence final verb, as a function of order and grammaticality (Experiment 4).

<table>
<thead>
<tr>
<th>word position</th>
<th>subject-object</th>
<th>object-subject</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>gramm.</td>
<td>ungramm.</td>
</tr>
<tr>
<td>12 auxiliary</td>
<td>1093 (4.5)</td>
<td>1698 (19.6)</td>
</tr>
<tr>
<td>had(den)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 verb</td>
<td>1034 (0.6)</td>
<td></td>
</tr>
<tr>
<td>gezien</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

gramm.: grammatical; ungramm. ungrammatical.

As one can see, very few errors (all less than 1%) are made before the auxiliary. Most errors were made at the auxiliary, where number information determined whether the sentence was grammatical or not. In Table 4.9 the mean absolute number and corrected percentage of errors for this word position is shown for all conditions.

As the number of errors on other word positions was negligible, errors were statistically analyzed for the auxiliary only. Analysis was done on the absolute numbers of errors in each condition. As can be seen from Table 4.9 the
Chapter 4

Table 4.9
Mean absolute number (corrected percentage) of errors on the auxiliary in Experiment 4. Conditions with a plural auxiliary are underscored.

<table>
<thead>
<tr>
<th>number of wh-phrase</th>
<th>subject-object gramm.</th>
<th>subject-object ungramm.</th>
<th>object-subject gramm.</th>
<th>object-subject ungramm.</th>
</tr>
</thead>
<tbody>
<tr>
<td>singular</td>
<td>.25 (6.3)</td>
<td>.85 (23.8)</td>
<td>.08 (2.1)</td>
<td>.31 (8.5)</td>
</tr>
<tr>
<td>plural</td>
<td>.08 (2.3)</td>
<td>.58 (15.6)</td>
<td>.33 (9.2)</td>
<td>.56 (14.4)</td>
</tr>
</tbody>
</table>

The absolute number of errors patterns with the percentage of errors relative to the number of non-missing responses. Analyzing the absolute rather than the relative numbers of errors should not have substantially affected the results.

Since the error data consisted of relatively small, positive discrete numbers, a model was fitted assuming a Poisson distribution (Van Duijn, 1992, 1993, cf. Experiment 1). First a saturated model was fitted including as factors: three subject lists (or: three item groups in the analysis by items), number of the wh-phrase, order, grammaticality and interactions among these factors. To estimate the contribution of each of the factors number, order, grammaticality and of the interactions among these factors, several models were fitted, each time excluding one of the relevant factors. For each such model the difference in deviance was calculated relative to the saturated model. This difference in deviance was then tested using a χ² test with the degrees of freedom corresponding to the number of factors dropped. In the analysis by items, the four-way interactions of number, order, grammaticality and each of the item groups were not significant. These effects were then dropped from the model. The deviance of models excluding the factors number, order, grammaticality or an interaction of these three factors was evaluated relative to this reduced model.

Errors at the auxiliary show a main effect of order [difference in deviance: by subjects: 5.6, p < .025; by items: 8.6, p < .01]. More errors were made in the subject-object than in the object-subject conditions. Furthermore, the factor order interacted with number [difference in deviance by subjects: 3.9, p < .05; by items: 12.9, p < .001]. In the conditions in which the wh-phrase was singular, more errors were made in the subject-object conditions than in the object-subject condition. In the plural wh-conditions, the pattern was reversed. In addition, there was a three-way interaction of order by number and grammaticality [difference in deviance by subjects: 4.8, p < .05; by items: 13.6, p < .001]. This interaction appears to be related to the presence of plural number marking on the auxiliary. In Table 4.9, the conditions containing a plural auxiliary are underscored. In the ungrammatical conditions, more errors were made when the auxiliary was plural than when it was singular. In the grammatical conditions, in
contrast, less errors were made with plural than with singular auxiliaries. It seems that plural inflection on the verb facilitates the detection of a licit subject-verb agreement, whereas it detracts from the recognition of an illicit agreement relation. This finding is in contrast with the difficulty found for plural auxiliaries in grammatical subject-object and object-subject main clauses (reading times, Experiment 1).

In order to take a closer look at the possible effects of order on the errors, a separate analysis was done for the grammatical conditions only. The analysis by items showed significant effects of order [difference in deviance 10.7; \( p < .001 \)] and an interaction of order by number [difference in deviance 25.4; \( p < .001 \)]. No significant effects were seen in the analysis by subjects [all \( p's > .1 \)], probably due to two- and three-way interactions with subject lists.

### 3.3.4 Summary of the findings

The self-paced grammaticality judgment task showed the following results.

- At the disambiguating pronoun and a few positions downstream, response times were shorter for object-subject \( wh \)-clauses (nominative pronoun) than for subject-object \( wh \)-clauses (accusative pronoun).

- At the auxiliary, response times were longer to ungrammatical than to grammatical clauses.

- A substantial number of errors was made at the auxiliary, the point at which grammaticality was manipulated. On average more errors were made in the subject-object than in the object-subject conditions. However, order interacted heavily with grammaticality and number. This is probably related to the number of the auxiliary: fewest errors were made in grammatical conditions with a plural \( wh \)-phrase (and a plural auxiliary); and most errors in the ungrammatical conditions with a singular \( wh \)-phrase (and a plural auxiliary).

### 3.3.5 Discussion

The results of the continuous judgment task thus confirm the off-line data of Experiment 3: decision times and, to some extent, the error data show an advantage for the object-subject \( wh \)-clauses relative to the subject-object clauses. In the decision times, this effect was seen already at the disambiguating pronoun. This is what is predicted by the Discourse Hypothesis: the pronoun bias for an object-subject order has a stronger effect than the syntactic bias for the subject-object order. Furthermore, the information induced by the pronoun is made use of immediately: the advantage for the object-subject order appears as soon as the pronoun is encountered in the input.
However, the Syntactic Hypothesis cannot be rejected on the basis of the present data. Grammaticality judgments are a fairly coarse measure: due to the long decision times, responses at a particular word position may reflect both early and later stages of sentence processing. It is therefore not excluded that at the disambiguating pronoun, initially the subject-object was preferred, but that this effect was obscured by other processes occurring somewhat later. This is exactly what the Syntactic Hypothesis would predict.

Furthermore, the task itself may have introduced some confounding effects. Judging each word for grammaticality is a rather artificial way of processing sentences. In addition, the grammaticality manipulation at the auxiliary may have induced specific processing strategies. This may have influenced the results obtained.

3.4 Experiment 5

3.4.1 Introduction

Experiment 5 was conducted to see whether the immediate advantage for the object-subject clauses in Experiment 4 could be replicated using a self-paced reading task. Self-paced reading is somewhat more natural as a task than a grammaticality judgment task. Furthermore, reading times may give a better view of how order preferences develop over time. Reaction times are generally somewhat shorter than decision times, and may thus be a more genuine reflection of early stages of processing. The early stages of processing is where the two hypotheses vary: the Syntactic Hypothesis predicts an initial preference for the subject-object order; the Discourse Hypothesis predicts an immediate advantage for the object-subject clauses. A self-paced reading study may therefore help decide between the two.

3.4.2 Methods

3.4.2.1 Materials

The experimental conditions

Twenty-four sentence sets used in Experiment 4 were selected (cf. Appendix 3).\(^6\) Twelve sets contained a singular pronoun and a plural wh-phrase and matrix clause subject; the twelve remaining sets contained a plural pronoun and a

---

\(^6\) The reason for not including all 32 sentence sets of Experiment 4 was that one set had appeared to be somewhat problematic (cf. Section 3.3.3), and therefore had to be excluded. To balance the number of sets with a singular or a plural wh-phrase, seven other sets were also dropped. The only selection criteria was that on the basis of the remaining sets two item groups could be formed that were matched as closely as possible.
The effect of the second NP

singular wh-phrase and matrix clause subject. The mean plausibility scores, as obtained by the materials rating in Experiment 4, were the following. Clauses with singular wh-phrases: subject-object: 6.13 (SD .33); object-subject: 5.95 (SD .42); clauses with plural wh-phrases: subject-object: 6.17 (SD .27); object-subject: 6.17 (SD .30). The plausibility scores did not differ between the two number conditions \[ F(1,22) = 1.31, p = .264 \], or the two order conditions \[ F(1,22) = 1.34, p = .260 \], and neither showed an interaction of number and order \[ F(1,22) = 1.29, p = .269 \].

As only grammatical sentences were used in this experiment, each sentence set contained only two versions of the sentence: one in which the pronoun was accusative (subject-object order), and one in which the pronoun was nominative (object-subject order). These versions correspond to the (a) and (c) versions in Table 4.5.

Item groups and subject lists
Two item groups of twelve sets each were created such that the number of the wh-phrase, the type of matrix clause subject and the matrix clause verb were balanced across the groups. Furthermore, the two groups were matched as closely as possible on plausibility and reversibility of the predicate, and on length and logarithmic frequency of the lexical items involved. On the basis of the two item groups, two subjects lists were created according to a Latin Square design, such that no member of one set appeared more than once in each list, and each list contained an equal number of sentences in each of the two experimental conditions.

Fillers
In addition, 56 grammatical filler sentences were created to yield a total of 80 sentences. Fillers were of various types: simple main clauses declaratives, main clause questions starting with either waar ('where'), waarom ('why') or wat ('what'), sentences containing embedded declaratives, and sentences containing subject or object relative clauses. In the majority of fillers, one or more pronominal forms were used.

Comprehension questions
In order to have the subjects perform the tasks attentively and read the sentences carefully, one half of the sentences was followed by a comprehension question with two alternative answers displayed at the left and the right of the screen (corresponding to the position of the two response buttons). Twelve experimental sentences and 28 filler sentences were probed, yielding a total of 40 questions. Half of these questions (six in the experimental items and 20 in the fillers)

\footnote{In addition, conditions were matched in the mean z-scores of the ratings, cf. footnote 3).}
concerned the interpretation of the thematic relations in the preceding sentence. For instance, the experimental sentence *Ik vroeg welke architecten hem door het centrum van de stad hadden rondgeleid* (‘I asked which architects guided him through the city center’) was followed by the question: *Wie werd(en) rondgeleid? Architecten / hij* (‘Who was (were) guided? Architects / he’). In the other half of the questions lexical items were probed. For instance, the filler sentence *De opzichters lieten ons de jonge panda’s in het hok zien* (‘The supervisors showed us the young pandas in the cage’) was followed by the question: *Wat lieten ze zien? Panters / beren* (‘What did they show? Panthers / bears’). The questions in the experimental items are provided in Appendix 3. The number of questions, type of question, and position (left/right) of the correct answer was balanced across the experimental conditions and the various types of fillers.

*The order of presentation*  
Fillers were pseudorandomly interspersed among the experimental items in the manner described for Experiment 1. The order of presentation was automatically scrambled for each subject, except that each subject first saw the dummy trial *Let op: de eerste zin begint nu* (‘Attention: the first sentence is starting now.’), and two filler sentences, one of which was followed by a question. These two fillers were the same for all subjects.

### 3.4.2.2 Subjects

Forty-six students of the University of Groningen took part (14 male, age 17-27, 4 left-handed). All were native speakers of Dutch, and all had normal or corrected to normal vision. None of them had participated in either of the experiments or ratings reported above. Subjects were paid €10,- for participation in this test and a memory task conducted at least one day before the reading experiment, cf. footnote 4.

Subjects were pseudorandomly assigned to the two subject lists such that each list was read by an equal number of subjects, and the number of males and left-handers was the same for both lists.

### 3.4.2.3 Procedure

The task was a self-paced word-by-word reading task using a moving window display. The experimental procedure was largely the same as described for Experiment 1. Differences were the following. First, subjects always used their dominant hand to operate the "Go"-key. Second, not every sentence was followed by a comprehension question. Third, if it was, the question was preceded by an arrow at the left hand side of the screen; the question itself only appeared at the next button press. This was to reduce the chance that subjects would miss a question by automatically pressing the response button. Fourth, feedback concerning the correctness of the response to the comprehension questions was provided by the stimulus presentation program. After an answer was given, the question disappeared and a feedback message was displayed: *Ok* if the answer
was correct; *fout! (‘wrong!’)* if the answer was false. The next sentence could
then be initiated by pressing the bar on the response box.

Before the experiment began, subjects were given a practice block of ten
sentences. Six of these sentences were followed by a comprehension question. In
three cases, the correct answer was on the left hand side; in three other cases, the
correct answer was on the right. The actual experiment was started only when it
was clear that the subject understood the task and could operate the button box
appropriately.

Subjects were instructed to read for comprehension and answer the
questions as accurately as possible. Subjects took about 20 minutes to complete
the task.

### 3.4.2.4 Analysis

Analysis was done on both residual reading times and raw data. The procedure
was the same as described for Experiment 1, except that the linear regression
equation for each subject was estimated on the basis of that subject’s reading
times on all word positions of all experimental and filler sentences in the
experiment. Absolute and relative cutoff-procedures were the same as described
for Experiment 1. Cutoff procedures affected about 2.6% of the residual and raw
data.

On both residual and raw data, MANOVAs were conducted for each
word position separately, starting at word position 3 (*welke*). In the analysis by
subjects (*F*1), order (subject-object, object-subject) and number of the *wh*-phrase
(singular, plural) were the within subject factors; subject list (2 levels) was the
between subject factor. In the analysis by items (*F*2), order was the within item
factor; the between item factors were item group (2 levels) and number of the *wh-
phrase*. Below I will report the *F*- and *p*-values for the analyses on residual
reading times only.

### 3.4.3 Results

#### 3.4.3.1 Comprehension questions

The main purpose of including comprehension questions in the materials was to
force subjects to pay attention to what they were reading. Since the number of
questions is too small to test statistically, I will only report the mean numbers of
errors.

---

8 A quadratic regression \( y = b_0 + b_1 x + b_2 x^2 \) was fitted as well. Single group t-tests
on the coefficients showed that \( b_0 \) and the quadratic component \( b_2 \) were significantly
different from zero; the linear component \( b_1 \) was not. Hence, analysis was also done using
a quadratic regression to compute the residuals. Results however did not differ from the
results reported in the main text.
Chapter 4

The results on the questions show that subjects were reading attentively: on average, only 8% of all comprehension questions (questions concerning experimental as well as filler sentences) were responded to incorrectly. Taking only questions following experimental sentences into consideration, 13% errors (1.5 out of 12) were made. Of these experimental questions, 22% (that is, 1.3 out of 6) of the questions that probed thematic relations were responded to incorrectly. Order (subject-object, object-subject) appeared to have effect as well: on average 35% (1.04 out of 3) of the thematic questions following a sentence in the subject-object condition was answered incorrectly, versus 9% (.26 out of 3) in the object-subject condition. This suggests that subjects have more difficulty with the subject-object clauses, although of course no reliable conclusions can be drawn on the basis of such a small number of responses.

3.4.3.2 Reading times
Mean residual and raw reading times are given in Table 4.10. Figure 4.2 depicts the residual reading times for the subject-object and object-subject conditions. Table 4.11 is an overview of the effects found significant for the residual and raw data.

For word positions up to and including the disambiguating pronoun, reading times showed no differences between the conditions. At the word following the pronoun (word position 6), object-subject ordered sentences were read slightly faster (17 ms in the raw reading times) than subject-object ordered sentences. This effect of order was marginally significant \[F(1,44) = 3.96, p = .053; F(2,120) = 4.86, p < .05\].

Also at the second preposition (word position 9) and following determiner (word position 10), the object-subject clauses were read faster than the subject-object clauses: \[F(1,44) = 6.51, p < .025; F(2,120) = 6.75, p < .025; \text{word position 10: } F(1,44) = 4.53, p < .05; F(2,120) = 8.91, p < .025\]. For the noun in the second prepositional phrase (word position 11), the effect of order was significant only in the analysis by subjects \[F(1,44) = 4.77, p < .05; F(2,120) = 2.60, p = .123\].

At the auxiliary, the interaction of order and number of the \textit{wh}-phrase was significant \[F(1,44) = 26.99, p < .001; F(2,120) = 38.71, p < .001\]. This is probably due to the different form of the auxiliaries used. In Table 4.12, raw and residual reading times at the auxiliary and sentence final word are displayed as a function of the number of the \textit{wh}-phrase. The conditions in which the plural form of the auxiliary was used are underscored.

In the raw reading times, the subject-object condition was read more slowly when the auxiliary was plural than when it was singular. The same was true for the object-subject condition. However, when length is corrected for, the pattern reverses: in the residual reading times, conditions containing plural auxiliaries were responded to faster than conditions containing singular auxiliaries. This facilitating effect for plural auxiliaries was also shown at the sentence final position: here, for both residual and raw reading times, conditions containing a
The effect of the second NP

Mean residual and raw reading times in ms as a function of order for word positions 3 through 13, collapsed over number (Experiment 5).

<table>
<thead>
<tr>
<th>word position</th>
<th>residual RTs</th>
<th>raw RTs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>subject-object</td>
<td>object-subject</td>
</tr>
<tr>
<td>3  welke ‘which’</td>
<td>-22.8</td>
<td>-16.3</td>
</tr>
<tr>
<td>4  man ‘man’</td>
<td>-7.0</td>
<td>-14.4</td>
</tr>
<tr>
<td>5  pronoun</td>
<td>51.5</td>
<td>44.4</td>
</tr>
<tr>
<td>6  bij ‘near’</td>
<td>27.4</td>
<td>8.9</td>
</tr>
<tr>
<td>7  het ‘the’</td>
<td>-8.0</td>
<td>-14.0</td>
</tr>
<tr>
<td>8  schuurtje ‘shed’</td>
<td>-41.5</td>
<td>-45.6</td>
</tr>
<tr>
<td>9  achter ‘behind’</td>
<td>0.2</td>
<td>-19.1</td>
</tr>
<tr>
<td>10 het ‘the’</td>
<td>-14.2</td>
<td>-29.6</td>
</tr>
<tr>
<td>11 huis ‘house’</td>
<td>-41.9</td>
<td>-56.0</td>
</tr>
<tr>
<td>12 had(den) ‘had’</td>
<td>-44.5</td>
<td>-45.6</td>
</tr>
<tr>
<td>13 gezien ‘seen’</td>
<td>124.9</td>
<td>112.9</td>
</tr>
</tbody>
</table>

plural auxiliary were read faster. However, this effect only reached significance in the analysis by subjects [$F(1,44) = 5.23, p < .05; F(2,1,20) = 3.59, p = .073$]. No other significant effects of order and number were found at the auxiliary or sentence final verb.

Note, first, that the effect of number patterns with the error data at the auxiliary in Experiment 4 (cf. Table 4.9): in the grammatical conditions in Experiment 4, less judgment errors were made when the auxiliary was plural than when it was singular. Furthermore, the present effect of verb number is the opposite of what was found for main clauses: in Experiment 1, conditions with plural auxiliaries were reacted to more slowly than conditions with singular auxiliaries. It is as of yet unclear what causes this difference between singular and plural forms, and the direction of this difference.

As in Experiment 3, several word positions showed a significant interaction of order and/or number with the factor subject list or item group. Since the main reason to include the factors list and group was to control for the variance due to list and group differences, I will not further discuss these effects here.
Chapter 4

**TABLE 4.11**
Significant effects for residual and raw reading times by subjects (F1) and items (F2) (Experiment 5).

<table>
<thead>
<tr>
<th>effect</th>
<th>word position</th>
<th>residual RTs</th>
<th>raw RTs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>F1</td>
<td>F2</td>
</tr>
<tr>
<td>order</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 bij ‘near’</td>
<td>&lt;.1</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>9 achter ‘behind’</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>10 het ‘the’</td>
<td>+</td>
<td>**</td>
<td>&lt;.1</td>
</tr>
<tr>
<td>11 huis ‘house’</td>
<td>+</td>
<td>&lt;.1</td>
<td></td>
</tr>
<tr>
<td>order x number</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 had(den) ‘had’</td>
<td>***</td>
<td>***</td>
<td>**</td>
</tr>
<tr>
<td>13 gezien ‘seen’</td>
<td>+</td>
<td>&lt;.1</td>
<td>+</td>
</tr>
</tbody>
</table>

The word positions correspond to the ones in Table 4.10. **Order**: subject-object vs object-subject order; **number**: singular vs plural wh-phrase. *** p < .001; ** p < .01; * p < .025; + p < .05.

**TABLE 4.12**
Mean raw and residual reading times at the auxiliary and the sentence final verb as a function of number and order in Experiment 5.

<table>
<thead>
<tr>
<th>word position</th>
<th>singular wh-phrase</th>
<th>plural wh-phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>subject-object</td>
<td>object-subject</td>
</tr>
<tr>
<td>residual RTs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>auxiliary</td>
<td>-30.1</td>
<td>-58.3</td>
</tr>
<tr>
<td>verb</td>
<td>142.8</td>
<td>99.0</td>
</tr>
<tr>
<td>raw RTs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>auxiliary</td>
<td>429</td>
<td>442</td>
</tr>
<tr>
<td>verb</td>
<td>687</td>
<td>645</td>
</tr>
</tbody>
</table>
3.4.4 Summary and discussion

The self-paced reading study on embedded wh-clauses showed the following results:

- Object-subject wh-clauses were read faster than subject-object clauses. This effect was weakly significant at the word position following the disambiguating pronoun, was absent at the next two word positions, and became significant again at the fourth and fifth word after the point of disambiguation.

- The factor number had an effect at the auxiliary: reading times were faster for plural auxiliaries than for singular ones. This pattern resembles the error data of Experiment 4.

The aim of Experiment 5 was to see whether the results from the judgment task in Experiment 4 could be replicated in a task in which only reading was required. The judgment and the reading studies showed roughly the same effects: an object-
subject preference was seen around the point of disambiguation and at several word positions later in the sentence. A preference for the subject-object order was not attested in either experiment. The advantage for the object-subject order found in the grammaticality judgment task, and the early onset of this preference can therefore not entirely be attributed to task specific strategies.

The judgment and the reading data differ in some respects. First, compared to the judgment task, the object-subject preference appears to be somewhat delayed in the reading task: in the judgment task, the effect of order was seen immediately at the disambiguating pronoun; whereas in the reading times, it appears one position later. This delay can be attributed to the effects of spill-over that are commonly attested in self-paced reading studies (cf. Holmes et al., 1989). A second difference is that in the reading experiment the object-subject preference showed a more pronounced bi-phasic pattern: it disappeared at the second word after disambiguation and reappeared again later. It is not clear whether this two-phase effect is an artefact or really reflects two processing stages. I will therefore not further discuss this effect.

Turning now to the two hypotheses, we see that the predictions of the Syntactic Hypothesis are still not borne out. According to this hypothesis, initially an advantage should be seen for a subject-object interpretation. However, as in the judgment data, no signs of an initial subject-object preference were attested; rather a preference for an object-subject order was seen directly following the point of disambiguation. The data are therefore somewhat problematic for the Syntactic Hypothesis. The Discourse Hypothesis, on the other hand, is once again confirmed. According to this hypothesis the object-subject bias introduced by the pronoun should immediately affect the order preferences. This is indeed what the present data suggest, although the object-subject preference immediately after the point of disambiguation is weaker than expected.

3.5 Discussion: order preferences with case-marked pronouns

Experiments 3 through 5 show comparable results: *wh*-clauses containing a case-marked pronoun have an off- and on-line preference for the object-subject order. The on-line studies show that this preference becomes apparent at or immediately after the disambiguating pronoun. No hint of a subject-object preference is attested in any of the studies.

This pattern of results is not expected if the syntactic bias dominates, as stated by the Syntactic Hypothesis. The results from Experiment 2 suggest that there indeed is a preference for a subject-initial order for embedded *wh*-phrases before the second NP is encountered. However a preference for this order was not attested in any of the experiments in which the *wh*-phrase was followed by a pronoun. The data thus support the Discourse Hypothesis: the discourse-related bias for a pronoun in subject position, and hence, an object-subject order for the *wh*-clause, is apparently stronger than the syntactic preference for a subject-object order.
The Syntactic Hypothesis cannot be rejected on the basis of the present data, however. The absence of a subject-object preference and the advantage for the object-subject clauses could have been enhanced by the properties of the materials. First, the early object-subject preference may have been partially due to the lexical differences between the two conditions. The nominative form of the pronoun may have been easier to access than the accusative form, independently of word order or discourse biases. Such low-level processes are usually seen at the word itself (Just, Carpenter and Woolley, 1982); however it cannot be excluded that the object-subject preference found right after the point of disambiguation in Experiment 5 is nevertheless due to formal differences between the disambiguating pronouns.

Second, the type of disambiguating information may have had an effect. In the experiments above, clauses were disambiguated by case information at the second NP. Studies from German show a less robust subject-object preference when clauses are disambiguated by case rather than by number information. For instance, Meng (1995) reports a strong subject-object preference for main clause *wh*-clauses with number disambiguation, but no preference for either order when clauses are disambiguated by case marking on the second NP. Since the case-marked NPs were full, definite NPs, a discourse explanation is not available for these results. Differences between case and number disambiguation have also been reported by Friederici *et al.* (1996); see Chapter 2, Section 3.2.2 for an overview of their results. The German data suggest that an object-subject reading is harder to obtain when number rather than case information is used to disambiguate the clause. Although the reason for this difference between case and number disambiguation is not clear (but cf. Chapter 6 for some speculations), the use of case information in the present experiments may have reduced the difficulty for the syntactically less preferred object-subject order. This may account for the absence of a subject-object preference in Experiments 4 and 5.

Third, the position of the disambiguating information may have affected the order preference. The *wh*-clauses were disambiguated immediately after the *wh*-phrase. The disambiguating information may have come in before a subject-interpretation was actually assigned to the *wh*-phrase (cf. Frazier and Clifton, 1989). The disambiguating information in the object-subject conditions, then, would not have caused any processing difficulties: a subject-initial reading was not yet been established, and hence was not contradicted by the input. Note, however, that this cannot be the full explanation of the present data. Recall that the main clause questions in Experiment 1 were also disambiguated immediately after the *wh*-phrase. Yet, some processing difficulty was attested for the object-subject order relative to the subject-object in this experiment. This is not expected if the subject preference for the *wh*-phrase is not established yet before the point of disambiguation. So, although the early disambiguation may have had an effect in Experiments 4 and 5, it cannot be the only explanation for the results obtained.

Yet another potential reason for the lack of a subject-object preference at the disambiguating second NP is the use of matrix verbs such as *vragen* and
weten. These verbs may also take an (non-wh) NP as their direct object, instead of a sentential complement. If the processor is not immediately sensitive to the nature of the NP, the welke-N phrase might have initially been incorrectly analyzed as the direct object of the matrix verb. This analysis may have been revised only later when the properties of the first NP become available to the syntactic parser, or when the presence of a second NP signals that the initial analysis is not correct. Hence, it might be the case that at the moment the second NP is read, the wh-phrase has not yet been interpreted either as the subject or as the object of the embedded clause. Under these assumptions, no increase in reading times is expected for object-subject clauses compared to subject-object clauses at the disambiguating second NP.

The effect of the matrix verb may be tested, for example, by comparing wh-clauses introduced by vroeg ‘asked’ and ones introduced by wilde vragen ‘wanted to ask’. The latter construction does not allow the wh-phrase to be analyzed as the direct object of the matrix verb, as a direct object NP would always precede the infinitive. Although it remains an interesting issue for future research, the influence of the matrix verb will not be further investigated in the present work.

In sum, the absence of a subject-object preference in Experiments 4 and 5 may be due to the properties of the materials used. The Syntactic Hypothesis can therefore still not be rejected: a preference for a subject-object interpretation may be seen when a different manner or point of disambiguation is used. This is tested in the next experiment.

4 Order preferences using number disambiguation

4.1 Experiment 6

In Experiment 6, the wh-clauses were disambiguated by number information at the verb in penultimate position, rather than by case information immediately following the wh-phrase. If the quick disambiguation and the use of case disambiguation indeed obscured the syntactic bias for the subject-object order in Experiments 4 and 5, a preference for this order should be attested in the present experiment.

In addition, the potential effects of the length of the ambiguous region were investigated. Two different length conditions were used: a "short" condition in which the wh-phrase was separated from the disambiguating finite verb by two words; and a "long" condition in which seven words intervened. Comparing subject-object and object-subject clauses in these two length conditions may give a better view of how the syntactic and discourse bias develop over time. The two hypotheses make different predictions. According to the Syntactic Hypothesis, the syntactic subject-object bias dominates, at least initially. It is only later that discourse-related information gains strength and may override the syntactic bias (cf. Mitchell, Corley and Garnham, 1992; Rayner, Garrod and Perfetti, 1992).
According to this view, an interaction is expected between order and length of the ambiguous region. In the "short" conditions, a preference for a subject-object interpretation is expected; if the ambiguous region is fairly long, the reverse order is favored.

According to the Discourse Hypothesis, on the other hand, the discourse properties of the pronoun should dominate immediately. This hypothesis therefore predicts an object-subject preference for both length conditions.

4.2 Methods

4.2.1 Materials

4.2.1.1 The experimental conditions

Twenty-eight sentence sets were created of the format illustrated in Table 4.13.

<table>
<thead>
<tr>
<th>short ambiguous region</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. subject-object</td>
<td>Ik vroeg me af welke man jullie gisteren heeft gezien.</td>
</tr>
<tr>
<td></td>
<td>I wondered which man you-PL yesterday has-SG seen</td>
</tr>
<tr>
<td>b. object-subject</td>
<td>Ik vroeg me af welke man jullie gisteren hebben gezien.</td>
</tr>
<tr>
<td></td>
<td>I wondered which man you-PL yesterday have-PL seen</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>long ambiguous region</th>
</tr>
</thead>
<tbody>
<tr>
<td>c. subject-object</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>d. object-subject</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

In all sets and conditions, the wh-phrase was singular, and jullie (plural ‘you’) was used as the second NP in the embedded clause. Jullie is the only pronoun in Dutch that has the same form for the nominative and accusative case and is unambiguous with respect to number. The indirect question was introduced either by Ik vroeg me af (‘I wondered’) or by Wij vroegen ons af (‘we wondered’).
For each set, four versions were created. First, the order of subject and object was manipulated by the number of the auxiliary: the auxiliary was either the singular heeft (‘has’) yielding a subject-object order, or the plural hebben (‘have’) yielding an object-subject order. Second, the length of the ambiguous region was manipulated. In the short conditions, an adverb intervened between the pronoun jullie and the auxiliary. In the long conditions, the pronoun and the auxiliary were separated by a prepositional phrase of six words in length. These prepositional phrases were of the same kind as the ones used in the previous two experiments.

Note that the short condition could have been even shorter by having the auxiliary immediately follow the pronoun. However, an adverb was inserted for two reasons. First, the pronominal form jullie can also be used as a possessive pronoun (e.g. jullie huis ‘your house’). The adverb following the pronoun rendered this reading less likely (though not impossible, cf. Section 4.4.3). The second reason for inserting the adverb was to avoid potentially confounding syntactic differences between the two length conditions. According to some syntactic accounts (e.g. Bennis and Hoekstra 1984), object NPs that immediately precede the verb are in their functional base position; objects that are separated from the lexical verb by adverbial material are not, but instead have been syntactically moved from their base position to a position to the left of the adverbial material. This syntactic movement may potentially affect word order preferences, as will be discussed in Section 5.4.4. Since the experiment was aimed at testing the effects of linear rather than structural distance between the NPs and the point of disambiguation, the two length conditions were made as syntactically equal as possible by having adverbial material follow the second NP in both cases.

4.2.1.2 Plausibility rating
The plausibility and reversibility of the embedded proposition of the 28 experimental and twelve additional sentence sets was checked in the same way as described in Experiment 4, except that embedded instead of main clauses were used. For instance, object-subject short version of the sentence set in Table 4.13 was pretested as Ik weet dat jullie gisteren één van de mannen hebben gezien (‘I know that you saw one of the men yesterday.’). Sentences were pseudorandomly intermixed with 40 semantically more or less implausible filler items. Four subjects lists were created such that each list contained an equal number of items in each condition and no list contained more than one version of each set. Three different orders of presentation were used. Sixty subjects, 15 per list, were asked to rate the items with respect to plausibility.

On the basis of the scores thus obtained, 28 sentence sets were selected to be used in the self-paced reading task. The complete set of experimental materials is provided in Appendix 4. Versions in which jullie was the object, corresponding to the eventual subject-object conditions, had a mean plausibility score of 6.5 (SD .29) in the short conditions and 6.4 (SD .32) in the long;
Versions in which the pronoun was the subject (corresponding to the eventual object-subject conditions) scored 6.4 (SD .34) in the short conditions and 6.3 (SD .27) in the long. The versions corresponding to the eventual subject-object versions were rated somewhat more plausible than the ones corresponding to the object-subject versions \[F(1,27) = 4.51, \ p < .05\], and there was a tendency for short ambiguous conditions to be somewhat more plausible than long ambiguous conditions \[F(1,27) = 2.96, \ p = .097\]. There was no interaction of length and order, however \[F(1,27) < 1, \text{ n.s.}\].

4.2.1.3 Items groups and subject lists
The sentence sets were divided into four item groups of seven sentences sets each such that the four groups did not differ in length, (word form) frequency and logarithmic frequency of the sentence final verb and the word preceding the disambiguating auxiliary. On the basis of these four item groups, four subject lists were created according to a Latin Square design, such that no list contained more than one member from each sentence set, and each list contained an equal number of sentences in each condition. Care was taken that the conditions were matched in plausibility of the sentence versions within and across lists.

4.2.1.4 Fillers
Fifty-seven filler items were created to yield a total of 85 items per list. Fillers included main clause questions, declaratives containing a relative clause, embedded declaratives introduced by forms of *zeggen* (‘to say’), *vertellen* (‘to tell’), or *weten* (‘to know’), and embedded yes/no questions introduced by *zich afvragen* (‘to wonder.’).

4.2.1.5 Comprehension questions
To ensure that the subjects would attentively read the sentence, about one half of the sentences (16 experimental items, 26 fillers) was followed by a comprehension question and two alternative answers displayed at the left and the right of the screen (cf. Experiment 5). Half of these questions (8 in the experimental sets and 13 in the fillers) concerned the thematic relations expressed by the sentence; the other half of the questions probed a lexical item. Questions to the experimental sentences are provided in Appendix 4. The number and type of questions, and position (left/right) of the correct answer was balanced across the various types of experimental sentences and fillers.

4.2.1.6 Order of presentation
The experimental sentences and fillers were presented in a pseudorandom order. To this aim blocks were created consisting of four experimental items and eight

---

9 In addition, conditions were matched in the mean z-scores of the ratings, cf. footnote 3.
fillers (cf. Experiment 1). The order of the blocks and of the items within each block was automatically scrambled for each subject, except that each subject first saw an introductory trial (‘Attention: the first sentence starts now’), followed by a filler sentence which was the same for all subjects.

4.2.2 Subjects

Forty-eight subjects participated (13 male, 6 left handed, age 18-28). All were native speakers of Dutch and all had normal or corrected to normal vision. Most of them were students at the University of Groningen; a few attended a school for higher education (HBO). None of them had participated in any of the previous experiments or materials ratings. Subjects were paid ƒ10,- for participation in this test and a memory task conducted at least one day before the reading experiment (cf. footnote 4).

4.2.3 Procedure

The procedure was the same as described for Experiment 5.

4.2.4 Analysis

Analysis was done on the reading times for three word positions: (1) the word preceding the auxiliary; (2) the disambiguating auxiliary itself; and (3) the sentence final lexical verb.

Analysis was done on residual as well as raw reading times. Residual reading times were computed as described for Experiment 5. Again, a linear regression was used.Trimming procedures were the same as described for Experiment 1 and affected less than 3% of the residual reading time data, and 3.5% of the raw data.10

For both the residual reading times and the raw data, MANOVAs were conducted for each word position separately, starting at the word position preceding the point of disambiguation. Order (subject-object, object-subject) and length of the ambiguous region (short, long ambiguous region) were within subject (item) factors; subject list (4 levels) was the between subject factor in the subjects analysis; item group (4 levels) was the between items factor in the items analysis. In the main text below $F$- and $p$-values will be reported for residual reading times only.

---

10 For the same reasons as in Experiment 5 (cf. footnote 8), an analysis was also carried out using a quadratic regression. The results did not differ substantially from those reported in the main text.
4.3 Results

4.3.1 Comprehension questions

Results of the comprehension questions show that subjects were reading attentively: on average only three out of the 42 comprehension questions (7%) were responded to incorrectly. Most of the errors were made on comprehension questions following experimental items: on average, 9% errors were made (1.5 of the 16 questions). Considering only experimental questions regarding thematic relations, 1.2 out of 8 questions (15%) were responded to incorrectly. Errors were equally distributed across the four conditions: on average, 15% errors (.29 out of 2) were made in the short subject-object condition; 16% (.31 out of 2) in the short object-subject condition; 16% (.31 out of 2) in the long subject-object condition; and 13% (.25 out of 2) in the long object-subject condition.

4.3.2 Reading times

Mean raw and residual reading times for the disambiguating auxiliary, and the words immediately preceding and following it are given in Table 4.14. Residual reading times are plotted in Figure 4.3. Table 4.15 gives an overview of the significant effects.

At the word preceding the disambiguating auxiliary, and the auxiliary itself, only a main effect of length was found: the nouns and auxiliaries in the long condition were responded to faster than the adverbs and auxiliaries in the short condition [Aux-1: $F(1,44) = 12.06, p < .001$; $F(2,124) = 9.91, p < .01$. Aux: $F(1,44) = 5.48, p < .025$; $F(2,124) = 6.48, p < .025$]. This main effect of length cannot be due to differences in frequency of the words being compared, or differences in plausibility between the conditions. The adverbs in the short condition were even slightly more frequent than the nouns in the long condition (mean logarithmic frequency: 1.68 and 1.47 for adverbs and nouns, respectively. Source: CELEX, Nijmegen). Furthermore, the short conditions were rated somewhat more plausible than the long conditions (cf. Section 4.2.1.2). On the basis of frequency and plausibility information one would rather expect the short conditions to be responded to faster.

One potential explanation for the main effect of length is the difference in linear position of the words compared. It is a well-known fact that reading times decrease as a function of the linear position of the words in the sentence (cf. Ferreira and Henderson, 1993; Trueswell, Tanenhaus and Garney, 1994). In the short conditions, the auxiliary and the preceding word occupied the ninth and eighth position in the sentence; in the long condition, the auxiliary and the preceding word were in the 14th and 13th position, respectively. Reading times
may therefore have been shorter for the long than for the short conditions because of this difference in linear position.

Let us now turn to the effects of order and the interaction of order and length which are more interesting for our purposes. At the auxiliary, the order of the subject and object in the embedded clause was disambiguated. However,

### Table 4.14

Mean raw and residual reading times for the disambiguating auxiliary, the preceding and following word positions, as a function of order and length (Experiment 6).

<table>
<thead>
<tr>
<th>word position</th>
<th>short ambiguous region</th>
<th>long ambiguous region</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>subject-object</td>
<td>object-subject</td>
</tr>
<tr>
<td>residual reading times</td>
<td></td>
<td></td>
</tr>
<tr>
<td>adv/noun gisteren/boerderij</td>
<td>-12.3</td>
<td>-11.5</td>
</tr>
<tr>
<td>auxiliary heeft/hebben</td>
<td>13.5</td>
<td>8.9</td>
</tr>
<tr>
<td>verb gezien.</td>
<td>89.2</td>
<td>116.3</td>
</tr>
<tr>
<td>raw reading times</td>
<td>adv/noun</td>
<td>421</td>
</tr>
<tr>
<td>auxiliary</td>
<td>419</td>
<td>425</td>
</tr>
<tr>
<td>verb</td>
<td>585</td>
<td>612</td>
</tr>
</tbody>
</table>

**adv**: adverb (short condition).

### Table 4.15

Significant effects for subjects (F1) and items (F2) in Experiment 6.

<table>
<thead>
<tr>
<th>word position</th>
<th>effect</th>
<th>residual reading times</th>
<th>raw reading times</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>F1</td>
<td>F2</td>
</tr>
<tr>
<td>adv/noun</td>
<td>length</td>
<td>***</td>
<td>**</td>
</tr>
<tr>
<td>auxiliary</td>
<td>length</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>verb</td>
<td>order x length</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

**adv**: adverb (short conditions); **length**: short v. long ambiguous region; **order**: subject-object v. object-subject; *** $p < .001$; ** $p < .01$; * $p < .025$; + $p < .05$. 
neither the factor order, nor the interaction of order with length was significant [all Fs < 1]. At the sentence final verb, a main effect of order was not seen, either [all Fs < 1]. The interaction of order and length was significant, however [F1(1,44) = 6.47, p < .025; F2(1,24) = 5.87, p < .025]. The means show that in the long conditions, the subject-object sentences are responded to more slowly than the object-subject sentences. In the short conditions, the subject-object sentences were faster. Separate tests for the long and the short condition showed that the effect of order for the long condition was significant [F1(1,44) = 4.33, p < .05; F2(1,24) = 4.76, p < .05]. In the short condition, the difference between the two orders was not significant [F1(1,44) = 2.45, p = .125; F2(1,24) = 1.81, p = .191]. This suggests that there is a preference for an object-subject interpretation for the long conditions only.

4.4 Discussion

4.4.1 Summary of the results

Experiment 6 showed the following results.
Before the sentence final position, clauses with a short ambiguous region were read more slowly than clauses with a long ambiguous region. This may be due to a difference in linear position in the sentence of the words compared.

At the sentence final verb, which immediately followed the disambiguating auxiliary, an interaction of order and length was seen: in the long condition, reading times were significantly longer for the subject-object than for the object-subject condition; in the short condition, reading times were longer, although not significantly so, for the object-subject clauses.

Experiment 6 differed from Experiments 4 and 5 concerning the manner and point of disambiguation. In contrast to the previous experiment, sentences were disambiguated by number information at the finite verb instead of by a case-marked pronoun. In addition, the disambiguating information did not immediately follow the ambiguous *wh*-phrase, but was separated from the first NP by either two or seven words. The results from the previous experiments were at least partly replicated: a preference for the object-subject order was attested, at least if the ambiguous region was fairly long. This suggests that the object-subject bias introduced by the pronoun ultimately is stronger than the syntactic bias for the subject-object order, irrespective of the type of disambiguating information.

4.4.2 The two hypotheses

Now let us turn to the two hypotheses and the effect of the length of the ambiguous region. The Syntactic Hypothesis predicts an initial preference for the subject-object order. As opposed to Experiments 4 and 5, the subject-object preference could not have been obscured by the immediate disambiguation and the use of case marking. A subject-object preference was therefore expected at least in the short conditions. The Discourse Hypothesis, on the other hand, predicts a preference for the object-subject order, irrespective of the length of the ambiguous region. According to this hypothesis, the object-subject bias introduced by the pronoun immediately overrides the syntactic bias.

The results are problematic for both hypotheses. The Syntactic Hypothesis is only weakly supported. In the short condition, the subject-object clauses were responded to faster, but the effect of order was not statistically significant. So, again, no strong evidence is found in favor of the Syntactic Hypothesis.

The Discourse Hypothesis is not fully supported, either. In contrast to Experiments 4 and 5, the object-subject preference was not immediately present, but only if the ambiguous region was fairly long. The absence of an object-subject preference in the short condition is not expected if discourse-related information is made use of when available. The Discourse Hypothesis cannot be rejected.
The delayed influence of discourse-related information may be attributed to the categorial ambiguity of the pronominal form used.

### 4.4.3 The categorial ambiguity of jullie

The delayed influence of discourse-related information may be attributed to the categorial ambiguity of the pronominal form *jullie*. This form is ambiguous between a personal pronoun and a possessive pronoun (‘your’). The form *jullie* can therefore also introduce a full definite NP such as *jullie huis* (‘your house’). Let us see how this may have influenced the word order preference.

In contrast to pronouns, full definite NPs do not have a strong bias to be interpreted as the subject of the clause (cf. Chapter 2, Section 5.3). The possessive reading therefore does not counteract the syntactic bias for the subject-object order. This means that as long as the possessive reading of *jullie* is available, the syntactic subject-object bias does not meet full opposition.

The availability of the possessive reading, and hence the relative strength of the order biases changes depending on the length of the ambiguous region. In the short conditions, *jullie* was followed by an adverb. One of the reasons for inserting the adverb was to render the possessive reading less plausible; however, the adverb does not exclude this reading of *jullie*, as shown in (5):

(5) jullie gisteren geschreven brief
    'the letter you wrote yesterday’

However, when the adverbial expression is rather long, for instance if it is a six-word PP as used in the long conditions, the NP becomes fairly difficult to process:

(6) jullie bij het schuurtje achter de boerderij geparkeerde auto
    'your car parked near the shed behind the farm’

Hence, in the short conditions, the possessive reading of *jullie* may still have been activated in addition to the personal pronoun reading; in the long conditions the possessive reading may have been almost completely inhibited. In the short conditions, the syntactic subject-object bias would therefore meet less opposition from the pronoun bias than in the long conditions. The net effect is an object-subject preference for the long conditions but not for the short. In Experiments 4 and 5, the pronominal forms were not ambiguous between a personal pronoun and a possessive reading. The object-subject bias introduced by the personal pronoun may therefore have been strong enough to immediately effect reading times in these experiments. Of course, more research is needed to test this potential explanation for the delayed object-subject preference.
4.4.4 Summarizing remarks

Briefly put, the Discourse Hypothesis cannot be rejected on the basis of the present data. The apparent delay of the pronoun bias in the present experiment may be attributed to categorial ambiguity of the pronominal form used (jullie). On the other hand, the data only weakly support the Syntactic Hypothesis. Whether the pronoun bias has an immediate effect on the order preferences (as is claimed by the Discourse Hypothesis) or is only applied somewhat later (Syntactic Hypothesis) therefore remains unclear. However, the data do show that eventually, the discourse bias is stronger than the syntactic bias.

5 Pronouns and definite NPs compared

5.1 Experiment 7

The overall goal of the experiments, as has already been made clear, is to test whether word order preferences can be influenced by the properties of the second NP. Experiments 3 through 6 tested clauses in which the second NP is a pronoun. The object-subject preference obtained suggests that the resolution of word order ambiguities is indeed affected by the presence of a pronoun. The discourse-related information induced by the pronoun presumably is strong enough to ultimately override the syntactic bias for the subject-object order.

However, the hypothesis that the resolution of word order ambiguities depends on the type of second NP as stated by the Discourse Hypothesis has not explicitly been tested. The object-subject preference obtained in the previous experiment may therefore just as well have been due to factors other than the use of a pronoun. The use of an adverbial expression, for instance, may have had consequences for the structure of the clause and the order preference, cf. Section 5.4.4. In Experiment 7 embedded questions containing an ambiguous pronoun were therefore contrasted with ones containing a definite NP as a second NP. As shown in Chapter 2, Section 5.3.3, full definite NPs are not as strongly biased towards a subject interpretation as pronouns are. If the nature of the second NP has an effect on the resolution of word order ambiguities, the preferred order or strength of the order preference should differ depending on the type of the second NP: wh-clauses containing a pronoun will show an object-subject preference; wh-clauses with a definite NP in second position will show a subject-object preference, as the syntactic bias for a subject-object order is not counteracted, at least not as much as in the pronoun conditions.
5.2 Methods

5.2.1 Materials

5.2.1.1 The structure of the experimental items
Twenty-eight sentences were constructed of the form given in Table 4.16.

<table>
<thead>
<tr>
<th>Second NP = definite NP</th>
<th></th>
</tr>
</thead>
</table>
| **subject-object** | Zij vroeg welke trainer de atleten ’s avonds na de wedstrijd  
She asked which trainer the athletes in the evening after the game  
heeft getrakteerd.  
has-SG treated |
| **object-subject** | Zij vroeg welke trainer de atleten ’s avonds na de wedstrijd  
She asked which trainer the athletes in the evening after the game  
hebben getrakteerd.  
have-PL treated |

<table>
<thead>
<tr>
<th>Second NP = pronoun</th>
<th></th>
</tr>
</thead>
</table>
| **subject-object** | Zij vroeg welke trainer jullie ’s avonds na de wedstrijd  
She asked which trainer you-PL in the evening after the game  
heeft getrakteerd.  
has-SG treated |
| **object-subject** | Zij vroeg welke trainer jullie ’s avonds na de wedstrijd  
She asked which trainer you-PL in the evening after the game  
hebben getrakteerd.  
have-PL treated |

The second NP could be either a plural noun preceded by the definite determiner *de* ‘the’ (definite NP condition), or the plural pronoun *jullie* ‘you’-PL (pronoun condition). The *wh*-phrase was always singular. The *wh*-clauses were disambiguated by number information at the auxiliary. In the subject-object conditions the auxiliary was the singular *heeft* (‘has’), agreeing in number with the *wh*-phrase; in the object-subject conditions, the auxiliary was the plural *hebben* ‘have’, which agreed in number with the second NP.

As in previous experiments, the subject of the matrix clause introducing the *wh*-clause was such that coreference with the embedded clause pronoun was
excluded. Main clause verbs were the appropriate simple past forms of *vragen* 'to ask', *zich afvragen* 'to wonder' and *weten* 'to know'.

The second NP was separated from the finite verb by four words in order to inhibit the inappropriate possessive reading of *jullie*. Recall from the discussion in Section 4.4.3 that the form *jullie* is categorically ambiguous between a personal pronoun and a possessive. The longer the adverbal expression following *jullie*, the less plausible a possessive interpretation becomes. A six-word adverbal region as used in the long conditions in Experiment 6 was excluded for practical reasons, however. Up to and including the second NP, the experimental items were already longer than the ones used in Experiment 6. Inserting more than four words between the second NP and the auxiliary would have violated the requirement that each sentence fit on one line, which could have consequences for the way sentences are processed.

The four-word region consisted of an adverb of time, and a PP consisting of a preposition, a definite determiner and a singular noun denoting an inanimate entity. The PP could either be an argument of the sentence-final verb or a temporal or locative expression. The reason for including the adverb was to prevent an incorrect interpretation of the PP in the definite NP conditions. A PP following a non-pronominal NP in an embedded clause can either be interpreted as a modifier of this NP or as a modifier of the VP. This ambiguity is avoided by inserting an adverb between the NP and the PP: the PP can then only be interpreted as a VP-modifier.

### 5.2.1.2 Materials pretest

To ensure that the four experimental versions did not differ in plausibility, a paper-and-pencil rating task was carried out on 44 potential experimental sets. To avoid potentially confounding effects due to the temporal ambiguity of the wh-clause, the items appeared in their unambiguous declarative embedded equivalents. The matrix clause was omitted. For instance, the definite NP conditions in Table 4.16 were tested as *..dat de trainer de atleten 's avonds na de wedstrijd heeft getrakteerd* '..that the trainer treated the athletes in the evening after the game’ in the "subject-object" version, and as *..dat de atleten de trainer 's avonds na de wedstrijd hebben getrakteerd* '..that the athletes treated the trainer in the evening after the game’ in the version corresponding to the eventual object-subject conditions. The "subject-object" and "object-subject" versions for the pronoun conditions were derived in a similar way, except that the pronoun *jullie* replaced de plural full NP (*de atleten*).

Four subject lists were created such that each list contained an equal number of sentences in each of the four conditions, and no list contained more than one member of each sentence set. Experimental items were pseudorandomly intermixed with 45 fillers of the same structure as the experimental items. Twelve fillers contained a singular definite NP subject and a plural definite NP object; eleven a plural definite NP subject and a singular definite NP object; eleven contained a singular definite NP as the subject and *jullie* as the object; and eleven
had *jullie* as the object and a singular definite NP as the subject. Fillers differed in plausibility. An example of a rather implausible filler is ‘.that the cow milked the farmers with the machine yesterday’. A moderately plausible filler is e.g. ‘.that you sadly sawed the branch off the tree’. Four different orders of presentation were used, except that the first sentence was the same filler sentence for each subject. This filler was preceded by the preamble *Ze zeggen... ‘They say’*. Eighty subjects, twenty in each list, received a booklet containing the 89 sentence fragments. Subjects were all native speakers of Dutch and students at the University of Groningen. They were asked to indicate how common the situation expressed by each fragment was, using a scale from 1 (= very uncommon) to 7 (=very normal).

From the 44 sets thus tested, 28 sentence sets were selected to be used in the self-paced reading experiment. The complete set of materials is given in Appendix 7. Criteria for inclusion were that all four versions of the set had a mean plausibility rating of 5 or higher, and that the difference between the four versions was as small as possible. Mean plausibility rates for the 28 sets chosen were the following: Definite NP conditions: "subject-object": 6.15 (SD .40); "object-subject": 6.26 (SD .35). *Jullie* conditions: "subject-object": 6.24 (SD .42); "object-subject": 6.31 (SD 32). A two-way ANOVA showed no main effects of order $[F(1,27) = 1.66, p = .208]$, or type of plural NP (definite NP vs. *jullie*) $[F(1,27) = 1.99, p = .170]$, or an interaction of these two factors $[Fs < 1]$.

### 5.2.1.3 Item groups and subject lists
Subjects lists and item groups were constructed in the same way as in Experiment 6.

### 5.2.1.4 Fillers
Fifty-seven filler items were created to yield a total of 85 sentences on each list. All filler items consisted of a matrix clause and an embedded clause. The embedded clause could either be a declarative, yes/no question or a ‘when’, ‘why’ or ‘where’-question.

### 5.2.1.5 Comprehension questions
As in the previous two experiments, about half of the stimuli (16 experimental items, 26 fillers) were followed by a comprehension question with two alternative answers. Half of the questions (8 following experimental items, 13 following fillers) concerned thematic relations. The comprehension questions for the experimental items are given in Appendix 7.

---

11 In addition, conditions were matched in the mean z-scores of the ratings, cf. footnote 3.
5.2.1.6 The order of presentation

The order of fillers and experimental items was pseudorandomized in the way as described for Experiment 6.

5.2.2 Subjects

Sixty subjects were tested (14 male; age 18-30; 8 left-handed). All had normal or corrected to normal vision and were native speakers of Dutch. Most of the subjects were students at the University of Groningen. Two of the subjects attended a school for higher education (HBO). None of the subjects had participated in any of the last three experiments or in the material ratings.

5.2.3 Procedure

Procedure was the same as for Experiment 5.

5.2.4 Analysis

Statistical analyses were performed on raw and residual reading times on the noun preceding the disambiguating auxiliary, the auxiliary itself, and the sentence-final participle following the auxiliary. Residual reading times were computed as described for Experiment 5. Cutoff and trimming procedures were the same as described for Experiment 1 and affected 1% of the data to be analyzed for both residual and raw reading times.

For each of the three word positions, MANOVAs were conducted separately for the subjects and item means. Within-subject (item) factors were order (subject-object versus object-subject) and type of second NP (definite NP versus pronoun). In the analysis by subject, subject list (4 levels) was a between subject factor; in the analysis by items, item group (4 levels) was a between item factor. In the main text below, I will report $F$ and $p$-values for residual reading times only.

5.3 Results

5.3.1 Comprehension questions

As the number of comprehension questions was too small to conduct a statistical analysis, only the mean number of errors will be reported. On average, 4.8 out of 42 (12%) of all comprehension questions were responded to incorrectly. For questions following experimental sentences, the average number of errors was 3.1 out of 16 (19%). Most errors were made on experimental questions that concerned thematic relations: 2.6 out of 8 (32%). Performance on these questions was worst for the object-subject condition containing a definite NP: 1.13 out of 2 (57%). Error rates for the remaining conditions are the following: NP subject-object .48
out of 2 (24%); *jullie* subject-object .38 out of 2 (19%) and *jullie* object-subject .57 out of 2 (28%).

These data suggest that subjects have more difficulties with object-subject ordered clauses, especially when the clause contained a full definite NP in second position.

### 5.3.2 Reading times

Mean residual and raw reading times are given in Table 4.17. Figure 4.4 is a graph of the residual reading times. An overview of the significant effects for residual and raw reading times is given in Table 4.18.

<table>
<thead>
<tr>
<th>TABLE 4.17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean residual and raw reading times for the disambiguating auxiliary, the preceding and the following word as a function of order and type of second NP (Experiment 7).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NP2= definite NP</th>
<th>NP2= pronoun</th>
</tr>
</thead>
<tbody>
<tr>
<td>word position</td>
<td>subject-object</td>
</tr>
<tr>
<td>residual reading times</td>
<td></td>
</tr>
<tr>
<td>noun</td>
<td>-45.3</td>
</tr>
<tr>
<td><em>wedstrijd</em></td>
<td>-3.3</td>
</tr>
<tr>
<td>auxiliary</td>
<td></td>
</tr>
<tr>
<td>verb</td>
<td>185.4</td>
</tr>
<tr>
<td>raw reading times</td>
<td></td>
</tr>
<tr>
<td>noun</td>
<td>427</td>
</tr>
<tr>
<td>auxiliary</td>
<td>438</td>
</tr>
<tr>
<td>verb</td>
<td>704</td>
</tr>
</tbody>
</table>

NP2= second NP

At all three word positions clauses with a full definite NP as the second NP were read significantly slower than clauses containing the pronoun *jullie* [Aux-1: $F(1,56) = 8.62$, $p < .01$; $F(2,124) = 7.26$, $p < .025$. Aux: $F(1,56) = 16.74$, $p < .001$; $F(2,124) = 15.81$, $p < .001$. Aux+1: $F(1,56) = 9.35$, $p < .01$; $F(2,124) = 5.29$, $p < .05$]. This may be due to the fact that the full NP conditions always contained an extra content word (the noun of the second NP) compared to the pronoun conditions. Recall that half of the comprehension questions probed lexical items (cf. Section 5.2.1.5). Subjects might have tried to
memorize the content words while reading the sentence. The additional content
word in the full NP conditions may thus have introduced more processing load
compared to the pronoun conditions. This conjecture remains to be tested,
however.

At the noun preceding the auxiliary, no effects were significant except
for the one just discussed. At the disambiguating auxiliary itself, residual reading
times for object-subject clauses were shorter than for subject-object clauses
\[ F(1,56) = 5.16, p < .05; F(2,1.24) = 4.57, p < .05 \]. One should be cautious in
relating this effect to the difficulty of the subject-object order, however. Recall
that the auxiliary always was plural (hebben) in the object-subject conditions and
singular (heeft) in the subject-object ones. Experiment 5 had shown shorter
reading times for plural compared to singular auxiliaries, irrespective of the order
of subject and object. The present advantage for the object-subject condition at the
auxiliary may therefore be due to the different verb forms used.

Turning now to the sentence-final verb, mean reading times were longer
for the subject-object than for the object-subject conditions when the second NP is

\textbf{FIGURE 4.4} Residual reading times for the disambiguating auxiliary and preceding and
following word, as a function of order and type of second NP (Experiment 7).
The effect of the second NP

137

TABLE 4.18
Significant effects by subjects (F1) and items (F2) in Experiment 7.

<table>
<thead>
<tr>
<th>word position</th>
<th>effect</th>
<th>residual reading times</th>
<th>raw reading times</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>F1</td>
<td>F2</td>
</tr>
<tr>
<td>noun</td>
<td>type NP2</td>
<td>**</td>
<td>*</td>
</tr>
<tr>
<td>auxiliary</td>
<td>type NP2</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>order</td>
<td>order</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>verb</td>
<td>type NP2</td>
<td>**</td>
<td>+</td>
</tr>
</tbody>
</table>

Type NP2: definite NP v. pronoun; order: subject-object v. object-subject
***: p < .001; **: p < .01; *: p < .025; +: p < .05.

For a definite NP (41 ms in the raw reading times). When the second NP is a pronoun, the mean reading times for the two orders do not differ. The interaction of order and type was not significant [F1(1,56) = 2.45, p = .123; F2(1,24) = 1.99, p = .171]. When analysis was restricted to the definite NP conditions, the effect of order just failed to reach significance [F1(1,56) = 3.49, p = .067; F2(1,24) = 3.55, p = .072].

As in previous experiments, the analysis also included the factors item group and subject list. In the items analysis, significant interactions of item group with the factors order, type of NP, or both were found. This was probably due to the fact that the subjects on one list showed longer reading times overall than the subjects on the three other lists.

5.4 Discussion

5.4.1 Summary

In sum, Experiment 7 showed the following results:

- At the disambiguating auxiliary, the preceding, and the following word, clauses with a definite NP were read more slowly than clauses containing a pronoun as the second NP. This may be due to the additional content word in the definite NP conditions.

- At the disambiguating auxiliary, reading times for the object-subject conditions (plural auxiliary) were shorter than for the subject-object conditions (singular auxiliary). This may be due to the number properties of the auxiliaries used.
At the sentence-final verb (the word following the point of disambiguation), subject-object clauses were read faster than object-subject clauses when the second NP was a definite NP, although this effect failed to reach significance. No difference between the two order conditions was seen when the clause contained a pronoun.

The aim of this experiment was to explicitly test the effect of the second NP on order preferences. An interaction was predicted between the type of the second NP and order: an object-subject preference was expected in the pronoun conditions; a subject-object preference in the definite NP conditions. The prediction was only partly borne out: at the sentence-final verb, a weak subject-object preference was seen in the definite NP conditions, but order did not have any effect at all when the second NP was a pronoun. Although the effects are statistically weak, the different pattern found for the pronoun and the definite NP conditions suggests that the nature of the second NP indeed has an effect on the resolution of order ambiguities.

Furthermore, the results suggest that the object-subject preference found in the previous experiments is unlikely to be due to the use of adverbial material (cf. Section 5.4.4) or other properties of the experimental items. The clauses in the present experiment were syntactically similar to the clauses in the previous experiment; yet, a tendency for a subject-object preference was attested in the full NP conditions. This suggest that the preference for the reverse order found in the previous experiments is at least partly due to the nature of the second NP, and not an artefact due to the structure of the experimental items.

The tendency for a subject-object preference in the definite NP conditions is in accordance with the general subject-object preference previously found for Dutch and German (cf. Chapter 2). However, the results of the present experiment are much weaker than expected on the basis of the findings reported in the literature. Furthermore, the object-subject preference found for clauses containing pronouns in Experiments 3-6, was not replicated. Below, I will first deal with the absence of any order preference in the pronoun conditions. Next I will give some speculations as to why the subject-object preference is so weak in the definite NP conditions.

5.4.2 No preference in the pronoun conditions

The absence of any effect of order in the pronoun conditions is in apparent contrast with the object-subject preference found for jullie clauses in the long condition in Experiment 6. What could explain this difference? Note that the jullie conditions in the present experiment differ from the long conditions in Experiment 6 with respect to the length of the ambiguous region: in the present experiment, four words separated jullie from the disambiguating auxiliary; in Experiment 6, six words intervened between the pronoun and the auxiliary. The different results
The effect of the second NP could therefore be due to the categorial ambiguity of the pronominal form *jullie* and the length of the ambiguous region.

As discussed in Section 4.4.3, *jullie* is ambiguous between a personal pronoun and a possessive reading. The form *jullie* can therefore also introduce a definite NP. The discourse properties of a definite NP do not strongly compete with the syntactic bias for a subject-object order, as is suggested by the weak subject-object preference for the definite NP conditions in the present experiment. As long as the possessive reading of the pronoun remains activated, the pronoun bias for an object-subject order may be too weak to completely override the syntactic bias for the subject-object order. The four-word region in the present experiment may have been too short to deactivate the possessive reading. The pronoun and the syntactic biases may have counteracted each other, yielding a null-preference as net effect. In Experiment 6, the six-word region presumably was long enough for the pronoun bias to override the syntactic bias, yielding an object-subject preference after the disambiguating auxiliary. Further research is however needed to test this potential explanation.

5.4.3 The weak effect in the definite NP conditions

Let us now turn to the definite NP conditions. As predicted, *wh*-clauses containing a definite NP showed a (non-significant) advantage for the subject-object order. The subject-object preference for this condition is consistent with the findings of Experiments 1 and 2, and the subject-object preference generally reported for other structurally ambiguous clause types (cf. Chapter 2). However, the subject-object preference in the current experiment is weaker than expected on the basis of the previous findings. There are three factors that may have reduced the difference between the subject-object and object-subject definite NP conditions and introduced some noise in these conditions.

First, the order preferences at the sentence-final position may have been somewhat obscured by sentence wrap-up effects. Reading times at the sentence-final position are generally longer than at other positions in the sentence. This wrap-up effect (Just and Carpenter, 1980) may have introduced some noise in the data, and reduced the chance of obtaining a significant effect of order in the present experiment.

Second, the advantage for the subject-object order may have been negatively affected by spill-over from the previous word position, the auxiliary. At this position, object-subject clauses (plural auxiliaries) were read faster than subject-object clauses (singular auxiliaries). This effect may still have been present at the next word position. This may have reduced the actual advantage of the subject-object order in the definite NP condition recorded at the end of the clause.

Third, the subject-object preference for the definite NP conditions may have been weaker than expected due to the presence of adverbial material directly following the second NP. In previous experiments on embedded clauses reported
in the literature, the second NP was immediately followed by the verb. In the next section I will show how the presence of adverbial material may have reduced the processing advantage for the subject-object conditions in the present experiment.

5.4.4 Scrambling

5.4.4.1 The effect of an adverbial expression
Let us first consider the effect of the position of adverbial materials in declarative subject-object clauses. The position of adverbial material may have consequences for the syntactic structure of the clause and the discourse status of the object referent. Consider for instance the embedded subject-object declaratives in (7)

(7) a. dat de man gisteren de vrouw zag.
   \text{that the man yesterday the woman saw}\n   \text{‘that the man saw the woman yesterday’}
 b. dat de man de vrouw gisteren zag.
   \text{that the man the woman yesterday saw}

In (a), the adverb \textit{gisteren} precedes the object de vrouw; in (b), the adverb immediately follows the object NP. According to some syntactic accounts (e.g. Bennis and Hoekstra, 1994; but see Vanden Wyngaerd, 1989; Neeleman, 1994), the syntactic representation of (b) is more complex than (a). In (a), the object de vrouw is assumed to be in its base position immediately to the left of the lexical verb, cf. (8a). In (b), the object has been moved (‘scrambled’) from its base position to a position left of the adverb, leaving a trace in its base position, cf. (8b)

(8) a. dat de man gisteren de vrouw zag
 b. dat de man de vrouw gisteren \textit{t}i zag

Subject-object clauses in which the object appears to the left of adverbial material may thus be syntactically more complex than sentences in which the object appears left adjacent to the lexical verb.

In addition, the position of the object NP relative to the adverbial material, has consequences for the discourse status of the object referent: an object NP to the left of an adverbial expression generally refers to an entity that is already present in the discourse. When the clauses are uttered with a neutral intonation, that is, with main stress on the element immediately preceding the finite verb, the discourse status of the referent of the object NP de vrouw is different in (7a) and (b). The clause in (a) can be used in a context in which the
NP referent has not been mentioned very recently. In (b) on the other hand, the NP referent is assumed to have been just mentioned. Hence, when the sentence in (b) is presented in isolation, the presence of a woman in the immediate discourse must be inferred. More discourse inferences have to be made when the object is separated from the verb by adverbial material, than when the verb follows the object immediately.

5.4.4.2 Implications for order preferences
Now let us see what this implies for the subject-object and object-subject welke-clauses used in the experiment. In the experimental materials, the second NP was always followed by an adverbial expression. Let us first consider subject-object clauses, cf. (9a). In such clauses, a syntactic dependency relation must be created between the second NP, *de atleten* and its functional base position near the verb, denoted by $t_i$ in (9b). The *wh*-clause now contains two dependency relations: the one just mentioned, and the dependency between the *wh*-phrase and the functional subject position, denoted by $t_j$:

(9) a. ...welke trainer *de atleten* ’s avonds [...] heeft getrakteerd.
   ...which trainer the athletes in the evening [...] has treated
   b. *welke trainer* $t_j$ *de atleten* $t_i$ ’s avonds $t_i$ V

In addition, the object *de atleten* is interpreted as referring to given information. This implies that the existence of an NP-referent in the discourse context must be inferred. Due to the adverbial material, subject-object clauses may thus have been relatively hard to process.

Object-subject *wh*-clauses, in contrast, are hardly affected by the presence of adverbial material. Consider for instance the following examples.

(10) a. ..welke trainer *de atleten* hebben getrakteerd.
   ..which trainer the athletes have treated
   b. ..welke trainer *de atleten* ’s avonds [...] hebben getrakteerd.
   ..which trainer the athletes in the evening [...] have treated

In (10b) the second NP (the subject) is followed by an adverbial expression; (10a) does not contain any adverb. The presence of the adverb does not force an additional dependency relation between an NP and the verb. The examples in (11a) and (b) represent the structures of (10a) and (b), respectively.

---

12 In the non-scrambled cases, the object is part of the focus of the utterance (cf. Reinhart, 1995). For the consequences of scrambling on the interpretation of quantified objects, cf. e.g. De Hoop (1992).
In both (11a) and (b), the wh-phrase is related to the object position to the left of the verb. This dependency relation is the same for (a) and (b). In both (a) and (b), the second NP is in the functional subject position, which is the same position in both cases. Furthermore, the (a) and (b) examples do not differ concerning the givenness of either the wh-phrase or the definite NP. The adverbial material in object-subject clauses therefore does not trigger inferences concerning the presence of the NP-referents in the discourse model.

The presence of the adverbial material to the right of the object NP may therefore have increased processing difficulty for subject-object clauses, whereas it hardly affected the object-subject clauses. This may have reduced the structural advantage of the subject-object order over the object-subject order in wh-clauses containing a definite NP in the present experiment. The preference for the subject-object order may therefore have been weaker compared to previous experiments in which the second NP was not followed by an adverbial expression.

One could argue that the same mechanisms have played a role in the processing of wh-clauses containing a pronoun (Experiments 3-7), leading to an increased processing difficulty for the subject-object clauses in these experiments. Object pronouns generally appear to the left of adverbial materials (perhaps due to their inherent givenness). Consider for instance the subject-object clauses in (12).

The object pronoun precedes in adverb gisteren in (12a), and follows the adverb in (b). The clause in (b) is quite marked, and requires heavy stress on the pronoun. One could therefore assume that object pronouns always move from their base positions, even if the clause does not contain any adverbial expressions and the movement is not overtly visible. In the experimental subject-object conditions, then, a dependency relation may have been initiated between the object pronoun and its base position, even before the adverbial material was encountered in the input. This may have increased the processing difficulty for the subject-object clauses, starting already at the pronoun (Experiment 4).

Note, however that the advantage for the object-subject order attested for the pronoun conditions in Experiments 3-6 cannot be entirely due to the structural complexity of the subject-object clauses: the definite NP conditions in Experiment 7 are structurally similar to the pronoun conditions in the previous experiments. However, the clauses containing a definite NP still showed a weak advantage for the subject-object order. Hence, in addition to, or instead of the structural
complexity of the subject-object clauses, the discourse properties of the second NP *per se* are likely to affect the order preferences, as well.

5.4.4.3 Predictions

If scrambling indeed plays a role in determining word order, one can make the following prediction: the subject-object order will be harder to process when the second definite NP is followed by an adverbial expression than when it is not. As a consequence, the subject-object preference should be less pronounced in clauses in which the second NP is followed by an adverb than when it is not.

Data from Meng (in preparation) are suggestive in this respect. Meng conducted an end-of-sentence grammaticality decision task on subject-object and object-subject embedded *wh*-clauses in German. The second NP either was immediately followed by the disambiguating verb, or was separated from the verb by adverbial material. In both cases, more errors were made in the object-subject than in the subject-object conditions. However, more errors were made in the subject-object condition when the second NP was separated from the verb than when it was not. The presence of the adverbial material did not affect accuracy in the object-subject conditions. This confirms the prediction above that subject-object, but not object-subject clauses are harder when the second NP is followed by adverbial material. Note that the effect of the adverbial expression in the Meng experiment may have been confounded by a difference in length between the conditions. Future research should determine whether the prediction is also borne out when differences in length are controlled for, and whether Dutch and German behave alike in this respect.

6 Discussion: The effect of the second NP

The aim of the studies presented in this chapter was to investigate whether order preferences in ambiguous clauses could be influenced by the discourse-related properties of the second NP in the clause. In most of the studies on order preferences reported in the literature, the second NP was a non-pronominal definite NP. The present experiments tested the effect of using a pronoun in second position. In Chapter 2 it was shown that pronouns generally refer to given, salient entities in the discourse, which are frequently encoded as the subject of the clause. The discourse-related properties of a pronoun therefore bias towards a subject reading for this pronoun. When a pronoun is used in second position, it favors an object-subject reading for the entire clause. A full definite NP, on the other hand, does not bias, or to a lesser extent biases an object-subject reading when used as second NP. Order preferences were therefore expected to differ when the clause contained a pronoun rather than a full definite NP in the second NP position. Below, I will first summarize the results obtained. After that, I will turn to the two hypotheses.
6.1 Summary of the results

In the six experiments reported in this chapter, the effect of the second NP on order preferences was investigated using embedded *wh*-clauses. Briefly put, the results relevant to the present purposes were the following.

- When no information concerning a second NP was provided, embedded *wh*-clauses showed an (off-line) preference for a subject-first order (Experiment 2). This confirms the view that there is a general syntactically driven bias for this order.

- When the *wh*-phrase was followed by a case-marked pronoun, the object-subject order was preferred. First, in an off-line forced decision study (Experiment 3), a nominative pronoun was chosen significantly more often than an accusative pronoun. Two subsequent on-line studies showed an increase in processing difficulty for the subject-object clauses relative to the object-subject. In the continuous grammaticality judgment study (Experiment 4), this effect was visible at the disambiguating pronoun and a few later positions. In the self-paced reading study (Experiment 5), an object-subject preference was found at the word immediately following the disambiguating pronoun, although the effect was not strongly significant. A significant object-subject preference was seen rather late in the sentence.

- When the *wh*-phrase was followed by the case-ambiguous pronoun *jullie* and the sentence was disambiguated by number information at the auxiliary in penultimate position, an object-subject preference was seen when the ambiguous region was seven words in length (Experiment 6). When only one word separated the pronoun from the verb, however, subject-object clauses were read (non-significantly) faster than object-subject clauses (Experiment 6). When four words intervened between the pronoun and the verb, no preference was attested for either order (Experiment 7).

- Finally, when the *wh*-phrase was followed by a full definite NP, a weak preference was attested for the subject-object order (Experiment 7).

6.2 The two hypotheses

In Chapter 2 two hypotheses, the Syntactic and the Discourse Hypotheses, were formulated. These hypotheses represent two extreme, opposite views concerning the impact of lexically mediated, discourse-related information on order preferences. According to the Syntactic Hypothesis, precedence is given to
syntactic biases only. A subject-object preference is therefore expected, independently of the nature of the second NP. The Syntactic Hypothesis does not exclude the influence of discourse-related information, however this kind of information is predicted to affect the order preference only at later stages. The Discourse Hypothesis, on the other hand, states that if available, discourse-related biases dominate syntactic biases. In this view, discourse information may have an immediate effect on order preferences.

The results are not conclusive between the two hypotheses. The data suggest that order preferences are indeed affected by the discourse-related properties of the second NP: the object-subject bias introduced by a pronoun in second position is strong enough to even override the structural subject-object bias. This immediately falls out from the Discourse Hypothesis; the Syntactic Hypothesis does not clearly predict this effect, but is not incompatible with these facts, either.

The time course of the object-subject preference is somewhat problematic for both hypotheses, however. On the one hand, the absence of a significant subject-object preference in the pronoun conditions and the immediate object-subject preference in the case-marked pronoun conditions suggests that the pronoun bias can immediately override the syntactic subject-object bias. This is in accordance with the predictions of the Discourse Hypothesis. On the other hand, when the pronoun was ambiguous and clauses were disambiguated by number information at the verb, an object-subject preference was seen only if the ambiguous region was quite long. This suggests that the pronoun bias does not have an immediate effect, confirming the predictions of the Syntactic Hypothesis.

Neither of the two hypotheses can be rejected, either. On the one hand, the Syntactic Hypothesis can be saved by claiming that the immediacy of the discourse information in the case-marked conditions is only apparent: the use of case marking, the almost immediate disambiguation and perhaps the nature of the matrix verb used may have obscured the subject-object preference predicted by the Syntactic Hypothesis. On the other hand, the Discourse Hypothesis may be saved by claiming that the delay of the use of discourse information in the number disambiguation conditions is due to the categorial ambiguity of the pronominal form used.

A less confounded view of the time course of discourse and syntactic information would be obtained by using pronouns that are ambiguous with respect to case, unambiguous with respect to number, and do not allow a possessive reading. Unfortunately, such pronouns are not available in Dutch. What the present data do suggest is that the factual pattern of the word order preferences is determined by a number of factors in addition to the syntactic and discourse biases. I will return to this in the General Discussion (Chapter 6).