The development of Dutch connectives

Change and acquisition
as windows on form-function relations

Published by
LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6006
fax: +31 30 253 6000
e-mail: lot@let.uu.nl
http://wwwlot.let.uu.nl/

Cover illustration: photograph of train rails with sidetracks as a representation of developments in language.

ISBN 90-76864-77-2
NUR 632

The development of Dutch connectives

Change and acquisition as windows on form-function relations

De ontwikkeling van Nederlandse connectieven
Verandering en verwerving
als vensters op vorm-functierelaties
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de Rector Magnificus, Prof. dr. W.H. Gispen,
ingevolge het besluit van het College voor Promoties
in het openbaar te verdedigen
op dinsdag 24 mei 2005
des middags te 16.15 uur

door

Jacqueline Evers-Vermeul

geboren op 9 december 1973 te Purmerend

Promotoren: Prof. dr. T.J.M. Sanders
Faculteit der Letteren, Universiteit Utrecht

Prof. dr. F.P. Weerman
Faculteit der Geesteswetenschappen, Universiteit van Amsterdam

Gelukkig is de mens die wijsheid heeft gevonden,
een mens die inzicht wint.
Wijsheid levert meer op dan zilver,
geeft meer profijt dan goud,
is kostbaarder dan edelstenen.
Alles wat je ooit zou kunnen wensen
valt bij de wijsheid in het niet.

(Spreuken 3: 13-15, Nieuwe Bijbelvertaling)

Happy is the man who finds wisdom,
the man who gets understanding.
For her good profit is better than getting silver,
and her return is better than fine gold.
She is more precious than rubies.
None of the things you can desire
are to be compared to her.

(Proverbs 3: 13-15, World English Bible)

Contents

Part I – General introduction

Chapter 1

Introduction

In this chapter I introduce the topic, the main research question, and the aim of this dissertation. Furthermore I give a motivation for both the selection of the connectives under investigation and the choice of the two research areas that are at the heart of this study: language change and language acquisition. To conclude, I give an overview of the remaining chapters.

> *"Cross-linguistic comparison, developmental and diachronic research are then needed before we can claim a full understanding of the principles underlying the marking of discourse structure."* (Redeker 1990: 380)

1.1 Research question

Children aged three already use words like *and*, *because*, and *but* to indicate relations between sentences. They may say things like the constructed utterance in (1).[1]

(1) *Ik ben een beetje moe,* ***want*** *ik ben te laat gaan slapen.*
I am a bit tired, because I am too late go sleep.
'I'm a bit tired, because I went to bed too late.'

In this utterance, the coordinating connective *want* 'because' is accompanied by the typically Dutch word order of a main clause: the finite verb *ben* 'am' directly follows the subject *ik* 'I', resulting in so-called *verb second*. As the grammaticality of (2) shows, the causal relation between being tired and going to bed late can just as easily be expressed with the connective *omdat* 'because' in combination with a subordinating word order.

(2) *Ik ben een beetje moe,* ***omdat*** *ik te laat ben gaan slapen.*
I am a bit tired, because I too late am go sleep.
'I'm a bit tired, because I went to bed too late.'

Using this subordinator, the clauses can be combined in a different order: that is, instead of placing the connective clause after the clause to which it connects (as in (1) and (2)), it is equally possible to position the connective clause in front of it (as in (3)). This ordering is not possible with *want*.

[1] Throughout this book, examples without a reference are constructed. Other examples in this thesis are taken from my historical and child language connective corpora (see Chapter 4 (section 4.1 and 4.2) and Chapter 10 (section 10.4.1) respectively for a description of these corpora). After each fragment of child language, an indication of the age of the child is given in (years;months.days). For each example from the historical corpus, information on the name and date of the specific source text is mentioned. If the ordering of constituents is relevant, a word-by-word translation is given. For the majority of examples, however, I only supply free English translations.

(3) ***Omdat*** *ik te laat ben gaan slapen, ben ik een beetje moe.*
'Because I went to bed too late, I'm a bit tired.'

A last option mentioned here is illustrated in (4), where the connective *daarom* 'that's why' heads the result or consequent clause instead of the cause or antecedent clause. *Daarom* is not a coordinator or a subordinator, but an adverb, functioning as a constituent within the clause. When topicalized, it triggers inversion of the finite verb *ben* 'am' and the subject *ik* 'I', resulting in the proper verb second word order of a main clause.

(4) *Ik ben te laat gaan slapen.* ***Daarom*** *ben ik een beetje moe.*
I am too late go sleep. Therefore am I a bit tired.
'I went to bed too late. That's why I'm a bit tired.'

My discussion of these four utterances illustrates that connectives can be described at two linguistic levels. The first is the text-linguistic or discourse level, which deals with combinations of clauses and larger spans of spoken or written text.[2] In (1)-(4), basically the same textual information is expressed, namely that a causal coherence relation exists between going to bed late and being tired. Dutch language users have several connectives at their disposal for explicating this conceptual causal relation. The second level is the sentence-linguistic level, which focuses mainly on properties of individual clauses and sentences. Connectives differ in syntactic properties such as grammatical status (*want* is a coordinator, *omdat* a subordinator, and *daarom* an adverb) and linearization (with the connective clause before or after the clause to which it is linked).[3]

How do these two linguistic levels relate to each other? For example, does the choice of a subordinating or a coordinating connective (with the accompanying word orders) make any difference at the text-linguistic level? And what is the effect of presenting two causally related clauses in the reverse order? This kind of question has led to the connective research reported in this book. The main question of this dissertation is presented in (5):

(5) <u>Main research question of this thesis:</u>
What is the relationship between the text-linguistic and the sentence-linguistic properties of connectives?

For ease of reference I will use the term *conceptual properties* to refer to the text-linguistic properties of connectives; the term *syntactic properties* is reserved for connective characteristics at the sentence-linguistic level.

The main research question is inspired by the hypothesis that conceptual and syntactic status of linguistic elements are – to a certain extent – related. This hypothesis, which will be referred to as the hypothesis about *form-meaning interactions* or *form-function relations*, can be found in a variety of theoretical frameworks. For example, Lambrecht argues "there is a relationship between the form of a sentence and its function in discourse, and that grammatical form is *in part* determined by the pragmatic circumstances under which the sentence is used as a unit of information" (Lambrecht 1988: 138).

---

[2] In this thesis, I use the term *text* to refer to both spoken and written discourse.

[3] For ease of exposition, I only treat the linear order of the combined clauses as a syntactic phenomenon here. Chapter 2 will show that the linearization of the clauses can also be described at the text-linguistic level if the antecedent- or consequent-status of the clauses is taken into account.

In this thesis, I will not restrict myself to one specific theoretical framework, since theoretical frameworks radically differ in the way they approach form-function relations. In order to investigate the interaction between conceptual and syntactic properties of connectives, I want and need to benefit from insights from both functionally oriented accounts such as cognitive and functional linguistics, and formally oriented accounts like generative grammar. Despite the fact that formal and functional approaches to the study of language differ in their basic assumptions, I will start from the idea that the approaches should complement rather than contradict each other. The two approaches represent "two sides of the same coin" (Van Kemenade 1999: 999), since they concentrate on different aspects of language phenomena. In the case of connectives, formal approaches often concentrate on syntactic properties, whereas most functional studies focus on conceptual properties, while also paying attention to the linguistic context of the connective clause and its interpretation at the level of discourse as a whole. The goal of my research is to develop a more precise relationship between the two disciplines of text and sentence linguistics in the field of connectives. Therefore, I will make use of relevant notions and ideas from both traditions, as laid out in Chapter 2.

1.2 Method

In this thesis I adopt a developmental approach to the investigation of the relation between conceptual and syntactic properties of connectives. The rationale behind this approach is as follows: if conceptual property A and syntactic property B interact, and conceptual property A shows a certain development, then it is very likely that something will change in syntax B as well. Therefore, the analysis of developing connectives can shed light on the interaction between their conceptual and syntactic properties.

Developing connectives can be found in several research areas. Two of them are selected here: language change and first language acquisition.[4] The first research area focuses on diachronic developments. In line with Cuenca (1997: 4), I believe that "it is generally possible to discover some kind of correlation between the form and meaning of a certain construction at some moment of its diachronic evolution, even though from a synchronic point of view form and meaning could be understood as being arbitrarily related."[5] Thus, by making historical comparisons it should be possible to derive conclusions on the interaction between conceptual and syntactic properties of connectives. In this thesis, I investigate connective data from the 13th to the 20th century, covering a period of 800 years.

The second field that might function as a window on form-meaning mappings is the first language acquisition of connectives. The acquisition data add to the historical data, since ontogeny need not recapitulate phylogeny (cf. Traugott & Dasher 2002: 42-44). The suitability of this kind of developmental data is also stressed by Byrnes & Gelman (1991): "A developmental approach is inherently concerned with process, and so is especially suitable for detecting the relations among the ingredients of a complex system (…). Studying development is thus a powerful methodological tool and potentially more revealing about the organization of a system than any attempt to infer it from the adult end product" (p. 4). Thus, we may safely consider child language data as a source of insights that are not transparent in

[4] Other areas are second language acquisition, and the rise of pidgins and creoles.

[5] Ehlich (1994) – who works in a different framework – also stresses the importance of investigating the diachronic development of words in order to gain more insight into their present-day characterization. He introduces the term 'functional etymology' for this type of research.

adult discourse (cf. Braunwald 1997: 121-122). In this study I examine longitudinal connective data from very young children with ages ranging from 1;6 to 5;6.

To conclude, I investigate the main question of this thesis by looking at four types of connective development. These types are shown in Table 1.1.

Table 1.1. Four types of development studied in this thesis

	Connective acquisition	**Connective change**
Conceptual	acquisition of a meaning/function	change in meaning/function
Syntactic	acquisition of a form/syntax	change in form/syntax

To be more precise, I focus on the historical change and the acquisition of several Dutch connectives in such a way that the following subquestions can be answered.

(6) Subquestions in this thesis:
   a. How can the conceptual development of the connectives be described?
   b. How can the syntactic development of the connectives be described?
   c. How can the conceptual and syntactic development of the connectives be related?

By answering these questions in both research areas, I investigate the interdependency between form and function of connectives and the clauses they are contained in.

1.3 Connective selection

For each of the two research areas at the heart of this study, several Dutch connectives have been selected. For other languages, and especially for English, several studies are available – both on connective acquisition (see Diessel 2004 for an overview) and connective change (see especially the work of Traugott, the trendsetter in this field: e.g. Traugott & König 1991; Traugott & Dasher 2002). However, for Dutch, relatively little is known about the two types of development and the interaction between conceptual and syntactic properties of connectives. The acquisition of Dutch connectives is especially a relatively virgin territory. Concerning the diachronic development, several general studies are available (see, among others, Bouman 1918 and Heersche 1991), but the more detailed studies into the history of individual connectives (cf. Burridge 1993 and Van Megen 2002 on *want* 'because') focus mainly on syntactic properties of these connectives and the clauses in which they appear.[6] Note that my focus on Dutch connectives does not mean that other languages are totally disregarded; where possible, a parallel will be drawn with relevant connective data from other languages.

In the following subsections I give a motivation for the selection made from the repertoire of Dutch connectives. Section 1.3.1 presents an account of the selection of the four causal connectives that are investigated in both developmental areas. Section 1.3.2 and 1.3.3 respectively introduce the remainder of the connective selection for the diachronic research and the acquisition research.

1.3.1 Connectives selected for both the diachronic and the acquisition research

As the core object of investigation, four Dutch causal connectives have been selected for both the historical and the acquisition analysis: *daarom* 'that's why', *dus* 'so', *omdat* 'because',

[6] An exception is the work by Van Es (1954, 1955), who investigated several temporal connectives thoroughly, both at the syntactic and the conceptual level.

and *want* 'because'. The selection of these causal connectives can be motivated as follows. First of all, these connectives represent different syntactic properties, which have been partially illustrated by the examples in (1)-(4). Table 1.2 gives an overview of these modern Dutch syntactic characterizations.

Table 1.2. Syntactic characterization of the four causal connectives

Connective	**Translation**[7]	**Categorical status**	**Positioning**
daarom	that's why	adverb	in the 2nd clause
dus	so	adverb / coordinator	in front of or in the 2nd clause
omdat	because	subordinator	in front of the 1st or 2nd clause
want	because	coordinator	in front of the 2nd clause

As mentioned earlier, *daarom* can be characterized as an adverb, *omdat* as a subordinator, and *want* as a coordinator. The connective *dus* can function both as an adverb (triggering inversion) and as a coordinator, as can be inferred from the position of the underlined finite verb in examples (7) and (8) respectively.

(7) *Iedereen krijgt een hele stapel vouwblaadjes,* ***dus*** *<u>heb</u> je er een heleboel van.*
'Everybody receives a whole pile of folding leaves, so you have many of them.'

(8) *Jij hebt de puzzel uit de kast gehaald,* ***dus*** *jij <u>moet</u> hem er weer in doen.*
'You took the puzzle out of the cupboard, so you must put him back again.'

The four causals also differ with respect to their positioning. The adverbials *daarom* and *dus* are positioned within the second of the two combined clauses (see (4) and (7) respectively), functioning as a constituent that is topicalized. The coordinators *want* and *dus* should be placed in front of the second clause (see the examples in (1) and (8)), and the subordinator *omdat* can be placed either in front of the second or the first of the combined clauses (see (2) and (3) respectively).

A second reason for selecting these four causals is that the modern-Dutch conceptual characterization of these connectives is relatively well documented. In contrast to connectives expressing other types of relations, Dutch causal connectives have been studied extensively, especially in the last decade (cf. Van Belle 1989; Degand 1996, 2001; Lagerwerf 1998; Oversteegen 1997; Pander Maat & Degand 2001; Pander Maat & Sanders 1995, 1996, 2000, 2001; Pit 2003; Pit, Pander Maat & Sanders 1997; Stukker, Sanders & Verhagen 1999). These studies show that the connectives at hand do not only differ in their syntactic characteristics, but also in the precise nature of the causal relations they can or cannot express. The most important distinction between the four connectives is that *daarom* and *dus* mark antecedent-consequent or forward causal relations, whereas *omdat* and *want* explicate consequent-antecedent or backward causal relations (see Chapter 2 for a more detailed description of these and other conceptual and syntactic distinctions between the four causals).

A third motive for choosing four connectives within one conceptual dimension, namely that of causality, is that it is to be expected that interactions between conceptual and syntactic properties will be very subtle. Given the scarce availability of theories on specific interactions

[7] Table 1.2 only introduces the most general translations of the connectives at hand, which are the ones that will be used throughout this thesis. *Daarom* can furthermore be translated as 'for that reason', *dus* equates 'thus' as well, and *want* also captures the meanings of English 'since' and 'for' (although it does not have the same stylistic formality).

(but see Degand 2001 for an exception), it is likely that the interactions will show up at more fine-grained levels of analysis, such as the more detailed distinctions within the function of causal connective that will be discussed in Chapter 2.

A final consideration for selecting *want*, *omdat*, *dus* and *daarom* is that they are the most frequent causal connectives in present-day adult Dutch (cf. Uit den Boogaart 1975; Degand 1996: 87). This meets an extra criterion that is imposed upon the connective selection because I also investigate the acquisition data of very young children. Since these children are only beginning to use connectives, it is not likely that they will produce highly formal, register-specific connectives, or other connectives that are relatively infrequent in the language of adults. Therefore, the selected connectives should be relatively frequent, as causal connectives are.

1.3.2 Additional connectives selected for the acquisition research

The process of connective acquisition is different from diachronic connective developments in that young children still have to acquire all of the basic notions needed for a proper connective use. This makes it interesting to adopt a somewhat broader view in the selection of connectives here. In order to compare the development of the causals to connectives expressing other basic notions, three additional connectives have been selected for the acquisition analysis: *toen* 'then/when', *maar* 'but', and *en* 'and'. These connectives meet the criterion of high frequency, since they are among the most regularly used connectives, both in adult and in child Dutch. *Toen* marks past temporal sequence or overlap and it can function both as an adverb (see (9)) and as a subordinator (see (10)).

(9) *Ze sliep en* ***toen*** *werd ze wakker.*
'She slept and then she woke up.'

(10) ***Toen*** *ik naar de kapper moest, had ik ook een beetje pijn in m'n buik.*
'When I was at the hairdresser's, I had some pain in my belly too.'

As the utterance in (11) illustrates, the Dutch connective *maar* functions just like its English equivalent *but*: it is a coordinator expressing contrastive relations.

(11) *'k Wou bij oma een molen maken,* ***maar*** *dat kon ik niet.*
'I wanted to make a mill at grandma's, but I couldn't.'

The connective *en* 'and' has been added to the acquisition selection because it is the first and most frequent connective in child language and because it can be used to express temporal or causal relations – although its primary meaning is additive. The utterances in (12), (13) and (14) show examples in which *en* can be interpreted as marking an additive, temporal and causal relation respectively.

(12) *Nathan houdt van tennissen* ***en*** *hij houdt van voetballen.*
'Nathan likes to play tennis and he likes to play soccer.'

(13) *Mirjam werd vroeg wakker* ***en*** *ging naar het bed van haar ouders.*
'Mirjam woke up early and went to the bed of her parents.'

(14) *Bob hoorde dat het een leuk boek was* ***en*** *heeft het gelezen.*
'Bob heard it was a nice book and read it.'

The complete selection of Dutch connectives for the acquisition research is presented in Table 1.3. Taken together, these connectives cover a range of syntactic properties (in terms of categorical status and positioning) as well as different conceptual notions (additive, temporal, contrastive and causal, including more fine-grained distinctions within the causal domain), allowing me to investigate the relation between form and function of connectives.

Table 1.3. Connective selection for the acquisition study

Connective	**Conceptual classification**	**Syntactic classification**
daarom	causal	adverb
dus	causal	adverb, coordinator
omdat	causal	subordinator
want	causal	coordinator
en	additive	coordinator
maar	contrastive	coordinator
toen	temporal	adverb, subordinator

1.4 Preview

When writing this book, I kept in mind two kinds of readers: readers who want to work through the whole dissertation and readers who are only interested in certain parts of this book (e.g. the chapters on acquisition or diachronic change). Because of this, and in order to prevent readers and myself from 'mental juggling', I divided my thesis into relatively short chapters, which can be read almost entirely independently as well.

This thesis consists of four parts. Part I (Chapters 1-3) introduces the necessary theoretical background. Chapter 2 presents definitions of the items I analyze as connectives. This chapter also treats the conceptual and syntactic primitives that are generally used to characterize and classify the selected connectives in more detail. Chapter 3 shows the plausibility of interactions between the conceptual and syntactic properties of connectives, and introduces several hypotheses on form-function relations concerning connectives.

Part II (Chapter 4-8) follows a top-down approach. Four chapters (Chapters 5-8) provide a first test of particular form-function hypotheses put forward in Chapter 3, by dealing with the diachronic developments of the selected causal connectives. This part starts with a chapter containing several methodological considerations (Chapter 4). Chapter 5 and 6 focus on the diachronic development of the complementizers *want* and *omdat*; the former chapter presents a first test of a hypothesis on the interaction between word order and the hierarchical structure of discourse segments, the latter investigates the relation between word order and domains of use. Chapter 7 discusses the diachronic development of the connectives *dus* and *daarom*, thus providing data to test the hypothesis on the interaction between positioning and conceptual properties of these connectives. With the diachronic data mentioned in Chapters 5 to 7, it is relatively easy to gain insight into a related discussion: the occurrence of subjectification in connective development and its relation to syntactic changes (see Chapter 8).

Part III (Chapters 9-13) treats the acquisition of connectives. This part of the thesis follows a bottom-up approach. Given the scarce availability of detailed analyses of connective acquisition by young Dutch children, I will first discuss my findings concerning this acquisition process in detail (see Chapters 10-13). Chapter 10 focuses on the very beginning of connective acquisition; it introduces acquisition orders based on first emergence. Chapter 11 treats the overall developments of the connectives discussed in Chapter 10, showing quantitative analyses and developmental curves of these connectives. Chapter 12 also deals

with the overall developments of connectives, but the focus here is on qualitative changes during the acquisition process. Finally, Chapter 13 pays special attention to the acquisition of causal connectives. I investigate whether young children already make conceptual distinctions based on domains of use. It is not until this chapter that I will explicitly re-address the main question of this thesis.

Part IV (Chapter 14) presents a summary of the diachronic results and the acquisition results, thus providing a basis for conclusions on the hypotheses on form-function relations in the use of Dutch connectives. The studies in this thesis indicate that several form-function mappings can be distinguished, but that these mappings are often not one-to-one.

Chapter 2

Connectives and their classifications

This chapter has two aims. The first is to present definitions of the items I analyze as connectives. The second is to characterize and classify the selected connectives in more detail, both at the conceptual and at the syntactic level. I do not aim for an exhaustive description of possible conceptual and syntactic primitives discriminating between different types of connectives. I will only treat the most basic primitives, which are the ones that will be used in my analyses of the diachronic and the acquisition data.

> *"[Connectives] provide the linguistic means to organize one's knowledge of the world into larger interrelated chunks of information and to express qualitative distinctions in the nature of that interrelationship."* (Braunwald 1985: 513)

2.1 On the definition of *connectives*

This study focuses on so-called *connectives*. But what are connectives and what is their function in communication? It is clear that connectives play a role in making a text coherent, that is, making it more than just a random set of simply juxtaposed sentences. According to Halliday & Hasan (1976), the connectives in (1)-(3) can be regarded as cohesive devices (see also Martin 1992 for a more recent and more elaborate version of this idea). They function as a kind of cement, keeping the building blocks of the texts – in these cases clauses and sentences – together.

(1) Bob came home. **Subsequently** he unpacked his backpack.
(2) Nathan wanted to go out and play soccer. **So** he put on his shoes and grabbed his football.
(3) Mirjam was told to go to bed, **but** she wanted to watch TV.

As Hobbs (1979) indicates, connectives are not necessary to make a text coherent. Coherence relations, the meaning relations that connect two text segments, can also be construed in the absence of connectives. Thus, even without *subsequently*, *so* and *but*, the temporal, causal and contrastive relations in (1) to (3) can be inferred.

(4) Bob came home. He unpacked his backpack.
(5) Nathan wanted to go out and play soccer. He put on his shoes and grabbed his football.
(6) Mirjam was told to go to bed. She stayed downstairs to watch TV.

Connectives can be said to "provide the linguistic means to organize one's knowledge of the world into larger interrelated chunks of information and to express qualitative distinctions in the nature of that interrelationship" (Braunwald 1985: 513).

The examples in (1) to (3) illustrate that connectives are not restricted to one specific grammatical category. *Subsequently* and *so* can be regarded as adverbials, whereas *but* is

considered as a complementizer. Simple connectives are not the only cohesive devices for explicating coherence relations; speakers or writers may also resort to lexical signals or idiom chunks such as *on the other hand*, *for that reason* or *as a result*. As Knott & Sanders (1998: 172) note, these and other lexical signaling phrases are often the more specific – but also the less frequently used – cue phrases for the relations under consideration. As I will discuss in section 2.3, connectives may also differ from each other in other syntactic respects, such as positioning within the clause.

In psycholinguistic terms connectives can be treated as linguistic markers that instruct readers and listeners how to connect the new discourse segment with the previous one (cf. Britton 1994: 643-646). Evidence for this claim comes from various sources (see Sanders & Spooren 2005 for an overview). Firstly, recent psycholinguistic work has shown that connectives and other lexical signals seem to affect the process of constructing a text representation (cf. Cozijn 2000; Millis & Just 1994; Noordman & Vonk 1998; Sanders & Noordman 2000). Secondly, connectives affect the quality of that text representation, as becomes apparent from higher numbers of correct answers to text comprehension test (see Degand & Sanders 2002 for a discussion of relevant literature).

The processing instructions connectives give can be very explicit, but they may as well be underspecified. This means "the semantics of the connective that is used to indicate the link does not fully match the semantics of the relation that is intended by the speaker or writer" (Spooren 1997: 150). Underspecified relations can be recognized by exchanging the underspecified connective for an explicit counterpart (cf. also Givón 1990: 828). To illustrate this point, the a-variant in (7) shows an underspecified coherence relation between the last two sentences in a child-directed conversation from my corpus (Childes Database, Wijnen Corpus). The constructed b-version provides an explicit counterpart of this underspecified relation. In the explicit b-version *daardoor* 'as a result' signals to the child that a cause-consequence relation holds between the act of shooting and the breaking of the houses. The a-version is underspecified: *en dan* 'and then' only indicates a temporal succession relation between these events. In this case, the child has to infer the more specific causal relation himself.

(7) *Hier zit een geweer. Kan 'ie mee schieten. (...) Kan zo in de rondte draaien, hè?*
'Here is a gun. He can shoot with it. (…) (It) can turn around like this, can't it?'
a. *Kun je alle kanten op schieten.* ***En dan*** *gaan alle huizen kapot.*
'You can shoot in all directions. And then all the houses break.'
b. *Kun je alle kanten op schieten.* ***Daardoor*** *gaan alle huizen kapot.*
'You can shoot in all directions. Because of this all the houses break.'
(a-version spoken to Niek, 3;4.26)

Note that the existence of underspecification does not imply that connectives are completely interchangeable (see Spooren 1997: 154 for two restrictions). A so-called substitutability test can show which connectives are and which connectives are not appropriate for signaling a certain type of coherence relation (cf. Knott & Dale 1994; Knott & Sanders 1998). For example, the coherence relation in (7) cannot be expressed by the connectives *hoewel* 'although' and *maar* 'but'. Substituting *daardoor* for *hoewel* or *maar* would change the nature of the underlying causal coherence relation into a contrastive relation.

I only investigate those fragments in which the connective combines clauses or sentences. This restriction is necessary, since certain words functioning as connectives may

also serve other functions. For example, *and* and *but* may also function as words coordinating constituents within a clause (e.g. "Bert *and* Ernie", "short *but* brave").

Connectives can be regarded as markers of local discourse coherence. However, certain words functioning as connectives (e.g. *and*, *but* and *so*) may also be used at a more global level. For example, they can play a role in organizing and maintaining discourse topics, in managing the constantly changing information status of discourse segments or they might serve as some kind of turn-taking device (cf. Schiffrin 1987). To distinguish these global kinds of coherence from the more local coherence expressed by connectives, I will use the term *discourse marker* to refer to the global use of these words (cf. Schiffrin 1987, 2001; Lenk 1998; Traugott & Dasher 2002; Auer & Günthner 2003).[1]

In conclusion: language users have a variety of connectives at their disposal to explicate the type of local coherence between two combined text segments that consist minimally of clauses. This variety of connectives can be classified both at the conceptual level (see 2.2) and at the syntactic level (2.3).

2.2 Conceptual primitives in the characterization of connectives

So far, the basic idea is that "a connective serves to relate the content of the connected segments in a specific type of relationship" (Sanders & Spooren 2005). It seems fruitful then, to take the classification of coherence relations as a starting point for the conceptual characterization of connectives, even though there is no one-to-one mapping between relations and connectives. However, selecting an appropriate set of coherence relations is not as easy as it seems. Although much work has been done on the categorization of coherence relations (cf. Kehler 2002; Knott & Dale 1994; Mann & Thompson 1988; Pander Maat 1998; Sanders, Spooren & Noordman 1992, 1993), the alternative sets that have been put forward are very different from one another and no consensus has emerged about a single set of relations (see Hovy 1990 for a comparison).

In my study I make use of the cognitive account of coherence relations advocated by Sanders, Spooren & Noordman (1992, 1993). They work out the idea that the set of coherence relations should be organized (cf. Kehler 2002 for a similar approach) and give a parameter-based account of relations. According to them, coherence relations can be described with or decomposed into four cognitive primitives which are common to all relations. As Knott & Sanders (1998: 136) illustrate this point: "A particular relation, such as EVIDENCE, consists of more elementary notions, such as causality, and readers make use of their knowledge of this limited set of basic notions to derive the appropriate coherence relation." The advantage of this theory is twofold. First of all, it is more economic; Sanders et al. need fewer conceptual notions to describe all the different types of coherence relations than, for instance, Mann & Thompson (1988) do. Secondly, it results in a classification that is not only useful as an analytic tool for describing text structure, but that is also psychologically plausible. Coherence relations should be thought of in cognitive terms, as a set of conceptual relations used by language users when producing or processing a text. In this view, it is implausible that language users should have to work through a relatively long and unordered set of relations (as suggested by Mann & Thompson 1986) each time they process or produce one. The psychological plausibility of Sanders et al.'s approach to coherence relations turns it into

[1] In the literature many terms are used to refer to these markers of global coherence: discourse push/pop markers (Polanyi 1988), pragmatic particles (Östman 1981; Fraser 1996) and – leaving out even more terms – discourse operators (Redeker 1991). See Lenk (1998) for an evaluation of the terminology used to refer to the very heterogeneous class of 'discourse markers'.

an attractive model for the classification of connectives. After all, connectives are the linguistic counterparts of these cognitive entities.

Providing evidence from experimental studies, Sanders et al. (1992, 1993) argue for the saliency of four cognitive primitives: *basic operation*, *polarity*, *order of the segments*, and *source of coherence*.[2] In the subsections below I will illustrate how these primitives can be used to give a basic characterization of connectives at the conceptual level. Section 2.2.1 introduces the first three primitives; a separate section (2.2.2) is devoted to the fourth primitive, *source of coherence*, because there is quite a lot of discussion about the precise nature of the distinctions it makes.

2.2.1 Basic operation, polarity, and order of the segments

The first primitive, the *basic operation*, distinguishes between additive and causal relations. An *additive* operation exists if a relation of logical conjunction (*P & Q*) can be deduced between the two discourse segments. A *causal* operation exists if an implication relation (*P* → *Q*) can be deduced (cf. Sanders et al. 1992: 7; Sanders et al. 1993: 99). Connective examples of the two basic operations are given in (8) and (9) respectively.

(8) Mirjam is a girl. **And** Nathan is a boy.
(9) Bob fell, **because** he slipped over a banana.

Parallel to the two basic operations, two types of connectives can be distinguished. The additive relation is marked with the additive connective *and*; the causal relation is explicated with the causal connective *because*. Note that – because of the phenomenon of underspecification – the presence of an additive connective like *and* does not necessarily imply that the relation between the two clauses is also an additive one. Additive connectives are often compatible with a causal interpretation, although they do not themselves force such an interpretation (compare the constructed examples in (10) and (11)). Causal connectives on the other hand, are not compatible with purely additive relations. For instance, inserting the causal connective *so* in the additive example mentioned in (8) encourages language users to construct a context in which a causal relation is possible. In other words, *so* changes the originally additive relation into a causal one, which is indicated by the hash sign in (12).

(10) Bob slipped over a banana **and** he fell.
(11) Bob slipped over a banana. **That's** why he fell.
(12) #Mirjam is a girl. **So** Nathan is a boy.

The second primitive, *polarity*, distinguishes between positive and negative relations. A relation is *positive* if the two discourse segments S_1 and S_2 function in the basic operation as *P* and *Q* respectively. A relation is *negative* if not S_1 or S_2 but their negative counterparts, not-S_1 or not-S_2, function in the basic operation (Sanders et al. 1992: 10-11). The causal basic operation underlying the positive relation in (13) links the antecedent *slipping over a banana* with the consequent *falling*. S_1 and S_2 express the antecedent and consequent respectively. The coherence relation in (14) is an instantiation of the same causal basic operation. In this case, however, the second discourse segment (S2) expresses not-Q, that is, the negation of the consequent of the basic operation.

[2] Sanders (1992, Chapter 6) sketches how this set can be extended using so-called segment-specific criteria to cover all 23 relations in the descriptively adequate theory of Mann & Thompson (1988).

(13) [$_{S1}$ Bob slipped over a banana.] [$_{S2}$ **That's why** he fell.]
(14) [$_{S1}$ **Although** Bob slipped over a banana,] [$_{S2}$ he did not fall.]

Again, the distinction between positive and negative at the level of coherence relations can also be applied to the lexical markers of these relations. For example, (13) illustrates that *that's why* is a positive causal connective, whereas (14) shows that *although* serves as a negative causal connective.

The third primitive is called *order of the segments*. The order in a relation is *basic* if the information in the first clause (S_1) expresses *P* in the basic operation $P \& Q$ or $P \rightarrow Q$, and if the second clause (S_2), expresses *Q* (Sanders et al. 1992: 10). As additive relations are logically symmetric, the primitive *order of the segments* does not discriminate between different classes of additive relations (Sanders et al. 1992: 10). The causal relation in (13) displays this *basic order*. The order of the segments is *non-basic* if S_1 expresses *Q* and S_2 *P*. In (15) the second segment refers to the antecedent in the basic operation, so the relation displays the *non-basic* order.

(15) [$_{S1}$ Bob fell,] [$_{S2}$ **because** he slipped over a banana.]

In applying this primitive to connectives, it becomes clear that connectives can put specific constraints on the order in which segments can be realized (Sanders et al. 1993: 101). For instance, the Dutch connective *want* demands the non-basic order (see the Dutch equivalent of (15) in (16)), whereas *omdat* 'because' can be used to realize both basic and non-basic orders (see the two variants of the same clause in (17)).

(16) [$_{S1}$ *Bob viel,*] [$_{S2}$ ***want*** *hij gleed uit over een banaan.*]
(17) a. [$_{S1}$ *Bob viel,*] [$_{S2}$ ***omdat*** *hij over een banaan uitgleed.*]
b. [$_{S1}$ ***Omdat*** *Bob over een banaan uitgleed,*] [$_{S2}$ *viel hij.*]

As Sanders et al. (1992: 27) mention themselves, there is one group obviously absent in their proposal: the temporal relations. For instance, the example in (10) also involves a temporal relation between the slipping and the falling event, which can be marked with the temporal connective *then*.

(18) Bob slipped over a banana. **Then** he fell.

Sanders et al. (1992: 28) give two reasons why they do not consider temporality as a basic categorizing principle. The first is that the temporal meaning aspect is to a large degree determined by the propositional content (e.g. the tense and aspect) of the segments. The second reason for not including temporality is that it is not productive in the categorization of coherence relations like causality and additivity are. For these reasons, Sanders and colleagues claim "temporal relations belong to the classes of additive relations and that the properties distinguishing temporal relations from other additive relations concern the referential meaning of the individual segments" (p. 28). Although I acknowledge their conclusion that temporality is a segment-specific criterion rather than a basic primitive, I will make use of temporal subdivisions within the class of additive connectives. The reason for this is that it is not my goal to develop a basic tool for analyzing coherence relations in general, but rather to describe the conceptual behavior of individual connectives. For

example, segment-specific criteria like temporal overlap and temporal sequence are needed to discriminate between the Dutch connectives *terwijl* 'at the time/while' and *toen* 'then/when'.

2.2.2 Source of coherence

The fourth primitive, *source of coherence*, distinguishes between connectives expressing a semantic or a pragmatic relation.[3] A relation is *semantic* if the discourse segments are related because of their propositional content (Sanders et al. 1992: 7). In (19) the speaker or writer refers to the locutionary meaning of the segments; the coherence exists because the world that is described is perceived as coherent. A relation is *pragmatic* if the discourse segments are related because of the illocutionary meaning of one or both of the segments; the coherence exists because of the language user's goal-oriented communicative acts.[4] In pragmatic relations the coherence relation concerns the speech act status (e.g. conclusion, directive, question) of the segments. In cases like (20) the content of the second clause is presented as an argument supporting the conclusion mentioned in the first clause.

(19) Theo was exhausted because he had run to the university. (Sanders et al. 1993: 99)
(20) Theo was exhausted, because he was gasping for breath. (Sanders et al. 1993: 99)

Again, this distinction appears to be relevant for the classification of connectives as well. For example, two Dutch causal connectives (*daardoor* 'as a result' and *doordat* 'because (of the fact that)') can only be used to express semantic content relations (Pander Maat & Sanders 1995; Pit, Pander Maat & Sanders 1997; Stukker, Sanders & Verhagen 1999).

There is a lot of discussion about the exact definition of a distinction like the *source of coherence* (see e.g., Bateman & Rondhuis 1997; Degand 2001; Hovy 1990; Knott & Dale 1994; Knott & Sanders 1998; Martin 1992; Oversteegen 1997; Pander Maat 1998; Sanders & Spooren 1999; Verstraete 1998). At the same time, several researchers have come up with very similar distinctions, and there seems to be agreement on the characteristics of the prototypical relations (Sanders 1997). However, in this thesis I will often resort to an extended version of the semantic-pragmatic dichotomy as the basic tool for conceptual analyses: Sweetser's (1990) trifurcation into domains of use (see below). The main reason for this choice is that in literature concerning form-function relations, Sweetser's domains of use show up as the text-linguistic factors interacting with certain syntactic properties of adverbial clauses. In the remainder of this section, then, I will present the three domains of use, explain how they relate to the semantic-pragmatic dichotomy, and I will also introduce two additional subdivisions, which seem to be needed to cover the variety of connectives in this area.

Sweetser (1990: 77) has introduced a distinction between the content (example (21)), epistemic (see (22)) and speech-act domain (see (23)).[5] She has shown that these domains are

[3] See Van Dijk (1979) for an early observation concerning this distinction in the use of connectives.
[4] Pander Maat (1994a: 119) comments on this definition that uttering semantic relations is also goal-oriented; the goal of semantic relations – describing – just differs from the goal of pragmatic relations.
[5] Sweetser's multiple semantic domains roughly correspond, respectively, to the predicational, the propositional and the illocutionary layers within the layered structure of the clause in Functional Grammar (cf. Dik, Hengeveld, Vester & Vet 1990). In Schiffrin's (1987: 202) terms, the use in these three domains can be called fact-based, knowledge-based, and action-based respectively.

relevant to describe the use and meaning not only of connectives[6], but also of other linguistic devices like verbs of perception and modal elements.[7]

(21) We went out in the garden because the sun was shining.
(22) The temperature is probably going to rise, because the sun is shining.
(23) Let's have dinner in the garden, because the sun is shining.

The content clause in (21) describes a causal relation between two events in reality.[8] In the epistemic utterance (22) there is no CONSEQUENCE-CAUSE relation, but a CONCLUSION-ARGUMENT relation: the speaker gives an argument for his claim that the temperature is going to rise. Fragment (23) gives an illustration of a relation in the speech-act domain. The clause *the sun is shining* offers a reason for an action performed through speech: making the proposal *let's have dinner in the garden*. The three domains are also relevant for the classification of connectives. For example, the Dutch causal *want* can be applied in all three domains (just like English *because* in (21)-(23)), whereas the connective *omdat* can only be used to explicate the content and epistemic relations mentioned in (21) and (22).

How does Sweetser's trifurcation relate to the semantic-pragmatic dichotomy? The semantic category distinguished by Sanders et al. (1992) can be equated with Sweetser's content domain; the pragmatic category includes both her epistemic and speech-act domain. This does not imply that the pragmatic category can be disposed of. Verstraete (1998: 199-206) convincingly argues that a basic trifurcation (like Sweetser's) does not give a satisfactory account of the similarities between epistemic and speech-act relations. He claims that the bipartition semantic-pragmatic (in his terms external-internal) is still necessary, and explains the conceptual difference between the two by showing that they relate to different components of the matrix clause. If an adverbial clause is used semantically it relates to the so-called *instantiated type* of the matrix clause, the description of a state of affairs as expressed by the lexical verb and its arguments. If an adverbial clause is used pragmatically it relates to the 'ground' of the matrix clause, "the speech event, its setting, and its participants" (Langacker 1985: 113). This ground is expressed by the speaker-encoding categories of tense, modality, and illocutionary value. Verstraete furthermore argues that the distinction between epistemic (in his terms *modal*) and speech-act (in his terms *speech-functional*) relations should be seen as a subdivision within the pragmatic category.[9] In this thesis, then, the semantic-pragmatic distinction will be used in addition to the domains-trifurcation.

---

[6] This claim holds for all types of connectives, but has been worked out and used most often in research into causal connectives.

[7] Several authors claim a fourth layer is required (e.g. Crevels 2000; Lang 2000). However, there is no agreement about the operationalization of this extra "domain". For example, Crevels (2000) suggests, in line with Hengeveld (1997), a so-called *text* layer should be added. Connectives pertaining to this layer have to do with the organization of the discourse and, therefore, apply to text units that may contain more than one sentence (Crevels 2000: 315). Although it appears to be useful at the syntactic level to investigate whether connectives may combine discourse segments larger than one sentence or not, I see no use in distinguishing such a separate conceptual text layer if the relation between the connective clause and the preceding text unit (larger than one sentence) can still be characterized on the basis of the other three domains.

[8] The examples in (21)-(23) were originally constructed by Ninke Stukker.

[9] This solution of maintaining the intermediate *pragmatic* category is also attractive for authors like Keller (1995: 24), who stresses the similarities between the epistemic and speech act domain and is reluctant to extract a separate third domain.

Two further subdivisions will be made: one within the semantic or content domain and one within the pragmatic category. The first extension of Sweetser's trifurcation is the subdivision of the content domain into *volitional* and *non-volitional* relations (cf. also Mann & Thompson's (1988) distinction between *volitional result* and *cause*). The notion of 'volitionality' distinguishes content relations in which human intentions do play a role (volitional relations like (21)) from content relations in which this is not the case (non-volitional relations like (24)).

(24) The temperature rose quickly because the sun was shining.

Earlier research has shown that this notion of volitionality is useful in describing the meanings of Dutch causal connectives (see Pander Maat & Degand 2001; Pander Maat & Sanders 1995; Pit et al. 1997; Stukker et al. 1999). This holds true especially for the contrast between *daarom* 'that's why' versus *daardoor* 'as a result' and the connectives *doordat* 'because (of the fact that)' versus *omdat* 'because': *doordat* and *daardoor* only occur in non-volitional contexts, whereas *omdat* and *daarom* are preferred for use in volitional relations (cf. Pander Maat & Sanders 1996, 2000, 2001; Pit 2003; Stukker et al. 1999).

A final subdivision to be added in relation to the semantic-pragmatic distinction is the distinction between so-called *abductive* and *non-abductive* relations within the pragmatic category (cf. Degand 1996, 2001; Pander Maat & Degand 2001).[10] The utterance in (25) can be classified as abductive: the argument *I saw her lights were out* can only be given in support of the conclusion and cannot be taken as the real-world cause for the action mentioned in the conclusion at the same time. Abductive relations, then, are exclusively pragmatic. Non-abductive examples (like (26)), on the other hand, appear to be ambiguous between a semantic and a pragmatic interpretation (see Sanders 1997). Without context, the *because*-clause in (26) can be taken both as the reason for Mirjam's returning home (resulting in a semantic content relation) and as an argument in support of the speaker's conclusion that Mirjam is at home (resulting in a pragmatic epistemic relation). As Sanders (1997: 130) notes "relations which appear to be ambiguous with respect to the semantic-pragmatic distinction turn out *not* to be ambiguous in context." The utterance in (26) clearly receives a semantic interpretation as an answer to the question in (27); the reason that Mirjam returned home was that she realized her lights were still on. A pragmatic interpretation is preferred if the same utterance serves as an answer to the question in (28). In that case, (26) can be interpreted as "I am sure; I saw her lights were still on when I passed her home."

(25) Mirjam is not at home, because I saw her lights were out.
(26) Mirjam is home again, because her lights were still on.
(27) Why did Mirjam return home?
(28) Are you sure Mirjam came back home?

Within the speech-act domain the same subdivision can be made: in the non-abductive example in (29) the argument given in support of asking the question may also be the real-

[10] Pander Maat & Degand (2001: 221-225) present different labels for this distinction. In their terminology *non-abductive* relations are 'causality-based epistemic relations'; *abductive* relations are labeled 'noncausal epistemic relations'. I prefer the term *abductive* to refer to this class of relations, which remains 'causal' in the general sense.

world cause for Mirjam's coming home. In contrast, in the abductive utterance in (30) seeing that the lights were out cannot be taken as the reason for Mirjam's being home.

(29) Did Mirjam come home? Because her lights were still on.
(30) Is Mirjam at home? Because I saw her lights were out.

The distinction between abductive and non-abductive relations is relevant for the conceptual classification of connectives as well. For example, it is known from the literature (cf. Degand 1996; Pit 2003) that *omdat* can only mark non-abductive epistemic relations, whereas *want* can mark both abductive and non-abductive epistemic relations. This is illustrated in (31) and (32), the Dutch equivalents of (25) and (26) respectively.

(31) a. *Mirjam is niet thuis,* ***want*** *ik zag dat haar lampen uit waren.*
b. #*Mirjam is niet thuis,* ***omdat*** *ik zag dat haar lampen uit waren.*
'Mirjam is not at home, because I saw her lights were out.'
(32) a. *Mirjam ging terug naar huis,* ***want*** *haar lampen waren nog aan.*
b. *Mirjam ging terug naar huis,* ***omdat*** *haar lampen nog aan waren.*
'Mirjam returned to her house, because her lights were still on.'

To summarize this section, Table 2.1 gives an overview of the different categories distinguished within the *source of coherence*.

Table 2.1. Overview of the different categories within the *source of coherence* primitive

Source of coherence	**Domain**	**Further subdivision**	**Example**
semantic	content	non-volitional	(24)
		volitional	(21)
pragmatic	epistemic	non-abductive	(26)
		abductive	(25)
	speech act	non-abductive	(29)
		abductive	(30)

Recently, some researchers have rejected Sweetser's original model as a means of accounting for the distribution and interpretation of specific causal markers (cf. Pander Maat & Sanders 2000; Pander Maat & Degand 2001; Pit 2003) and have argued instead in favor of a notion of subjectivity. For example, Pander Maat & Degand (2001) put forward a subjectivity scale, one of *speaker involvement*. This scale is a continuum on which non-volitional content relations are maximally objective, whereas epistemic relations are very subjective, and volitional content relations hold an intermediate position. Despite certain merits of the subjectivity approaches (cf. Pander Maat & Sanders 2000: 77-78), I will resort to the original three-domain distinction in this thesis. The two main reasons for this choice are the following. First of all, in literature concerning form-function relations, it is Sweetser's domains of use rather than the subjectivity scales that show up as text-linguistic factors which interact with certain syntactic properties of adverbial clauses (see especially the literature in section 3.4 and 3.5 in Chapter 3). Secondly, the different authors adhering to subjectivity do not completely dispose of the distinction based on domains; they often assign a relative position to the different domains on their subjectivity scales. In this thesis, then, the majority of causal data will be analyzed on the basis of the three-domain distinction. In Chapter 8, however, the

subjectivity approach will also be taken into account in the diachronic analysis of four Dutch causals. This chapter may shed some light on the competition between the two theories, since it allows for a comparison of the two approaches.

2.3 Syntactic primitives in the characterization of connectives

At the sentence-linguistic level, I examine both the syntactic properties of the connectives themselves and the different syntactic environments in which connectives occur. In order to arrive at a syntactic characterization, four primitives are combined, namely a) the position of the connective within the connective clause, b) the word order of the core constituents (subject, object, finite verb and main verb) in the connective clause, c) the position of the connective clause as a whole, and d) the word order within the clause to which the connective clause is related. These primitives are taken up for further discussion in section 2.3.1 to 2.3.4 respectively.

My syntactic analysis deviates from traditional approaches in that I do not start with a classification based on categorical notions such as adverb and subordinator. I have two reasons for preferring the four syntactic primitives instead of categorical status. First of all, primitives like positioning and word order seem to be the more basic notions, since the categorical status of connectives is often (in part) derived from their behavior along these parameters. Furthermore, the distinction between different syntactic categories is not always clear-cut, at least not in Dutch. For example, both the *Van Dale* dictionary and the descriptive grammar *Algemene Nederlandse Spraakkunst* (henceforth *ANS*) 'General Dutch Grammar', (edited by Haesereyn et al. 1997) grammar not only distinguish between *voegwoorden* 'conjunctions' and *bijwoorden* 'adverbials'; they also use the intermediate category of *voegwoordelijke bijwoorden* 'conjunctive adverbials'. The occurrence of intermediate categories as such need not be problematic. However, the specific grammatical category of conjunctive adverbials does pose a problem, since its definition in the *ANS* conflates syntactic and semantic notions (cf. also Smessaert & Beeken 1995: 376). The first two syntactically oriented sentences in this definition (see (33)) show why conjunctive adverbials should be seen as adverbials. This part of the definition sets the conjunctive adverbials apart from conjunctives. The third, semantically oriented sentence illustrates the 'conjunctive' nature of this type of connective.

(33) <u>Definition of conjunctive adverbials in the *ANS* (1997: 462)</u>

Grammaticaal onderscheiden voegwoordelijke bijwoorden zich van nevenschikkende voegwoorden doordat ze volledig deel uitmaken van de zin. Ze kunnen op verschillende plaatsen in de zin voorkomen, terwijl nevenschikkende voegwoorden als zinsverbindende elementen als het ware tussen zinnen in staan. Net als de voegwoorden leggen voegwoordelijke bijwoorden een logisch verband tussen twee zinnen of delen van zinnen.

'Grammatically speaking conjunctive adverbials differ from coordinate conjunctions in that they are fully part of the sentence. They can occur in different places in the sentence, whereas coordinating conjunctions occur as clause combining elements in between two sentences, so to speak. Just like conjunctions, conjunctive adverbials establish a logical connection between two sentences or parts of sentences.'

The use of terms conflating conceptual and syntactic notions beforehand does not seem attractive to me, since I first want to investigate conceptual and syntactic properties of connectives independently and subsequently study the way they interact. Therefore, the four primitives will be used as the primary tools for the syntactic analyses. In section 2.3.5 the

analyses along the four primitives will be related to the more traditional characterizations based on categorical status.

2.3.1 Positioning of the connective

In order to describe the syntactic properties of connectives with respect to their positioning, three patterns are taken into account (cf. Lamiroy & Van Belle 1995: 402; Smessaert & Beeken 1995: 378). Although I refer to these positionings in linear terms, I do believe that these relative positionings can only be accounted for in hierarchical terms (which generative frameworks insightfully illustrate with tree diagrams).

(34) <u>Positioning possibilities of connectives:</u>
a. clause-initial position
b. clause-medial position
c. clause-final position

The clause-initial position is the position at the head of the connective clause (see (35)). The clause-final position refers to the position after all other constituents in the clause (see (36)). The clause-medial position refers to a position between the finite verb (or in case of subordinate clauses: the subject) and the remaining (parts of the) verb(s), such as a verb particle, a participle or an infinitive (see (37)a-c).[11]

(35) ***Dus*** *hij heeft dat boek uitgelezen.*
So he has that book out-read-participle.
'So he has finished that book.'
(36) *Hij heeft dat boek uitgelezen* ***dus****.*
He has that book out-read-participle so.
'So he has finished that book.'
(37) a. *Hij las* ***dus*** *dat boek uit.*
He read so that book out-particle.
'So he finished reading that book.'
b. *Hij heeft dat boek* ***dus*** *uitgelezen.*
He has that book so out-read-participle.
'So he finished reading that book.'
c. *Hij moest* ***dus*** *dat boek lezen.*
He had-to so that book read.
'So he had to read that book.'

The classification in (34) does not completely match the patterns introduced by Smessaert & Beeken (1995: 378), who argue for a subdivision within the a-category. This subdivision will

[11] In principle, it is possible to subdivide the clause-medial position into several categories (e.g. in the examples in (37) the connective can be positioned before or after the direct object *dat boek* 'that book'). There are two reasons to disregard these further subdivisions. Firstly, whether connectives are placed in one clause-medial position or another seems to depend on triggers for scrambling (e.g. focus, stress) rather than the syntactic possibilies of these connectives themselves. Furthermore, the various clause-medial positions do not seem to differentiate between individual connectives.

be taken up for discussion in section 2.3.2, since it involves a second syntactic primitive: the positioning of other elements – especially the finite verb – within the connective clause.[12]

2.3.2 Word order within the connective clause

My interest in the second primitive, word order within the connective clause, derives from the fact that word order in Dutch in certain respects shows more variation than in, for example, English or French. As the examples in (38) and (39) show, English and French do not differentiate between their word orders in main and subordinate clauses. The Dutch counterparts of these clauses (given in (40)) show that Dutch does exhibit a well-documented difference between the two clause types (cf. among many others De Haan 2001: 3-4; Lamiroy & Van Belle 1995: 401-402). In a typical declarative main clause (see the a-example in (40)) the finite verb or the finite part of a complex verb typically occupies the second position, preceded by the subject, an expletive or some topicalized element. The complements and adjuncts are placed between the finite verb and the remaining elements of the verb phrase (particle, infinitive). In subordinate clauses, the whole verb complex stands in final position (see the b- and c-version in (40); all examples taken from Smessaert & Beeken 1995: 377-378).[13]

(38) a. He has read the book.
b. (…) that he has read the book.
(39) a. Il a lu le livre.
b. (…) qu'il a lu le livre.
(40) a. Hij <u>heeft</u> het boek gelezen.
b. (…) dat hij het boek <u>heeft</u> gelezen.
c. (…) dat hij het boek gelezen <u>heeft</u>.

In other words, Dutch (and German) declarative main clauses show so-called *verb second* or *V2*, but subordinate clauses show so-called *V-late*.

How can this word order primitive be used to discriminate between connectives? The different patterns can be regarded as different syntactic configurations in which connectives may or may not occur. Certain connectives only occur at the head of clauses showing V2; other connectives only introduce clauses exhibiting V-late. A typical Dutch 'V2-connective' is the causal connective *want* (see (41)); its V-late counterpart is *omdat* (see (42)).

(41) a. *Bob weet het antwoord,* ***want*** *hij <u>heeft</u> het boek gelezen.*
'Bob knows the answer, because he has read the book.'
b. **Bob weet het antwoord,* ***want*** *hij het boek <u>heeft</u> gelezen.*
(42) a. **Bob weet het antwoord,* ***omdat*** *hij <u>heeft</u> het boek gelezen.*
b. *Bob weet het antwoord,* ***omdat*** *hij het boek <u>heeft</u> gelezen.*

---

[12] Smessaert & Beeken (1995) also distinguish between clause-initial position with and without intonational or punctuational break. Since my classification will mainly be used for analysis of historical data (which does not provide intonational nor punctuational information), this category will be disregarded here. As Smessaert & Beeken (1995: 382, Tables III and IV) show, the absence or presence of such a break is hardly discriminative between the connectives discussed here anyway.

[13] In *yes/no* questions, conditionals, and imperatives, the finite verb appears in first position.

Within the clauses showing V2, a further distinction can be made. The V2 word order also appears when – instead of the subject – a wh-element or another constituent (e.g. a topicalized object, adjunct or a fronted subclause) occupies the first position. The topicalization of these constituents (exemplified in (43)) triggers so-called inversion of the subject and the finite verb.[14]

(43) a. *Welk boek heeft hij gelezen?*
'Which book did he read?'
b. *Dat boek heeft hij gelezen.*
'That book did he read.'

The absence or presence of inversion is the syntactic factor used by Smessaert & Beeken (1995: 378) to make a subdivision between different types of connectives in clause-initial position. Certain connectives in this position function as topicalized elements, triggering inversion of the subject and the finite verb themselves. A typical Dutch example is *daarom* 'that's why' (see (44)). Other connectives in clause-initial position, like *want*, do not trigger inversion (see (45)).

(44) a. *Bob hoorde dat het een leuk boek was.* ***Daarom** hij heeft het gisteren gelezen.*
'Bob heard it was a nice book. That's why he read it yesterday.'
b. *Bob hoorde dat het een leuk boek was.* ***Daarom** heeft hij het gisteren gelezen.*
(45) a. *Bob weet het antwoord,* ***want** hij heeft het boek gisteren gelezen.*
'Bob knows the answer, because he has read the book yesterday.'
b. **Bob weet het antwoord,* ***want** heeft hij het boek gisteren gelezen.*

Connectives that trigger inversion should be regarded as constituents within the connective clause. This can be inferred from the fact that they cannot co-occur with other topicalized elements. For example, in (46) *daarom* cannot occur in combination with the topicalized adverbial *gisteren* 'yesterday'. Such a combination is possible with *want* (see (47)).

(46) *Bob hoorde dat het een leuk boek was.* ***Daarom** gisteren heeft hij het gelezen.*
Bob heard it was a nice book. That's why yesterday he read it.
(47) *Bob weet het antwoord,* ***want** gisteren heeft hij het boek gelezen.*
Bob knows the answer, because yesterday he has read the book.

An additional argument in favor or this analysis is that *daarom* can co-occur with a connective that does not trigger inversion (cf. Dik 1968: 34; De Vries 1971: 416 for a similar test), whereas *want* cannot (see (48)).

---

[14] Note that the two word order aspects (verb-second and inversion because of topicalization) are not completely independent of each other: the impossibility of V2 in certain clauses implies that topicalization triggered inversion is also not at stake. For example, topicalization of the object is not possible in subordinate clauses:
(i) (…) **dat dit boek heeft hij gelezen.*
(…) that this book has he read.
'(…) that he has read this book.'

(48) a. *Bob hoorde dat het een leuk boek was* ***en daarom*** *heeft hij het gelezen.*
'Bob heard it was a nice book. And that's why he read it.'
b. **Bob weet het antwoord,* ***en want*** *hij heeft het boek gelezen.*
'Bob knows the answer, and because he has read the book.'

So, although the position of the two connectives at the head of the connective clause looks the same at a superficial level (that is, only taking into account the linear order with respect to the rest of the connective clause), the contrast between connectives with and without the possibility of inversion shows that two clause-initial positions should be distinguished. Connectives that trigger inversion must be considered as constituents within the connective clause. These connectives are really contained in their host clause, functioning both as a brick within the clause and as cement holding two clausal bricks together. Connectives that do not trigger inversion themselves are not part of their host clause; they only function as cement between the combined clauses.

In conclusion, three word order patterns can be distinguished (see (49)). These patterns can be taken as syntactic configurations in which connectives may or may not occur.

(49) Word order possibilities within the connective clause:
a. V2 without inversion triggered by the connective
b. V2 with inversion triggered by the connective
c. V-late

2.3.3 Linearization of the connective clause

A third syntactic primitive is the linearization of the connective clause as a whole. Using this primitive, three positions of the connective clause can be distinguished. In (50) the bracketing marks out the two related clauses, which are labeled S_1 and S_2 respectively.[15]

(50) Linearization possibilities of connective clauses:
a. postposed position: [S_1] [connective S_2]
b. preposed position: [connective S_1] [S_2]
c. intraposed position: [S_1… [connective S_2] … S_1]

If the connective clause functions as S_2 (as in (51)), it occurs in the postposed position, which Lehmann (1988: 186) labels the 'central' position (cf. also Thompson & Longacre 1985).

(51) [$_{S1}$ *Na het eten moest hij de struiken opzoeken*], [$_{S2}$ ***omdat*** *alles er weer uit kwam.*]
(Erin de Enige, 1990)
'After dinner he had to seek out the bushes, because everything came out again.´

The connective clause may alternatively function as S_1, as in (52). In this case, it occurs in the preposed position, or in Lehmann's terms the 'marginal' position.

(52) [$_{S1}$ ***Omdat*** *alles er weer uit kwam*], [$_{S2}$ *moest hij na het eten de struiken opzoeken*.]
'Because everything came out again, he had to seek out the bushes after dinner.'

[15] Note that the bracketing need not coincide with the precise syntactic demarcation of the clauses.

The third possibility is that the connective clause functions as a kind of interjection, occurring in the middle of S_1. An example of such a connective clause in intraposed position is given in (53).

(53) [$_{S1}$ *Hij moest* – [$_{S2}$ ***omdat*** *alles er weer uit kwam*] – *na het eten de struiken opzoeken*.]
'He had – because everything came out again – to seek out the bushes after dinner.'

Note that the syntactic bipartition preposed/postposed does not coincide with the conceptual distinction between basic and non-basic order of the segments (mentioned in section 2.2.1). For example, clause combinations with the connective clause in postposed position can either show a non-basic order or a basic order of the segments. In (51) the *omdat*-clause functions as the antecedent, which results in the non-basic word order. In (54) the connective clause in postposed position functions as the consequent, which results in a basic order of the segments. The two distinctions, then, should indeed be seen as separate primitives.

(54) [$_{S1}$ *Na het eten kwam alles er weer uit*.] [$_{S2}$ ***Daarom*** *moest hij de struiken opzoeken*.]
'After dinner everything came out again. That´s why he had to seek out the bushes.´

2.3.4 Word order within the matrix clause

The final syntactic primitive to be discussed here involves the word order within the clause that follows the connective clause. This primitive, then, only applies to connective clauses in preposed position. König & Van der Auwera (1988: 127) argue there are three ways of combining Dutch adverbial clauses followed by a main clause: (i) mere juxtaposition, (ii) juxtaposition and linking through a resumptive element introducing the main clause, and (iii) complete integration (see examples (55)-(57), taken from König & Van der Auwera 1988: 102-103).

(55) <u>Non-integrative word order (adverbial clause - topic - finite verb - …)</u>
Zelfs als ze niet allemaal gekomen zijn, we <u>kunnen</u> met het bezoek tevreden zijn.
Even if they not all come are, we can with the attendance content be.
'Even if they haven't all come, we can be pleased with the attendance.'

(56) <u>Resumptive word order (adverbial clause - resumptive - finite verb - …)</u>
Als hij ziek is, dan <u>blijft</u> hij thuis.
If he sick is, then stays he at home.
'If he is sick, then he stays at home.'

(57) <u>Integrative word order (adverbial clause - finite verb - …)</u>
Omdat hij ziek is, <u>kan</u> Fred niet meekomen.
Because he sick is, can Fred not come along.
'Because he is sick, Fred cannot come along.'

The adverbial clauses in these examples all show two properties traditionally taken as indicators of subordination in Germanic languages: V-late placement and the presence of a connective. Looking at the position of the finite verb in the matrix clause, however, only (57) can be considered as a clear case of embedding in the sense of functioning as a constituent within another clause. In this example, the finite verb shows up in the initial position of the main clause, thus taking second position in the overall complex sentence. The other two adverbial clauses cannot be assumed to function as the first constituents of the second clause. In the matrix clause of (56), the finite verb is preceded by a resumptive (or correlative)

element, resulting in a kind of left dislocation of the adverbial clause. The matrix clause in (55) exhibits a non-integrative verb-second word order, which separate sentences generally have (topic – finite verb).

2.3.5 The four syntactic primitives in relation to categorical status

How do the four syntactic primitives relate to traditional syntactic classifications based on categorical status? Below I discuss the way in which the categories *coordinator*, *subordinator* and *adverb* can be characterized using the four syntactic primitives.

Connectives traditionally labeled 'coordinators' or 'coordinating conjuncts' are rather restricted in their syntactic possibilities.[16] As the grammaticality judgments in (58) indicate, the use of coordinators is restricted to the clause-initial position.[17]

(58) *Ik moet even dat mesje hebben…*
'I need to have that knife for a moment…'
a. ***want*** *ik moet brood snijden.*
'because I have to cut bread.'
b. **ik moet* ***want*** *brood snijden.*
c. **ik moet brood snijden* ***want****.*

In combinations of declarative clauses, coordinators need to occur in clauses showing V2 (see the a-variant of the *want*-utterance in (58), repeated in (59)a). *Want* in combination with V-late results in an ungrammatical clause (see the b-variant in (59)). Furthermore, coordinators do not trigger inversion themselves, as is indicated by the ungrammaticality of the placement of the finite verb in the c-version of (59).

(59) *Ik moet even dat mesje hebben…*
a. ***want*** *ik* <u>*moet*</u> *brood snijden.*
b. ****want*** *ik brood snijden* <u>*moet*</u>*.*
c. ****want*** <u>*moet*</u> *ik brood snijden.*

In terms of linearization, coordinators can only occupy the postposed position (see the grammaticality of the a-clause, and the ungrammaticality of the b- and c-clause in (60) respectively).

(60) a. [$_{S1}$ *Na het eten moest hij de struiken opzoeken*], [$_{S2}$ ***want*** *alles kwam er weer uit.*]
'After dinner he had to seek out the bushes, because everything came out again.'
b. *[$_{S1}$ ***Want*** *alles kwam er weer uit*], [$_{S2}$ *hij moest na het eten de struiken opzoeken*.]
'Because everything came out again, he had to seek out the bushes after dinner.'
c. */$^{??}$[$_{S1}$ *Hij moest* – [$_{S2}$ ***want*** *alles kwam er weer uit*] – *na het eten de struiken opzoeken*]
'He had to – because everything came out again – seek out the bushes after dinner.'

---

[16] I do acknowledge that "coordination is not a homogeneous phenomenon" (De Haan 2001: 29). For ease of exposition, however, I have choosen to leave out syntactic details that illustrate the distinction between so-called 'normal' and 'paratactic' coordination (cf. De Haan 2001: 29-33 and the references there). A coordinator like *and* belongs to the first type, Dutch *want* to the second.

[17] The examples in (58) and (59) are all variations of a clause combination inspired by an utterance of the child Josse (Childes Database, Groningen Corpus).

Since coordinators "occur obligatorily to the right of their matrix clause" (cf. De Haan 2001: 15) the fourth syntactic primitive – which discriminates different word orders within a postposed matrix clause (see section 2.3.4) – does not apply to coordinators.

The second traditional category is the class of 'subordinators' or 'subordinating conjunctions'. As the grammaticality judgments in (61) indicate, the use of subordinators is restricted to the initial position of the connective clause.

(61) *Ik moet even dat mesje hebben…*
'I need to have that knife for a moment…'
a. ***omdat*** *ik brood moet snijden.*
'because I have to cut bread.'
b. **ik moet* ***omdat*** *brood snijden.*
c. **ik moet brood snijden* ***omdat****.*

Unlike coordinators, traditional subordinators cannot show up in declarative clauses with a V2 word order. As the first two *omdat*-examples in (62) illustrate, subordinators must combine with the V-late word order. Just like coordinators however, they cannot trigger inversion of the subject and the finite verb themselves. This indicates that neither coordinators nor subordinators function as topicalized constituents within the clauses they head.

(62) *Ik moet even dat mesje hebben…*
a. ****omdat*** *ik <u>moet</u> brood snijden.*
b. ***omdat*** *ik brood <u>moet</u> snijden.*
c. ****omdat*** *<u>moet</u> ik brood snijden.*

In terms of linearization, subordinators have maximal freedom. The *omdat*-examples in (51), (52) and (53) have already illustrated that subordinators can occur in postposed, preposed, or intraposed position.[18] If the subordinating connective clause is preposed (see (57), repeated here as (63)), the modern Dutch word order within the matrix clause is typically integrative (see König & Van der Auwera 1988 for certain exceptions).

(63) *Omdat hij ziek is, <u>kan</u> Fred niet meekomen.*
Because he sick is, can Fred not come along.
'Because he is sick, Fred cannot come along.'

The inversion of the subject and the finite verb in the matrix clause indicates that the connective clause functions as a constituent within that matrix clause.

The third traditional syntactic category is that of 'adverbs'. In contrast to coordinating and subordinating conjunctions, adverbs have more positional possibilities within the connective clause. Adverbs can occur in a clause-initial position (see (64)a), but also in clause-medial position (see (64)b). Whether an adverb can occur in clause-final position,

[18] This claim does not hold for all subordinators; certain connectives may – because of their semantics – force the subordinate clause in a fixed position after the main clause (cf. Lehmann 1988: 188, examples (13) and (14)).

depends on the specific connective (see the difference between *dus* 'so' and *daarom* 'that's why' in (64)c).[19]

(64) *Ik moet dat brood snijden…*
'I have to cut that bread…'
a. ***dus/daarom*** *moet ik dat mesje even hebben.*
'so/that's why I need to have that knife for a moment.'
b. *ik moet* ***dus/daarom*** *even dat mesje hebben.*
c. *ik moet even dat mesje hebben* ***dus/*daarom****.*

If the adverb is in clause-initial position, it triggers inversion of the subject and the verb (see (65)a). In this position, it alone cannot trigger a V2 or a V-late word order (compare (65)b-c).

(65) *Ik moet dat brood snijden…*
a. ***daarom*** *moet ik even dat mesje hebben.*
b. ****daarom*** *ik moet even dat mesje hebben.*
c. ****daarom*** *ik even dat mesje moet hebben.*

Frequently, clause combinations containing an adverbial connective have the connective clause in postposed position (see (64)). However, the exact linearization possibilities depend on the semantics of the adverb (e.g. whether it has a cataphoric or anaphoric nature) and therefore vary per adverbial connective. If the adverbial connective clause is preposed, the word order within the related clause is not affected by the adverb.

The distinction between the three syntactic categories is not always as clear-cut as it seems. Certain connectives show characteristics of two different categories. For example, *dus* 'so' may show a V2 word order, but as (66) illustrates, it can also trigger inversion. It alone cannot trigger a V-late word order (cf. (66)c). This connective then shows word order characteristics of both a coordinator and an adverb; according to the *ANS* it should be classified as a conjunctive adverb, the disputable intermediate category mentioned earlier.

(66) *Ik moet dat brood snijden…*
'I have to cut that bread…'
a. ***dus*** *ik moet even dat mesje hebben.*
'so I need to have that knife for a minute.'
b. ***dus*** *moet ik even dat mesje hebben.*
c. ****dus*** *ik dat mesje even moet hebben.*

A different type of ambivalence can be illustrated by *toen* 'then/when'. As its double translation already suggests, and the different word orders in (67) show, this connective shows characteristics of both an adverb (triggering inversion) and a subordinator (showing V-late).

[19] Note that – *dus* and *daarom* being forward instead of backward causal connectives – the order of antecedent and consequent is reversed here.

(67) *Ik moest dat mesje even hebben…*
'I needed that knife for a minute…'
a. ***toen*** *moest ik dat brood snijden.*
'then/at the time I had to cut that bread.'
b. ***toen*** *ik dat brood moest snijden.*
'when I had to cut that bread.'
c. ****toen*** *ik moest dat brood snijden.*
'then/when I had to cut that bread.'

The existence of these ambivalent connectives again shows the need for a connective analysis based on the four syntactic primitives instead of an analysis based on categorical status.

2.4 Characterization of the selected Dutch connectives

In this section I give both a conceptual and a syntactic characterization of the Dutch connectives selected here for investigation. According to the first three conceptual primitives treated in section 2.2, the selected connectives can – at least in modern Dutch – be classified as in Table 2.2. The fourth primitive, source of coherence, is disregarded in this table, since it will only be applied to the four causal connectives *want*, *omdat*, *dus* and *daarom* (see Chapter 6 and 7).

Table 2.2. Classification of connectives according to three conceptual primitives

Connective	Basic operation	Polarity	Order of the segments
daarom	causal	positive	basic
dus	causal	positive	basic
en	additive	positive	(does not apply)
maar	additive	negative	(does not apply)
omdat	causal	positive	basic, non-basic
toen	additive (temporal sequence)	positive	(does not apply)
want	causal	positive	non-basic

According to the first three syntactic primitives treated in section 2.3, the modern Dutch connectives selected for investigation can be classified as in Table 2.3.

Table 2.3. Classification of connectives according to three syntactic primitives

Connective	Position	Word order[20]	Linearization
daarom	initial, medial	inversion	postposed
dus	initial, medial, final	V2, inversion	postposed
en	initial	V2	postposed
maar	initial	V2	postposed
omdat	initial	V-late	intra-, pre-, postposed
toen	initial, medial, final	inversion, V-late	intra-, pre-, postposed
want	initial	V2	postposed

[20] This column only shows the word order possibilities of declarative connective clauses in which the connective is in clause-initial position.

The fourth syntactic primitive, word order within the matrix clause, is disregarded in Table 2.3, since it only applies to subordinating connectives used in preposed position. However, it can be concluded that the two modern Dutch connectives (*omdat* and *toen*) in preposed subordinate clauses typically demand an integrative word order.

In the current chapter the various conceptual and syntactic primitives were treated as if they operate completely independent of each other. Chapter 3 will discuss theories and hypotheses about the interaction between the conceptual and syntactic properties of connectives.

Chapter 3

Theories on form-function relations

The current chapter has two aims. The first is to provide evidence for the idea that there is an interaction between form and function of connectives and the clauses containing them. The second aim is to introduce theories about specific form-function relations in the use of connectives. To facilitate this introduction, the four syntactic primitives presented in Chapter 2 will be discussed in relation to certain text-linguistic functions or conceptual properties. The current chapter provides a theoretical foundation for the subsequent chapters on specific connective developments in history and child language.

> "*(…) there is a relationship between the form of a sentence and its function in discourse, and (…) grammatical form is* in part *determined by the pragmatic circumstances under which the sentence is used as a unit of information*."
> (Lambrecht 1988: 138)

3.1 Introduction

The idea of a possible interaction between the form and function of a linguistic element is not new. For example, Lambrecht (1988) states: "there is a relationship between the form of a sentence and its function in discourse, and that grammatical form is *in part* determined by the pragmatic circumstances under which the sentence is used as a unit of information" (p. 138). The current chapter adheres to this view. In fact, it argues that each of the four syntactic primitives discussed in the previous chapter can be related to certain conceptual properties.

In this chapter, I discuss theories on specific form-function relations pertaining to the use of connectives. In the first four subsections, each of the syntactic primitives will be discussed in relation to particular conceptual properties. Section 3.2 studies the functional counterpart of the connective positioning, section 3.3 looks at the text-linguistic parallel of the linearization of connective clauses, section 3.4 will discuss the function of the word order within postposed matrix clauses, and section 3.5 examines the conceptual counterparts of the word order within connective clauses themselves. Section 3.6 introduces a form-function interaction specific to diachronic developments, namely, the interaction between so-called grammaticalization (a process that is often accompanied by syntactic changes) and subjectification.

3.2 Interactions with connective positioning

Certain connectives can be inserted into various syntactic slots (see section 2.3.1), as is illustrated by *dus* 'so' in (1).

(1) a. *Ik heb het altijd naar mijn zin gehad,* ***dus*** *wil ik graag opnieuw beginnen.* (MC, 1995)
b. *Ik heb het altijd naar mijn zin gehad; ik wil* ***dus*** *graag opnieuw beginnen.*
'I was always content, so I would gladly start again.'

At first sight, the different positioning possibilities may seem arbitrary. However, Traugott & Dasher (2002: 158) claim they are not: "As has long been noted, the position of an adverb is correlated with difference of meaning (e.g. Greenbaum 1969, Jackendoff 1972, McConnell-Ginet 1982, Ernst 1984, Quirk et al. 1985, Cinque 1999, to mention only a few)." The hypothesis underlying their claim is attractive: a multifunctional word "exploits" its syntactic possibilities in order to differentiate between its functions. In this section, I will show that the general idea of interaction between positioning and function of connectives can be, and in fact has already been worked out in more detail.

Several authors have illustrated of individual connectives that there is a relation between positioning and function. For example, Auer & Günthner (2003: 2-4) have recently claimed that German *jedenfalls* 'anyway' can function in two ways. These different uses are illustrated in (2) and (3). As a sentential adverb in topic (i.e. clause-initial) position, *jedenfalls* has a modal meaning. As a word 'outside' the clause – set apart by intonation as well – it functions as a discourse marker that marks the return from a digression to the main line of reasoning (cf. also Schiffrin 1987; Polanyi 1988).[1]

(2) *Er ist ein gutter Arzt.* ***Jedenfalls*** *sagt man das.* (Auer & Günthner 2003: 2)
'He is a good doctor. At least that's what people say.'

(3) (Fragment from a phone call to a radio program, in which a listener tells his story)
(…) ja wie gehts (…) irgendwie öhm kann er da öh ich weiss nich(t);
in Frauenkrankheiten bin ich nicht so eh bewandert;
jedenfalls *– wir fahren da zu eh nach Hause zurück (…)* (Auer & Günthner 2003: 3)
'(…) yes how are you (…) somehow ehm he can oh I don't know;
I'm not familiar with women's illnesses / anyway / we drive there back home (…)'

These examples illustrate that the positioning of the word *jedenfalls* sets the connective use apart from the discourse marker use.

Lamiroy & Van Belle (1995: 412-416) present a second Germanic example, showing that Dutch *toch* 'still/after all' fulfills distinct argumentative functions in different positions (cf. also Elffers 1992 for a more detailed study of the use of *toch*). They distinguish between stressed *toch* in topic position and unstressed *toch* in clause-medial position. As a stressed adverbial in topic position (see (4)), *toch* functions as a concessive connective, equivalent to English 'still' or 'nevertheless'.[2] As an unstressed adverbial in clause-medial position (see (5)), it turns the clause into a rhetorical question, comparable to the English 'after all'.[3]

(4) *Deze keer ga ik Bob eens flink de waarheid zeggen. Ik had hem gewaarschuwd dat er vandaag files zouden zijn.* ***Toch*** *is hij weer te laat.*
'This time I am going to tell Bob a few home truths. I warned him there would be traffic jams today. Still he is late again.'

---

[1] In Polanyi's (1988) terms, *jedenfalls* functions as a *discourse POP marker*. "Discourse PUSH/POP markers signal the embedding, continuation and returns to and from discourse constituents at the various levels" (Polanyi 1988: 605-606; see also Polanyi & Scha 1983).

[2] The constructed examples in (4) and (5) are inspired by examples (32b) and (35b) in Lamiroy & Van Belle (1995: 412-413).

[3] According to Lamiroy & Van Belle (1995: 416), these two distinct argumentative functions of *toch* correlate with two sets of words in French, namely *cependent, pourtant* and *quand même, tout de même* respectively.

(5) *Deze keer ga ik Bob eens flink de waarheid zeggen. Ik had hem* ***toch*** *gewaarschuwd dat er vandaag files zouden zijn en hij is weer te laat.*
'This time I am going to tell Bob a few home truths. After all I warned him there would be traffic jams today and he is late again.'

The analyses of *jedenfalls* and *toch* both illustrate a distributional pattern in which positioning singles out the connective function from other text-linguistic functions these words may have, particularly discourse markers.[4] A similar claim can be found in Ariel's (1988, 1999) work. Ariel focuses on words that can function both as connectives and as a specific type of discourse markers, namely markers signaling that the presented information is already accessible to the reader or hearer. Her accessibility theory can best be explained against the background of Schiffrin's (1987) claims about the marking of *information states*. Schiffrin (1987: 28-29) distinguishes between so-called knowledge and meta-knowledge. Speaker/hearer *knowledge* concerns what a speaker knows and what a hearer knows; speaker/hearer *meta-knowledge* concerns what speakers and hearers know about their respective knowledge, and what parts of each knowledge one knows (or assumes to know) the other to share. Because discourse involves exchange of information, knowledge and meta-knowledge are constantly in flux. This makes it useful for speakers to signal certain changes in these interactively emerging information states.

Schiffrin (1987: 205-207) argues that speakers can select different connectives to display changes to one another in either knowledge or meta-knowledge. She shows, for example, that a connective like *because* marks a shift from unshared to shared knowledge. This connective "can be used to preface information when the status of that information as shared background knowledge is uncertain and when that information is important for understanding adjacent talk" (Schiffrin 1987: 207). The connective *so*, on the other hand, marks a change in meta-knowledge. It can be used to preface information when the understanding of that information is supplemented by information that has just become shared background.

Schiffrin (1987) notes that words like *because* and *so* function on several 'planes of talk'. As connectives they serve as words expressing causal coherence relations between clauses. As discourse markers – words expressing transitions in knowledge or meta-knowledge – they function on the 'textual' level (cf. Halliday & Hasan 1976; Degand 1996), or the level where "the information structure of sentences or clauses interacts with that of the discourse containing them" (Lang 2000: 253).

Ariel (1988, 1999) contributes to this discussion by investigating how the connective use of specific words can be distinguished syntactically from their use as discourse markers. Unlike Schiffrin, Ariel explicitly tries to relate these different functions to different positioning possibilities. According to Ariel (1988: 570-571), both syntactic factors as well as intonation distinguish the discourse marker use from other uses words like *harey* might have. Although the most prominent example in her analyses is the Hebrew word *harey*, I will

[4] In reaction to observations by Noach (1952), Huijsinga (1953) discusses a third Germanic example concerning the positioning of the Dutch combination *dan ook* 'and so/not surprisingly'. This word often functioning as a causal connective may even occur in the clause preceding the connective clause if it functions as a kind of discourse marker. This example is not discussed in greater detail here, since I will focus on positioning possibilities within the connective clause.

An example concerning a Romanic language can be found in Rossari & Jayez (1996), who investigate the interaction between the different interpretations and positionings of the French connective *donc* 'therefore'.

present Ariel's analyses of English *after all* to illustrate this position-function interaction.[5] Ariel (1988: 579-581) discusses two different uses of *after all*: firstly, as a word expressing a contrastive meaning (see (6)), and secondly, as an accessibility marker (see (7)). To put it in Schiffrin's terms, *after all* can also function as a marker of meta-knowledge.[6]

(6) A: Did you know that John went to America **after all**? (Ariel 1988: 579)
B: No, I didn't. I am glad for him.

(7) The existence of such individuals is PRESUPPOSED (we cannot **after all** identify something that does not exist). (Ariel 1988: 579)

The two uses of *after all* have different positional and intonational properties. Contrastive *after all* is stressed, and typically occurs at the end of the clause. Functioning as an accessibility marker, *after all* typically occurs in a destressed position, either clause-initial or clause-medial (Ariel 1988: 580). These different distributional properties are also supported by the examples in (8): the intended contrastive meaning cannot be maintained in clause-initial position.

(8) a. Can I see Jane now, **after all**!
b. #**After all**, can I see Jane now? (Ariel 1988: 580)

Ariel's claim about the interaction between positioning of a word and its respective meanings seems promising. Because she (1988: 570) claims that accessibility markers can be found in a variety of languages (mentioning examples from languages like German, Russian, Swedish, Polish, Hungarian, and Armenian), it seems attractive to investigate whether the Dutch language also shows such a syntactic division of labor between the connective and discourse marker function of words. This interaction will be investigated further in Chapter 7, which discusses the conceptual and syntactic properties of the Dutch adverbials *dus* and *daarom*.

---

[5] A complicating factor in Ariel's (1988) discussion of *harey* and other Hebrew words that signal the accessibility of information is that her use of the term 'connective' is confusing. The quote in (i) as well as the syntactic tests on page 571 seem to support the idea that Hebrew 'connectives' behave differently from Hebrew discourse markers.

(i) "Another feature distinguishing *harey* words from the various conventional connectives is their relative freedom to occur in different positions within the sentence. Connectives are limited to initial position." (Ariel 1988: 571)

However, the quote in (ii) indicates that Ariel uses the term 'connective' to refer to the syntactic class of complementizers.

(ii) "Though they [= *harey* words] have a whole sentence as their scope (like the connectives) their syntactic behaviour is that of an adverb, and not that of a connective." (Ariel 1988: 571-572)

This quote implies that Ariel's syntactic tests on page 571 only illustrate that the words functioning as accessibility markers should be seen as sentential adverbials instead of complementizers; the quote does NOT imply that connectives, defined as markers of coherence relations, show a distribution different from discourse markers.

[6] Note that *after all* marks the meta-knowledge itself, and not a shift in this meta-knowledge. This means that Schiffrin's (1987) theory about markers of knowledge and meta-knowledge should be extended: there are words to mark shifts in knowledge or meta-knowledge, but also words that signal the (meta-)knowledge itself.

3.3 Interactions with the linearization of connective clauses

Several authors have claimed that the second syntactic primitive, the linearization of adverbial clauses, is also dependent on discourse factors (cf. Chafe 1984; Thompson 1985; Givón 1987; Ramsay 1987; Prideaux 1989; Ford 1993; Auer 2000). As Ramsay (1987: 384) puts it: "(…) the positioning of some types of adverbial clauses is dictated by the organization of information in discourse". But what is the precise nature of the relation between linearization and discourse? In this section I discuss the ideas of several authors who studied this interaction, and have come up with highly similar proposals.

Several authors have claimed that preposed or initial adverbial clauses represent a limitation of focus (cf. Chafe 1984; Thompson 1985; Ramsay 1987; Auer 2000). According to Chafe (1984) preposed clauses signal a path of orientation in terms of which the following information is to be understood. This 'grounding' function of initial clauses has also been described by Auer (2000). His analyses of German *wenn*-clauses show that *wenn*-clauses with a conditional interpretation are preferred in clause-initial position in spoken German (Auer 2000: 9). Auer (2000: 14) argues there seems to be some kind of cognitive 'naturalness' in the way in which initial conditionals "create the ground – or, in more recent but equally metaphorical parlance, set up a 'mental space' (Fauconnier 1985) – in which some hypothetical or factual proposition is located".[7] For cognitive reasons, it is the grounding that (iconically) precedes the focal proposition, and not the other way around. Similar findings are reported on the use of *if*-clauses in English conversations (cf. Ford & Thompson 1986: 362; Ford 1993: 24). Ford (1993: 15) suggests that "the prevalence of initially placed *if*-clauses may reflect the general tendency to signal (…) that the interpretation of the coming clause will be, in some general way, limited by the contents of the *if*-clause."

These authors also agree on the function of postposed or final adverbial clauses. A postposed clause only adds something to the assertion made by the main clause or modifies part of what is stated there. Thompson (1985: 69-70) formulates the contrast between final and initial (in her case purpose) clauses as follows: "while the final purpose clause performs a local semantic function, naming the motivation for some particular action, the initial purpose clause plays a much broader discourse role, guiding the reader's attention by serving as a link in an expectation chain." The two functions can be illustrated by the italicized purpose clauses in (9) and (10). The initial purpose clause in (9) signals how the reader is expected to associate the material following the purpose clause with the material preceding it. More specifically, the initial clause in (9) functions to state a 'problem' within the context of expectations raised by the preceding discourse, to which the following material (in this case more than one clause) provides a solution. The final purpose clause in (10) does *not* play this role (cf. Thompson 1985: 61). It plays the much more local role of stating the purpose for which the action named in the immediately preceding clause is performed.

(9) (Section on 'About Yeast Bread Making') Meanwhile get your pans ready. Glass and enamel pans require a lower temperature than darkened tin or dull aluminum ones. Any of these will give you a well-browned crust. *To form the loaf*, throw down onto the board one of the pieces of dough which have been resting. You may use a rolling pin or your palm to press it evenly before forming. (…) (Thompson 1985: 66)

[7] See Dancygier & Sweetser (1996, 2000) for a more detailed account of the linearization of *if*-, *since*-, and *because*-clauses based on mental spaces.

(10) George had always been my first choice for crew. Twenty-six years old, he had served in the army and later gone to the Middle East *to train soldiers for an oil-rich sheik*. With the money saved from this venture, he had decided to take a couple of years looking around the world and pleasing himself. (Thompson 1985: 68)

The different functions of initial and final clauses can explain why purpose clauses cannot always be freely transposed from one position to the other. For example, Givón (1990: 837) argues that the purpose in (11)a is not that of the main-clause subject (the intended reader of the page), but rather the implicit writer. Since post-posed purpose clauses typically code the intent of the agent of the main clause, (11)b is odd. Similarly, it is fairly clear that the purpose in (12)a is that of the main-clause subject. Presumably for that reason, (12)b is odd.

(11) a. *To illustrate this*, consider the following page.
b. #Consider the following page *to illustrate this*. (Givón 1990: 837)
(12) a. He went *to fix the plumbing*.
b. #*To fix the plumbing* he went. (Givón 1990: 837)

The conceptual difference between initial and final adverbial clauses is supported by several text-linguistic data. For example, Thompson provides two sources of data that support the more global discourse character of initial purpose clauses. Firstly, she observes that initial clauses can have many clauses or even sentences in their scope (Thompson 1985: 75-76). That is, as in example (9), the material describing the solution to the problem named by the purpose clause is often expressed by means of several clauses. In contrast, none of Thompson's final purpose clauses has more than a single clause in its scope, namely its immediately preceding main clause. Secondly, Thompson's claim that final purpose clauses are more closely tied to their main clause is supported by her finding that very few final clauses in her corpus (less than 1%) are separated from their main clauses by a preceding comma, while 78% of the initial purpose clauses are separated from their following main clause by a comma. Ramsay (1987) provides additional quantitative evidence in favor of the local nature of postposed clauses versus the more global nature of preposed ones. Her analyses of spoken and written *if*- and *when*-clauses show that (a) postposed clauses have a much higher referential continuity with the main clauses than preposed clauses (Ramsay 1987: 391); (b) the scope (the number of clauses needed in order to understand the whole clause) of final *if*- and *when*-clauses is much more restricted than that of initial clauses (p. 395); and (c) commas are used much more frequently with preposed clauses than with postposed clauses (p. 401).

From the previous discussion it can be concluded that the second syntactic primitive, linearization, can be related to the organization of discourse: initial adverbial clauses create a ground, restricting the way in which the upcoming clause(s) should be interpreted; final adverbial clauses, on the other hand, only modify the preceding clause. This interaction has been formulated on the basis of evidence from English and German. However, it has also been applied to the Dutch language. In Degand's (2001: 105-108) study of adverbial clauses containing the causal connective *omdat*, most of Thompson's observations were confirmed.[8] In 71% (12 out of 17) of the preposed *omdat*-clauses there is no immediate reference to the preceding clause. In contrast, most of the postposed *omdat*-clauses (84%, i.e. 97 out of 116)

[8] Another investigation into the linearization of Dutch connective clauses is reported in Noordman & Van Rijswijk (1997). They discuss the linearization of adverbial clauses headed by *hoewel* 'although'.

"demonstrate a tight semantic and referential dependency on the immediately preceding main clause" (Degand 2001: 107). For example, the preposed *omdat*-clause in (13) does not contain referential elements that are mentioned in the preceding sentence. The postposed *omdat*-clause in (14), on the other hand, repeats the subject (*het* 'it') mentioned at the beginning of its matrix sentence.

(13) *Omstreeks kwart voor vijf gistermorgen ontdekte een toevallig passerende politiepatrouille dat er rook kwam uit een woning op de vierde verdieping van het flatgebouw in de wijk Heerlerbaan. De agenten alarmeerden de brandweer, en gingen op onderzoek uit.* ***Omdat*** *herhaald kloppen en bellen geen reactie uit de woning opleverde, werd de voordeur geforceerd.* (Degand 2001: 107)
'Yesterday, around 4:45 AM, a police patrol discovered smoke coming out of a flat on the fourth floor of a building at Heerlerbaan. The officers called the fire brigade and went to investigate. Because repeated knocking and ringing didn't bring any reaction from the flat, the front door was forced open.'

(14) *Asbest is ongevoelig voor de sterkste zuren en basische stoffen en buitengewoon slijtvast. Het isoleert uitstekend warmte en electriciteit, is zeer trekvast en vooral onwaarschijnlijk goedkoop,* ***omdat*** *het zo gemakkelijk te winnen is.* (Degand 2001: 107)
'Asbestos is resistant to the strongest acids and alkalines and it is exceptionally durable. It is a very good insulator from heat and electricity, it has important tensile strength and it is incredibly cheap, because it is so easy to extract.'

Degand concludes from her observations that the linearization of *omdat*-clauses can be functionally explained: the postposed clauses play a more local role than the preposed ones.[9]

It can be concluded that the analyses of different types of adverbial clauses show ample support for an interaction between linearization and the organization of discourse. A clause in post-position is not merely a positional variant of a preposed one (and thereby always freely interchangeable with it) but a tool 'designed' for specific contexts of usage. Initial adverbial clauses serve a grounding function, guiding the interpretation of the upcoming clause(s), whereas final adverbial clauses serve a more local semantic function. Since this hypothesis is corroborated for Dutch *omdat* as well, the interaction between linearization and text-linguistic function will not be taken up for further investigation.

3.4 Interactions with the word order within matrix clauses

The third syntactic primitive, word order within postposed matrix clauses (see section 2.3.4), can be related to certain text-linguistic properties as well. A frequently cited formulation concerning this form-function parallelism can be found in the work of Givón (1990). He suggests of clause combining in general: "The more two *events/states* are integrated semantically or pragmatically, the more will the *clauses* that code them be integrated grammatically" (Givón 1990: 826).

---

[9] In other work (Degand 2000), Degand discusses contextual constraints on causal sequencing in general. In that article, she focuses on the text-linguistic factors that account for the choice of either cause-consequence or consequence-cause, irrespective of the connective that is used to signal the causal link. Her main conclusion is that causal sequencing cannot be linked to the temporal order of the denoted states of affairs, but that it is sensitive to discursive constraints of information flow in the discourse. See also Renkema (1996) on Dutch *omdat* versus *Omdat*.

König & Van der Auwera (1988) explicitly propose an interaction between word order within the postposed matrix clause and conceptual properties. In their cross-linguistic study they investigate the degree to which preposed conditional, concessive, and concessive conditional clauses are syntactically incorporated into their following main clause. Using word order to measure this degree of integration, they (1988: 127) argue that there are three ways of combining Dutch or German adverbial clauses with a following main clause (see section 2.3.4 and examples (15)-(17), taken from König & Van der Auwera 1988: 102-103).

(15) Non-integrative word order (adverbial clause - topic - finite verb - …)
D: *Zelfs als ze niet allemaal gekomen zijn, we kunnen met het bezoek tevreden zijn.*
G: *Selbst wenn sie nicht alle gekommen sind, wir können mit dem Besuch zufrieden sein.*
Even if they not all come are, we can with the attendance content be.
'Even if they haven't all come, we can be pleased with the attendance.'

(16) Resumptive word order (adverbial clause - resumptive - finite verb - …)
D: *Als hij ziek is, dan blijft hij thuis.*
G: *Wenn er krank ist, dann bleibt er zu Hause.*
If he sick is, then stays he at home.
'If he is sick, then he stays at home.'

(17) Integrative word order (adverbial clause - finite verb - …)
D: *Omdat hij ziek is, kan Fred niet meekomen.*
G: *Weil er krank ist, kann Fritz nicht mitkommen.*
Because he sick is, can Fred not come along.
'Because he is sick, Fred cannot come along.'

The three word orders show an increasing degree of syntactic integration. Looking at the position of the finite verb in the matrix clause, only (17) can be considered as a clear case of embedding in the sense that it functions as a constituent within another clause. In this example, the finite verb shows up in the initial position of the main clause, thus taking second position in the overall complex sentence. The other two adverbial clauses cannot be assumed to function as the first constituent of the second clause. In the matrix clause of (16), the finite verb is preceded by a resumptive (or correlative) element, resulting in a kind of left dislocation of the adverbial clause. The matrix clause in (15) exhibits a non-integrative verb-second word order, which is characteristic of independent sentences (topic – finite verb).

The data in König & Van der Auwera on present-day Dutch show that the different syntactic possibilities can be related to certain conceptual properties.[10] For instance, in

[10] König & Van der Auwera (1988) assume the three types of word order "are linked as stages in a historical development leading from non-integrative via resumptive to integrative word order" (p. 107). They also claim that German and Dutch have undergone such a process of successive clause integration (p. 108), which Lehmann (1988: 216) labels *compression*. Despite the diachronic range of their hypothesis, König & Van der Auwera focus on the present-day distribution of the three clause types.

The evidence they derive from other literature in support of their hypothesis is the observation that there is a diachronic decline in the frequency of the resumptive word order in both Dutch and German concessive conditionals and concessives (p. 107-108). However, a decrease in the resumptive word order pattern in favor of an increase in the integrative pattern does not necessarily prove their claim. This only proves that the distribution over the three word orders changes, but not that one of the word orders is prior to another. Decisive evidence should come from data showing (a) that there was a period in which the Dutch and German language did contain the resumptive, but not the integrative

conditional clauses the non-integrative word order can only occur in combination with adverbial clauses that are separately assertable or adverbial clauses that function as speech act qualifiers (König & Van der Auwera 1988: 110-115). An example of the latter type is given in (18), where the preposed adverbial clause relates to the speech act performed in uttering the following clause, rather than to the proposition expressed by it.

(18) *Als iemand mij zoekt, ik ben in de bibliotheek.* (König & Van der Auwera 1988: 110)
If someone me searches, I am in the library.
'If someone is looking for me, I am in the library.'

The absence of syntactic incorporation into the matrix clause is not surprising here, since the adverbial clause in (18) is *about* rather than *part of* the following clause.

The relation between domains of use and word order within the postposed matrix clause observed by König & Van der Auwera has also been reported in Van Belle (1997). His data on conditional clauses and connectives show that speech-act clauses can either occur with a non-integrative word order (see the a-example in (19)) or with a resumptive or integrative word order (see the b-sentence there). Content conditionals like (20) and epistemic conditionals as in (21), on the other hand, can either occur in combination with the resumptive or the integrative word order (see the a-variants), but not with the non-integrative word order (see the b-variants).

(19) a. *Als je hem wil spreken, hij is in de tuin.* (Van Belle 1997: 219)
'If you want to speak to him, he is in the garden.'
b. *Als je hem wil spreken, (dan) is hij in de tuin.*
(20) a. *Als je die draad aanraakt, (dan) krijg je een elektrische schok.* (Van Belle 1997: 218)
If you that wire touch, (then) get you an electric shock.
'If you touch that wire, (then) you'll get an electric shock.'
b. *#Als je die draad aanraakt, je krijgt een elektrische schok.*
(21) a. *Als de druiven zo goed gelukt zijn, (dan) moet het een warme zomer geweest zijn.* (Van Belle 1997: 219)
If the grapes so well succeeded are, (then) must it a warm summer been are.
'If the grapes came off so well, (then) it must have been a warm summer.'
b. **Als de druiven zo goed gelukt zijn, het moet een warme zomer geweest zijn.*

These data show that the word order within the postposed matrix clause can be linked to the interpretation based on domains, although this mapping is not bi-directional (a speech-act interpretation does not exclude an integrative or a resumptive word order).

In conclusion to this section, the word order within a postposed matrix clause can be used to measure the degree of syntactic integration of the preposed connective clause. The selection of this word order can – at least in the case of conditionals, concessives and concessive conditionals – be related to the text-linguistic characterization based on domains. Content and epistemic conditionals can only show an integrative or a resumptive word order, whereas speech-act conditionals allow for all three word order patterns. In Chapter 6, this interaction will be tested for the Dutch complementizer *omdat*.

---

word order, and (b) that the rise of the integrative word order is linked to the use of the resumptive word order before that time.

3.5 Interactions with the word order within connective clauses

The word order within the connective clause itself can also be related to certain text-linguistic properties. Section 3.5.1 introduces hypotheses about the interaction between word order and domains of use. Section 3.5.2 deals with the work of Verhagen (2001), who explicitly rejects this interaction and proposes in its stead an interaction between word order and the hierarchical organization of text segments. Section 3.5.3 shows a comparison of the two proposals.

3.5.1 Word order related to domains of use

The interaction between word order and domains of use can be illustrated by the recent development of the German connective *weil*, as discussed in Günthner (1993, 1996) and Keller (1995).[11] Formerly, *weil* could only be used in clauses showing a subordinating V-late word order (see (22)). In present-day spoken German, it can also occur at the head of a clause showing verb-second, the word order common to main clauses (see examples (23) and (24)).

(22) *Er ist nach Hause gegangen,* ***weil*** *er Kopfweh* <u>*hatte*</u>*.* (Keller 1995: 20)
He is to home gone, because he headache had.
'He went home, because he had a headache.'

(23) *Er ist nach Hause gegangen,* ***weil*** *ich* <u>*sehe*</u> *seinen Mantel nicht mehr an der Garderobe.* (Keller 1995: 22)
He is to home gone, because I see his coat not anymore on the hall tree.
'He went home, because I don't see his coat anymore on the hall tree.'

(24) *Er ist nach Hause gegangen,* ***weil*** *er* <u>*hatte*</u> *Kopfweh.* (Keller 1995: 21)
He is to home gone, because he had headache.
'He went home, because he had a headache.'

These analyses of German *weil* give concrete evidence in line with the hypothesis that there is indeed a relation between word order and use in a certain domain: the content use of *weil* appears in a subordinating construction (see example (22)), whereas the main clause word order is reserved for epistemic and speech-act use of this connective (see (23) and (24). Both Günthner and Keller attribute the word order change to specific text-linguistic properties of the new use of *weil*. As Keller (1995: 25) formulates it: "The shift from subordinate to main-clause order in *weil*-clauses (…) is the consequence of a semantic change from factual to epistemic meaning. The epistemic reading demands that the proposition of the *weil*-clause is not presupposed, but stated, and consequently it demands that word order is not hypotactic, but paratactic."[12]

In her article, Günthner (1996) also discusses recent observations on syntactic changes in clauses containing *obwohl* 'although' or *wobei* 'whereby'. Like *weil* these traditional subordinators sometimes show a main clause V2 word order in spoken German. Günthner argues that this non-standard word order is restricted to specific domains: the speech-act and the epistemic domain. The examples in (25) and (26) illustrate the speech-act usage of both connectives. In (25), Kati's speech act – the question *Kommt er eigentlich au(ch)* 'is he coming too?' – is withdrawn in the remainder of the utterance.

---

[11] See also Gohl & Günthner (1999), Pasch (1997), Uhmann (1998) and Wegener (1993).

[12] Keller (1995: 27) explains this new use as a form of compensation: *weil* adopts the function as well as the particular word order of *denn*, which belongs to the literary language.

(25) Kati: *Kommt er eigentlich au?* (Günthner 2000: 447)
Comes he actually too?
'Is he coming too?'
Kati: ***Obwohl**, s isch mir eigentlich egal.*
Although, it is me actually even.
'Although, I don't care really.'

(26) (In answer to the question "Can you come to the dinner?")
*Ja. **Wobei** ich hab am frühen Abend ne Univeranstaltung und weiß nicht genau wann die zuende ist.* (Günthner 1996: 348-349)
'Yes. Whereby I have a seminar at the university in the early evening and don't know exactly when it will be over.'

So, Günthner claims the syntactic difference between V-late and V2 is being used to discriminate between content use of *obwohl* and *wobei* on the one hand and epistemic and speech-act use of these connectives on the other.[13]

A similar form-function parallelism between word order and domains of use can be found in typological research concerning connective clauses. Crevels (2000) offers an implicational semantic hierarchy for concession, namely Content > Epistemic > Illocutionary > Textual. She hypothesizes that the higher the semantic level, the more likely it is that a clause becomes less and less integrated into its main clause and more and more paratactic-like (Crevels 2000: 321).

From the above data, it can be concluded that the hypothesis about the mapping between word order and domains of use (or source of coherence) is plausible. Recently, Verstraete (1998, 2000) and Haegeman (2001, 2003) have worked out the link between the source of coherence and the syntactic properties of adverbial clauses in more detail. Both authors discuss syntactic characteristics other than word order, which is the logical result of the fact that they study English, a language that does not discriminate between so-called 'subordinate' and 'main clause' word orders. Their ideas are highly similar (see Chapter 6 for a more elaborate discussion), despite the fact that Haegeman works within a generatively oriented framework and Verstraete within the tradition of functional grammar. Here, I will introduce Haegeman's theory as an indication of the form-function relations both authors propose.

Haegeman especially focuses on conditionals, but also applies her line of reasoning to adverbial clauses headed by connectives like *while*, *because*, *when*, *since*, *as*, and *so that* (Haegeman 2003: 318). Following standard generative assumptions, she (2003: 325-326) assumes that clauses consist of three layers: (i) the thematic domain, which is labeled Verb Phrase (VP) and is headed by the predicate; (ii) the functional domain, summarized as IP, in which functional categories such as time, mood, and aspect are licensed; and (iii) the periphery of the clause, centered around the subordinating conjunction, and labeled Complementizer Phrase (CP). She furthermore assumes that each layer (VP, IP, CP) decomposes into a sequence of layered projections (compare Rizzi's (1997) 'split CP' for similar claims concerning the CP layer). Haegeman uses this basic analysis of clauses to account for the different syntactic behavior of semantic and pragmatic adverbials. She argues

[13] In a more recent article, Günthner (2000) introduces a different analysis of the syntactic change of *obwohl*. She claims that the V2-use of *obwohl* is preserved for a new function as 'discourse marker' with a corrective instead of a concessive meaning. Because Günthner's definition of discourse marker is not clear to me (e.g. it is not clear whether it includes the category of connectives or not), this section resorts to her earlier (1996) analysis.

that the two types of adverbials differ in their external syntax (the layer of the matrix clause to which they attach) and their internal syntax (the layers they consist of themselves).

Haegeman uses different terms for the conceptual distinction based on source of coherence. The bipartition semantic-pragmatic turns up in her distinction between adverbials relating to event structure on the one hand (i.e. semantic use) and discourse structure on the other (i.e. pragmatic use). She provides several syntactic scope tests to argue that both types of adverbials differ in their external syntax (Haegeman 2003: 320-324). According to her, semantic adverbials are more closely integrated with their matrix clause than pragmatic adverbials. Semantic adverbials are 'central'; in generative terms they "are adjoined to a projection of the matrix V or to a higher projection between VP and the surface position of the subject", or in minimalist terms "they are inserted in (or merged with) the matrix clause early on in the derivation; they are merged before IP is completed" (Haegeman 2003: 326). Pragmatic adverbials, on the other hand, remain 'peripheral' to their associated CP; they "are adjoined to the associated CP" or "merged after the associated CP has been completed" (Haegeman 2003: 326).

In addition to the difference in external syntax, Haegeman (2003: 330-336) argues that semantic and pragmatic adverbials differ in their internal make-up. Her proposal is that the two types of clauses differ in the number of functional projections on top of the VP and IP. She claims that pragmatic adverbials contain a complete CP with independent illocutionary force. Semantic adverbials contain an incomplete or so-called truncated CP; there is no functional projection that encodes illocutionary force, and it also lacks the functional projections that enable topicalization and focalization (see also Haegeman 2001 for a more elaborate discussion).[14]

To conclude this section, there is evidence for an interaction between domains of use and specific syntactic properties of the adverbial clause. The German data presented by Günthner (1993, 1996) and Keller (1995) indicate a link with word order. In addition, Haegeman's proposal can be characterized as a successful attempt to show that there is a link between source of coherence on the one hand and the internal and external syntax of the adverbial clause on the other.

3.5.2 Word order related to hierarchical text structure

In the previous subsection, the word order properties of connective clauses were related to their interpretation based on domains. The domain interpretation attributes to the specific nature of the coherence relation that should be derived between two text segments. Apart from this relational aspect of text structure, Sanders & Van Wijk (1996: 93) also distinguish a hierarchical aspect of text structure (cf. also Matthiessen & Thompson 1988; Polanyi 1988). This aspect handles the structure of the text as a whole and the hierarchical relations between segments. Connectives also seem to signal both of these structural aspects. As Verhagen (2001: 118) formulates it, connectives incorporate two functions, which are both relevant for the construction of a text representation. Their first function is to provide information on the content of the coherence relation to be derived. The second function is to provide information on the possible text segments to which the connective clause can be related.

Verhagen (2001) provides evidence that – at least for two modern Dutch connectives – the word order properties should not be related to the content of the relation they express, but

[14] This idea of a truncated CP is in line with Hooper & Thompson (1973: 484-485, in Haegeman 2003: 331), who discuss 'reduced clauses' in which "main clause phenomena" like topicalisation and focalisation cannot be applied.

to the hierarchical nature of that relation. According to him, the syntactic difference between the Dutch causal coordinator *want* (which shows V2) and the subordinator *omdat* (which shows V-late) cannot be explained by reference to the way they are used in terms of domains. In line with a large amount of literature (starting with De Vries in 1971 and culminating in Pit's dissertation in 2003) Verhagen admits that the content of the coherence relation is not exactly the same for *want* and *omdat* (although both connectives express causal coherence relations). Nevertheless, he claims that the two word order constructions *want* and *omdat* are conventionally associated with are the indicators of the second type of information: information on the possible text segments to which the connective clause can be related.

Verhagen treats V2 and V-late as two specific construction types whose meaning can only be distinguished using combinations of more than two clauses. The contrast between the following examples illustrates this point.

(27) [**A** *Het had die nacht geijzeld;*] [**B** *'s ochtends stormde het hard,*] [**C** *waardoor veel mensen te laat op hun werk* <u>*kwamen*</u>.] (Verhagen 2001: 118)
'[**A** It had frozen over that night;] [**B** in the morning a storm was blowing up,] [**C** which is why many people arrived late at work.]'

(28) [**A** *Het had die nacht geijzeld;*] [**B** *'s ochtends stormde het hard;*] [**C** *daardoor* <u>*kwamen*</u> *veel mensen te laat op hun werk.*] (Verhagen 2001: 118)
'[**A** It had frozen over that night;] [**B** in the morning a storm was blowing up;] [**C** that's why many people arrived late at work.]'

The subordinating word order of the C-clause in (27) presents the late arrival as the consequence of the storm only. This is an example of the so-called 'late closure' processing strategy (Verhagen 2001: 114): the new element to be processed (the C-clause) is related to the last processed element (the B-clause). However, the main clause word order of the C-clause in (28) does not indicate this restriction in the interpretation of the coherence relation. Here, the late arrival can also be the consequence of the glaze and the storm together. Word order then can be used to indicate the level of attachment of the adverbial clause: V-late restricts this level of attachment to the previous clause (it indicates late closure), whereas V2 does not impose such a restriction. This also holds for the following two examples, containing *want* and *omdat* respectively.[15] In this clause combination [A-B-C], the subordinating *omdat*-clause (C) in (29) can only be related to clause B. This interpretation can be hierarchically represented as [A-[B-C]]. If the same information is connected by the coordinator *want*, as in (30), the C-clause can be interpreted as relating either to B, resulting in the hierarchy [A-[B-C]], or to the clause combination A-B, resulting in the hierarchy [[A-B]-C].

(29) [**A** *Getuigeverklaring X mag niet in de beschouwingen betrokken worden.*]
'Testimony X may not be taken into consideration.'
[**B** *Verder moeten de stukken Y en Z uit het dossier verwijderd worden,*]
'Furthermore, documents Y and Z have to be removed from the dossier,'
[**C** ***omdat*** *vroegere veroordelingen in de lopende zaak geen rol* <u>*mogen*</u> *spelen.*]
because earlier convictions in the current case no role may play
'because earlier convictions may not play a role in the current case.'

[15] The examples in (29) and (30) are slightly modified versions of the examples in Verhagen (2001: 117).

(30) [**A** *Getuigeverklaring X mag niet in de beschouwingen betrokken worden.*]
'Testimony X may not be taken into consideration.'
[**B** *Verder moeten de stukken Y en Z uit het dossier verwijderd worden.*]
'Furthermore, documents Y and Z have to be removed from the dossier,'
[**C** ***Want*** *vroegere veroordelingen* <u>*mogen*</u> *in de lopende zaak geen rol spelen.*]
Because earlier convictions may in the current case no role play-infinitive.
'because earlier convictions may not play a role in the current case.'

In conclusion, Verhagen's (2001: 111) proposal is that the syntactic difference between *want* and *omdat* has nothing to do with the content of the relation between two text segments, but rather that it is related to the structural organization of a text.

3.5.3 Comparison of the two word order proposals

In section 3.5.1 and 3.5.2 the fourth syntactic primitive, word order within the connective clause itself, has been related to two different text-linguistic properties of adverbial clauses: (a) their interpretation based on domains of use or source of coherence, and (b) their level of attachment to the preceding discourse. Haegeman (2003) has shown that a further link can be established between source of coherence and syntactic properties of the adverbial clause (see also Verstraete 1998). How do the two proposals relate to each other? Do they exclude each other or are they complementary?

It can be argued that the two theories do not necessarily exclude each other. In principle, it is perfectly possible that word order expresses both conceptual aspects. For example, the data on German *weil* allow for a combination of the two conceptual properties. As is already shown in section 3.5.1 the word orders accompanying this connective can provide information on the domain type. Further research is needed to test the hypothesis about the interaction between word order and level of attachment of the *weil*-clause. For example, an investigation of the German equivalents of (29) and (30) and their possible interpretations may provide evidence for this interaction.

This leaves open the question whether these interactions also apply to the Dutch complementizers that are selected in this thesis. Verhagen's recent account of the interaction between word order and the structural organization of the text includes some convincing examples on *want* and *omdat*, but his theory has not been supported with quantitative data yet. His theory, then, will be put to the test in Chapter 5, in which I present qualitative as well as quantitative data on the diachronic development of *want* and *omdat*.

The hypothesis about the interaction between word order and domains of use also demands further investigation. The data on German seem promising, but this does not necessarily imply that the interaction holds for the Dutch language as well. In fact, it is even explicitly rejected by Verhagen. On the basis of previous literature (e.g. Degand 2001; Pit 2003), it can be concluded that the interaction is not exactly the same as the German interaction between word order and domains of use. For example, Degand (2001: 104, 109) argues that both *want* and *omdat* can be used to express content (see (31)) and (non-abductive) epistemic relations (see (32)).[16] The connectives differ in that only the coordinator *want* can mark speech-act relations (compare the a- and b-version in (33)).

[16] Note that the antecedent-clause in (32) presents a real-world cause as an argument for the conclusion that it must be heating up.

(31) a. *Bob viel,* ***want*** *hij gleed uit over een banaan.*
b. *Bob viel,* ***omdat*** *hij over een banaan uitgleed.*
'Bob fell, because he slipped over a banana.'
(32) a. *Het moet wel warmer worden,* ***want*** *de zon schijnt.*
b. *Het moet wel warmer worden,* ***omdat*** *de zon schijnt.*
'It has to be heating up, because the sun is shining.'
(33) a. *Is Mirjam nog thuis?* ***Want*** *haar lampen zijn nog aan.*
b. *Is Mirjam nog thuis?* ****Omdat*** *haar lampen nog aan zijn.*
'Is Mirjam still home? Because her lights are still on.'

Chapter 6, then, will focus on the specific nature of the Dutch interaction between word order and domains of use. It will also follow up on the syntactic analyses of Haegeman and Verstraete.

3.6 Interactions with subjectification

The current section introduces a form-function interaction specific to diachronic developments: the interaction between so-called grammaticalization (a process that is often accompanied by syntactic changes) and subjectification. The diachronic development of the connectives can be placed in a grammaticalization framework, a theory that starts from the idea that lexical elements can develop into functional elements and that existing functional elements can develop new grammatical functions (see, among others, Heine, Claudi & Hünnemeyer 1991: 4; Hopper & Traugott 1993: 94; Bybee, Perkins & Pagliuca 1994: 13).[17] Such diachronic developments often involve one or more syntactic changes. Examples of such changes are the affixation or fusion of the grammaticizing material to surrounding material (cf. Bybee et al. 1994: 106; Hopper & Traugott 1993: 40), changes in the accompanying word order (cf. Günthner 1996; Keller 1995), and changes in categorical status (cf. Dubinsky & Williams 1995). This latter type of change is often considered a case of reanalysis (cf., among others, Abraham 1993; Langacker 1977; Weerman 1989). According to Hopper & Traugott (1993: 32), reanalysis modifies the underlying (syntactic or morphological) representations and brings about rule change. Typical examples in the area of connectives are the reanalysis of *by cause of* into *because* and the change of the adverbial phrase *þa hwile þe* ('at the time that') into the complementizer *while*.

A large number of these grammaticalization phenomena goes hand in hand with a specific change in meaning: so-called 'subjectification', an increase in subjectivity (cf. Traugott 1995: 31). In Langacker's words: subjectification "is a recurrent and highly important type of semantic extension and is often a central factor in the evolution from 'lexical' to 'grammatical' elements" (1990: 5).[18] Traugott (1995: 46-47) also argues that an increase in subjectification is a characteristic property of grammaticalization. As she formulates it in Traugott & König (1991: 198), there is a tendency to change "from meanings grounded in more or less objectively identifiable extralinguistic situations to meanings grounded in the speaker's attitude to or a belief about what is said" (cf. also Sweetser 1990). Thus, the subjectification hypothesis focuses on the speaker as the origin of linguistic change;

[17] See Traugott (2001: 1) for some remarks on the precise formulation of a definition of the term 'grammaticalization'.
[18] Keller correctly notices: "subjectification is the name of a descriptive generalisation of different kinds of semantic changes which need explanation" (1995: 17).

it "implies some degree of integration of the perceiver in the description of an object or a process" (Cuenca 1997: 5).

Is there any evidence for this idea in the area of diachronic connective developments? Traugott (1995) shows that a process of subjectification took place in the diachronic change of English *while*. The first case of subjectification is the change of the adverbial phrase *þa hwile þe* ('at the time that') into the temporal connective *while*. Instead of profiling a specific time, *while* now profiles discourse structure, which is the responsibility of the speaker. The second case of subjectification is the development of temporal *while* into concessive *while*. The new use construes a world that has no reference in the described situation, but only in the speaker's world or belief about coherence. Traugott (1995: 39) even claims: "historically almost all grammatical markers of clause combining have developed out of a more 'objective' function" (see also Dasher 1995). For example, many temporal, causal and conditional connectives have grown out of adverbial constructions (Genetti 1991). This subjectification theory can furthermore be applied to the changes German *weil*, 'because', underwent (see Keller 1995; Günthner 1996).

Looking at other connective data investigated so far, the subjectification hypothesis indeed seems to hold for three types of conceptual change connectives might go through: (1) the actual rise of connectives (cf. the rise of English *while*), (2) changes from one connective function into another (cf. the changes in the connective use of English *while* and German *weil*), and (3) shifts away from the connective usage. Two of these developments involving subjectification (1 and 3) are conceivable for changes in the use of the Dutch word *nu* 'now'.

(34) a. <u>*nu* as deictic element:</u>
Ik wil graag ***nu*** *over dat probleem praten.*
'I would like to talk about that problem now.'
b. <u>*nu* as temporal connective:</u>
Nu *ik je zie, denk ik er weer aan dat ik met jou over dat probleem wilde praten.*
'Now that I see you, I remember I wanted to talk to you about that problem.'
c. <u>*nu* as discourse marker:</u>
De oplossing van <u>*dat*</u> *probleem ken ik, maar hoe zit het* ***nu*** *met* <u>*dit*</u> *probleem?*
'I know the solution to that problem, but now what about this problem?'

In terms of Schiffrin (1987: 230), who gives a similar analysis for English *now*, the three uses of the word *nu* in (34) give a description of reference time, event time, and discourse time respectively. These three uses differ in their degree of subjectivity, i.e. in the degree to which they are grounded in the speaker. In the a-sentence *nu* refers to a specific point of time in reality; the word has a purely objective, deictic function, in which the attitude of the speaker does not play a role. In the b-sentence *nu* marks the temporal relation between two propositions. On the one hand, this temporal relation is a description of the temporal sequence in the real world, but, on the other hand, it is a temporal coherence relation constructed by the speaker for communicative purposes. Therefore, this use of *nu* is more subjective than the deictic use in (34)a. In the c-example *nu* does not express a semantically meaningful relation, but only refers to the textual level itself. In this sentence, *nu* functions as a discourse marker: it marks the shift to a new topic. Here, the coherence only exists in the mind of the speaker, which makes the discourse marker use of *nu* the most subjective. Now, subjectification of *nu* can be established empirically if it can be shown that the more subjective uses appear later

and develop out of the less subjective ones (cf. the procedure in Traugott & König 1991).[19] The change from (34)a to (34)b is an example of the rise of connectives; the change from (34)b to (34)c illustrates the shift away from the connective usage.

To conclude, the subjectification hypothesis seems to hold for the three types of conceptual change connectives might go through: the actual rise of connectives, changes from one connective function into another, and shifts away from the connective usage. However, the exact range of the subjectification hypothesis in relation to the diachronic development of these grammatical markers of clause combining is not yet clear. Does it hold for all types of connectives? And what about the relative subjectivity of different uses within a certain connective function? For example, causal connectives can be used either to mark cause-consequence relations occurring in reality, or to explicate argument-conclusion relations or even to mark relations between justifications and speech acts (cf. the domains of use in Sweetser 1990). Does the subjectification hypothesis even hold for subtle changes within the use as causal connective? And finally, how does subjectification relate to changes in categorical status (e.g. adverb vs. complementizer)? These last questions will be taken up as the starting point for the study in Chapter 8.

3.7 Conclusion and overview

The current chapter has shown that there is compelling evidence for the general idea that there is an interaction between form and function of connectives and the clauses they are contained in. On the basis of evidence from a variety of languages relatively specific form-function hypotheses can be formulated for each of the syntactic primitives.

The first syntactic primitive, the positioning of words used as connective, can single out the connective function from other (in particular: discourse marker) text-linguistic functions these words may have. The second primitive, linearization, can be related to the organization of discourse: initial adverbial clauses serve a grounding function, guiding the interpretation of the upcoming clause(s), whereas final adverbial clauses serve a more local semantic function. The third primitive, word order within the postposed matrix clause, can be related to the domains of use. The fourth syntactic primitive, word order within the connective clause itself has been related to two different text-linguistic properties of adverbial clauses: (a) their level of attachment to the preceding discourse, and (b) their interpretation based on domains or source of coherence. Verstraete (1998) and Haegeman (2003) have shown that a further link can be established between source of coherence and syntactic properties of the adverbial clause.

Evidence for the different interaction hypotheses often comes from languages other than Dutch. This leaves open the question whether these interactions also apply to the Dutch connectives that are selected in the current study. In the remainder of this thesis, the Dutch form-function relations are investigated from two developmental perspectives: a diachronic perspective (see Part II, Chapters 4-8) and an acquisition perspective (see Part III, Chapters 9-13).

Part II and III differ in their architecture. The acquisition part (Part III) follows a bottom-up approach. Given the scarce availability of detailed analyses of connective acquisition by young Dutch children (see Chapter 9), I will first pay much attention to my findings concerning this acquisition process (see Chapters 10-13). It is not until the final chapter of

[19] The development of *nu* from deictic element to connective to discourse marker is in line with the cline Traugott (1995a: 1) proposes: clause-internal adverbial > sentence adverbial > discourse particle (of which discourse markers are a subtype).

Part III (Chapter 13) that I will explicitly re-address the main question of this thesis. The diachronic part follows a top-down approach: each diachronic chapter puts a specific form-function relation to the test, with a subordinate role for the description of diachronic connective changes. Chapter 5 and 6 present a diachronic analysis of the complementizers *want* and *omdat*. Chapter 5 focuses on the interaction between word order and closure, whereas Chapter 6 investigates the interaction between word order and domains of use. The hypothesis about the interaction between text-linguistic function and positioning of connectives will be put to the test in Chapter 7, which discusses the diachronic development of the Dutch adverbials *dus* and *daarom*. To conclude the diachronic part of this thesis, Chapter 8 relates the syntactic development of the four causals *want*, *omdat*, *dus*, and *daarom* to an analysis based on subjectivity.

One important parallel between Part II and III is that both parts start with a methodological chapter (see Chapters 4 and 9). The next chapter, then, discusses several methodological considerations regarding the diachronic corpus research in this thesis.

Part II – Developments in the history of connectives

Chapter 4

On the methodology of the diachronic corpus studies

This chapter introduces the general methodology of the diachronic studies in this thesis. It begins with a description of the text material on which my study is based (4.1) and then goes on to discuss the way the connective samples are constructed (4.2). It also introduces the operationalization used for the conceptual analyses based on domains (section 4.3) and concludes (see 4.4) with a description of the statistical method used to test the hypotheses in the diachronic studies in the following chapters.

> *"(…) most of these works* [on text-linguistic properties, JEV] *are based almost exclusively on intuitive observations or present a minimum of empirical evidence but as Givón (1987) proposes, we need to devise operationalized discourse measurements in order to test the empirical validity of our hypotheses"*
>
> (Ramsay 1987: 384)

4.1 Sample of texts

The analyses of the historical developments of *want*, *omdat*, *dus*, and *daarom* in Chapters 5 to 8 are based on a corpus study, because a "corpus compiled from existing texts offers a more systematic and potentially more successful way of gathering data than dictionary searches" (Vismans 1994: 76). In order to perform this corpus study, I have assembled a sample of texts with the following characteristics (see (1)).

(1) <u>Characteristics of the sample of diachronic texts:</u>
   a. from the 13th, 16th, and 20th century
   b. from computerized sources
   c. from various Dutch regions
   d. from both literary and/or rhyming genres and non-literary and/or non-rhyming genres.

The characteristics in (1) are determined by several factors. (1)a is related to the choice of the research method; (1)b results from considerations of a more practical nature; (1)c and d are related to two specific criteria that are often applied to historical syntactic research.

In this diachronic study I perform an investigation in real time according to the trend model (see Gerritsen (1987: 5-7) for a more elaborate discussion of this method). This means that I track the historical developments of the four connectives by first making a synchronic description of the connective use in different stages of Dutch and then providing a diachronic description of the development of these connectives (cf. Koelmans 1979: 5). According to the trend model, one can compare the beginning and ending point of a certain time span, but this model also allows a researcher to take periods in between into account. In my study, I chose the latter option, because an investigation of more than two periods allows a smaller chance of mistakenly interpreting differences as change than a study of only two periods does.

The choice for the real time method restricted the time span to 800 years: the period in which written material is available. The earliest period with written Dutch texts is the

thirteenth century. The connective use in this century is contrasted with the use in very recent texts from the last decade of the twentieth century, which represent present-day (written) Dutch. One period in between has been selected: the sixteenth century.

The second factor restricting the choice of texts is not of a theoretical, but of a practical nature. For this study only computerized data have been used (e.g. historical texts from the *CD-rom Middelnederlands* 'Middle Dutch' (1998) or from the Internet). The reason for this is that it would have taken too much time to read through all the available manuscripts looking for fragments containing one of the four connectives at stake. The advantage of this criterion is that a survey based on texts that are readily available to other researchers can be checked easily. The disadvantage is that an uninteresting selection criterion has penetrated the research: why are certain texts computerized whereas others – from the same periods – are not? However, it may be argued that the effect of this criterion is comparable to the effect of the willingness of people to take part in a survey of social differences against the refusal of other people to do so (Gerritsen 1987: 18). In my study, the impact of this factor is less drastic. It is to be expected that the willingness to take part in a social survey interacts with social differences, whereas the fact that a text is computerized or not does not interact with conceptual or syntactic differences of the texts.

A third factor that could have restricted the choice of text material is a criterion often mentioned as specific to historical syntactic research, but which could equally apply to both the conceptual and the syntactic parts of this diachronic study. This is the criterion that a historical study should take regional variation into account, especially because there was no uniform, standardized Dutch language in the early centuries.[1] The language of these periods, so-called 'Middle Dutch'[2], consists of a group of dialects (Stoett 1977: 1).[3] In my study, regional variation is not taken into account, because I focus on similarities rather than differences between the medieval dialects. I am interested in all the possible uses of the words *want*, *omdat*, *dus*, and *daarom*, irrespective of dialectal source.[4] This choice is in line with Van der Wal's (1986: 140) main arguments against the necessity to take this criterion into account in order to gather a representative sample. She claims that most of the dialectal differences are phonological differences,[5] and that a parallel development between different dialects can be noticed for all kinds of syntactic phenomena (Van der Wal 1986: 232, footnote 104).

A fourth criterion, also specific to historical syntactic research, is that it should exclude rhyme texts because of the possible influence of stylistic factors such as rhyme and meter (see, among others, Lehmann 1972: 242; Werth 1970: 28). This is an important selection criterion, since many medieval texts are written in verse. As Blom (2002: 15) formulates it:

---

[1] From a purely linguistic point of view, the objection that Middle Dutch must be regarded as an idealization which abstracts from dialectal (and other kinds of) variation (Van der Wal 1986: 141) should be extended to all centuries. All three periodizations are relative constructs as language develops in a continuum (Vismans 1994: 81).

[2] The term 'Middle Dutch' is a construct of 19th-century researchers (Weerman 1988: 290).

[3] It should be noted that medieval Dutch texts are largely Flemish texts. Flanders was a wealthy region in the Middle Ages, which explains why so many texts were produced in that area (Blom 2002: 23).

[4] If it can be shown that certain uses only appear in part of the dialects, the worst consequence for my research is that my claims about early connective use should be restricted to those Dutch dialects (Weerman 1988: 291). As far as I know, however, the Dutch dialects do not differ in their use of the causal connectives at hand.

[5] Note that this outcome can at least partially be regarded as the result of a bias of researchers towards phonological characteristics of the Middle Dutch language.

"rhyming lines (…) should be treated with caution when studying word order phenomena", because "rhyme contains many unusual [= ungrammatical, JEV] word orders". Van der Wal (1986: 141-142) dismisses objections to using poetic texts as 'prejudice'. On meter, she claims that in medieval Dutch texts we "are dealing with free verse (…) which means that we do not have to be afraid of all kinds of unnatural (…) linguistic phenomena caused by forced adaptations to metrical form." In addition, Gerritsen (1987: 17) states: "exactly because of the rhyme (…) more of all the syntactic possibilities in a language are being used than in non-literary untranslated prose". Weerman (1988: 292) even argues that leaving out poetry because of stylistic factors is "not very principled", since stylistic factors always play a role. The current research follows his advice not to disregard any possible piece of evidence in advance, and to presume that all texts are built according to rules of the grammar, unless there are good grounds to presume the contrary.[6] In case of doubt about a certain word order, an additional analysis has been performed to investigate which word orders 'survive' if only non-rhyming texts are taken into account.

The characteristics in (1) determined the choice of the historical data that were gathered. These data were collected from a number of different sources, such as CD-roms (e.g. the CD-roms *Middelnederlands* 'Middle Dutch' and *Klassieke literatuur* 'Classical literature') and the Internet (e.g. the project *Laurens Jansz. Coster*).[7] For each period a corpus of at least 2000 pages text was arrived at (see Appendix A for an overview of the selected texts).

4.2 Samples of connective fragments

In accordance with the trend method, a randomly selected connective sample was constructed for each of the selected periods. In this study, one extra step was taken before actually constructing these samples. Since the total population of texts is rather heterogeneous, it is very possible that differences between, for example, the use of *want* in certificates of the 13th century and its use in modern Dutch youth novels do not reflect a real change, but just a difference in style. In this study I have tried to minimize the interference of genre-specific effects and other stylistic considerations, not by restricting my research to one genre or style (which would have led to the exclusion of various texts in Appendix A), but rather by incorporating one global genre difference into the data (cf. Gerritsen 1987: 24-29). To be more concrete, I divided the texts per period into two subcorpora: a corpus of 'rhyme texts' versus one of 'non-rhyming texts'.[8] For the thirteenth and sixteenth centuries this bipartition is indeed based on the distinction between rhyming and non-rhyming texts. During these periods this distinction coincided to a large extent with the one between literary and non-literary texts. In the twentieth century the genre 'rhyming literature' has almost vanished, which makes it impossible to keep the bipartition completely constant across ages. However, the parallel with literary rhyming versus mainly non-literary, non-rhyming texts can still be maintained by making a distinction between literary texts (novels) and non-literary texts (newspaper fragments) within the 20th-century corpus.[9] Therefore, the global bipartition is not

---

[6] See Fischer, Van Kemenade, Koopman & Van der Wurff (2002: 31-32) for other reasons not to exclude verse.

[7] The amount of work devoted to the collection of data from various digital sources shows that the DBNL-project is an excellent and necessary initiative. It also shows the need for tools that allow researchers to combine different texts into one file that can be searched easily.

[8] These labels do not fully cover all the text types in the two subcorpora (see my remarks on the 20th century), but they will be used for ease of reference anyway.

[9] The literary narrative corpus has kindly been given to me by Mirna Pit. It consists of the first 100 pages of 21 literary books (7 literary, 7 youth and 7 popular novels). The news paper fragments are

perfect, but a more refined subdivision is not possible given the restricted availability of computerized data from the early periods. An advantage of this bipartition is that it enables me to filter out obvious stylistic peculiarities and at the same time get a complete picture of the connective use in each period (by not excluding certain genres in advance).

Per period and per subcorpus all the available texts mentioned in Appendix A were put together into one file (resulting in six large text files). Per subcorpus I took a random sample of pages using a table with random digits.[10] For each connective in this study I selected the first 25 occurrences in all six samples. This procedure resulted in a corpus of 150 fragments per connective, that is 25 'rhyming' and 25 'non-rhyming' fragments for all three periods in this study (see Table 4.1).

Table 4.1. Number and nature of connective fragments selected for diachronic analysis

Period	**# rhyming and/or literary fragments**	**# non-rhyming and/or non-literary fragments**	**Total**
13th century	25	25	50
16th century	25	25	50
20th century	25	25	50
Total	75	75	150

As the spelling was not yet conventionalized in the early periods, different spellings of the same word can be found. Therefore, various search strings were used to select connective fragments (see Table 4.2). These strings are based on the spelling variants that are mentioned in the *Woordenboek der Nederlandsche Taal* (*WNT*) 'Dictionary of the Dutch language' (De Vries et al. 1882-1998) and the *Middelnederlands Woordenboek* (*MNW*) 'Middle Dutch dictionary' (Verwijs & Verdam 1885-1952).

Table 4.2. Search strings for *want*, *omdat*, *dus* and *daarom*

Connective	**Search strings**
want	*want*, *wante*, *waent*, and *went*
omdat	*omdat* and *dat* (per occurrence I determined whether *dat* was preceded by *om* or one of its variants, such as *ombe, um, omme*)
dus	*dus* and *dos*
daarom	*dar*, *daar*, *daer* and *der* (per occurrence I determined whether it was directly followed by *om* or one of its variants)

One additional selection criterion was invoked in order to obtain productive connective fragments. A connective sample containing many variants of the same formulaic expression (with no variation in conceptual or syntactic configuration), would give a limited picture of the use of this connective in the period under investigation. An example is the following *omdat*-fragment. The original non-rhyming sample of the 13th-century contained 19 variants

from the *Meppeler Courant* (MC), available in the *38 Miljoen Woorden Corpus 1996* '38 million words corpus 1996' from the *Instituut voor Nederlandse Lexicologie (INL)* 'Institute of Dutch Lexicology'.

[10] This procedure was not possible for the INL-corpus, because its electronic interface does not give access to the full texts (it only shows the results of queries on specific (strings of) words). Within the news-paper corpus I therefore selected every tenth connective from the odd months in the 1995-edition of the *Meppeler Courant*.

of this concluding administrative formula (which De Rooij 1982: 340 labels a 'corroboration formula'). The first of these fragments has been selected in my sample, whereas the other 18 fragments have been replaced by more productive ones.

(2) *Ende* ***omme dat*** *wie willen dat dese voreworde ende dese dinghe si vast ende ghestade so hebben wi dese lettren ende desen charter gheseghelt met onsen zeghelen wdhanghende.*
(CG 1-0217, 1278)[11]
'And because we want this agreement and this condition to become valid and have a binding effect, so we have sealed these letters and this charter with our seals hanging out.'

This extra selection criterion was also applied during the selection of *want*- and *daarom*-fragments from the non-rhyming 13th-century sample. Both connectives repeatedly appear in variants of the fixed expression *de wijze man zegt* 'the wise man says'. The fragments in (3) and (4) give examples of this formula from the *Nederrijns Moraalboek* 'Morality book from the Lower Rhine'. There is neither variation in the syntactic configuration *want* and *daarom* mark here, nor in the conceptual relation they express. Therefore, per connective only one of these cases has been selected in my sample. For *want*, two occurrences have been replaced by other more productive fragments. For *daarom*, seven cases have been replaced.

(3) ***Want*** *die wise man spriekt aldus. duo girigheit van di. inde angst sal die laten*
(NM, 1270-1290)
'For the wise man speaks thus: put greed away from you and fear will leave you'

(4) *Vort sal man huoden dat man nit alte lecker en si vop guode spise.*
dar vombe *segt die wise man. manlik die huode inde vorsiie sine magt inde sine mate.*
(NM, 1270-1290)
'Furthermore one has to take care that one is not too fond of good food.
That's why the wise man says: like an adult (is) the man who keeps his self-control and observes his moderation.'

4.3 Operationalization for the analyses based on domains

Each of the connective fragments was subject to a conceptual analysis based on domains (see Chapter 6 for the results for *want* and *omdat*, and Chapter 7 for *dus* and *daarom*). Since this analysis has been applied to all four connectives, its operationalization is presented here; other conceptual and syntactic methodologies are discussed in the relevant chapters.

In order to perform the conceptual analysis based on domains, I selected a large context per fragment and made a translation, if necessary. For each connective, I determined whether it was used as a causal connective or not by determining if an implication relation ($p \rightarrow q$) between the two clauses could be derived (cf. Sanders et al. 1992).[12] If the connective was indeed used to mark a causal relation, I also applied an analysis based on domains of use, in which I distinguished between content, epistemic, and speech act relations, as illustrated by the examples in (5)-(7) below (see also Chapter 2 for a more elaborate discussion of these domains).

---

[11] For three source texts, I use abbrevations: *CG = Corpus Gysseling*, *MC = Meppeler Courant*, and *NM = Nederrijns Moraalboek*.

[12] If both a causal and a temporal relation could be derived, the relation was still classified as causal on the basis of cumulative complexity (cf. Bloom, Lahey, Hood, Lifter & Fiess 1980, and the more elaborate discussion of this notion in Chapter 10).

(5) Content
Omdat *de pony de fietser bleef nalopen*
bracht deze hem maar onder bij de dieren in het park. (MC, 1995)
'Because the pony kept on following the cyclist he brought him to the animals in the park.'

(6) Epistemic
De sloffen maakten een wrang geluid, maar ze waren pure noodzaak, ***want*** *alleen dankzij die sloffen had ik het gevoel enige grip te hebben op de vloer.* (De vriendschap, 1995)
'The slippers made a dry sound, but they were pure necessity, because only thanks to those slippers did I have the feeling of having some grip on the floor.'

(7) Speech act
Kun je knollen rapen, ***want*** *ander werk is er niet.* (MC, 1995)
'Can you gather turnips, because there is no other work.'

To establish the domain type in an objective way, I used an adapted version of the 'Basic operation paraphrase test' (Sanders 1997: 126). This test works with paraphrases that make the different causal relations maximally explicit, independent of the connective used in the fragment. The paraphrase test consists of the following steps, exemplified by the content fragment mentioned earlier in (5).

(8) Paraphrase test
a. Determine the text segments for which the causal relation holds.
Omdat [$_{S1}$ *de pony de fietser bleef nalopen*]
[$_{S2}$ *bracht deze hem maar onder bij de dieren in het park*]
b. Reconstruct the propositions if necessary (explicate referential expressions, integrate implicit information that can be recovered from the conceptual and/or linguistic context).
Omdat [$_{S1}$ *de pony de fietser bleef nalopen*]
[$_{S2}$ *bracht de fietser de pony maar onder bij de dieren in het park*]
c. Remove the connective.
[$_{S1}$ *De pony bleef de fietser nalopen*]
[$_{S2}$ *De fietser bracht de pony maar onder bij de dieren in het park*]
d. Insert the paraphrases and determine which one fits best.
Situation P: *De pony bleef de fietser nalopen*
leads to the act/*[13]conclusion/*speech act:
De fietser bracht de pony maar onder bij de dieren in het park

An overview of the paraphrases is given in Table 4.3.[14] The a-versions of the paraphrases are 'forward-causal' (antecedent P and consequent Q are presented in the order in which they occur in reality: P precedes Q) and are applied to cases in which the *omdat*-clause precedes the main clause (e.g. (8)). The b-versions are 'backward-causal' (order of presentation is the

[13] An asterisk marks the unacceptibility of this paraphrase.
[14] Pander Maat & Sanders (1995: 352-353) argue that the occurrence of a connective in a certain domain is not a necessary condition for determining the text-linguistic meaning of the connective itself (or what it contributes to the total meaning of the fragment). According to them, an additional substitutability test is needed (cf. Knott & Dale 1994; Knott & Sanders 1998). In order to perform such a substitutability test, however, intuitions from native speakers are needed, which obviously are not available for the diachronic data at hand. In this study, then, only the occurrence in certain domains will be used to characterize the connectives.

reverse of the order in reality: consequent Q precedes antecedent P) and are applied to all *want*-clauses and to *omdat*-clauses that follow their main clause. Both non-causal use and non-connective use have been categorized as 'other use'.

Table 4.3. Overview of the paraphrases used in the domains analysis

Domain	**Paraphrase**
1. Speech act	1a. Situation P causes speech act Q
	b. Speech act Q is caused by situation P
2. Epistemic	2a. Situation P causes conclusion Q
	b. Conclusion Q is caused by situation P
3. Content	3a. Situation P causes act / situation Q
	b. Act / situation Q is caused by situation P
4. Other use	None of the paraphrases applies

The historical corpus contained fragments that proved to be more or less ambiguous in the categories mentioned above. These ambiguities are important for the diachronic development of linguistic elements; in fact, they can make evolution possible: a new meaning develops if the old one leaves room for an alternative interpretation, which then becomes conventionalized (cf. Traugott & König 1991). Although I acknowledge the possibility of ambiguity in certain fragments, for practical reasons I have chosen to set extra decision criteria to disambiguate such fragments. A first ambiguity occurred in cases that contain an explicated speech act, like the promise in example (9).

(9) *Heer vader,* ***om dat*** *ghi hem versonden hebt, so belove ic u, dat ic in kerstenrijck te Eggermont reisen sal ende claghent zijn vrienden* (Historie van Malegijs, 1556)
'Lord father, because you have banished him, I promise you, that I will travel to the church of Egmond and make a complaint to his friends'

In this fragment, the speaker chooses to mention his speech act explicitly. It is this speech act 'I promise you', he motivates in the preceding clause. Because of the great similarity with content relations in which the speaker motivates physical acts, I decided to categorize such cases as *content* rather than as *speech act*. This is in line with Martin's claim that the introduction of a projecting clause actually makes the originally internal relation external, "reconstructing the rhetorical structure of the text in experiential terms" (Martin 1992: 228). Or, as Verstraete (1998: 195-196) states it: "the meanings that were originally encoded subjectively in the grounding resources of the main clause are now 'objectified' as part of the instantiated type. (…) When the objectified ground elements are brought into the instantiated type, they lose their grounding function and become part of the description of the 'world' (…)". My category *speech act*, then, only includes fragments with implicit speech acts like example (10), in which an advice is given.

(10) *o etsijtes lief dat en suldie niet doen.* ***want*** *dat waer grote sonde dat ghi uwer moeder mesdeet* (Historie van Margrieta v. Lymborch, 1516)
'o Etsijtes sweetheart, you shouldn't do that, because it would be a great sin if you harmed your mother'

A different kind of ambiguity is the one between the categories *speech act* and *epistemic*. This concerns fragments like the following:

(11) *Vonuersiin doit kuompt gerne den gienen die lange wenen leuen. Jnde in dien hope vome laten te beteren. Jnde* ***dar vombe*** *sal man alle die dage die kuomen setten gelik den lesten dage* (NM, 1270-1290)
'An unexpected death often happens to those who expect to live long, and in that hope refrain from mending their ways. And that's why one should treat each day to come as ones final day.'

In this example, the *daarom*-clause can be interpreted in two ways: as a piece of advice (resulting in a speech-act relation) or as a claim that a certain conduct is recommendable (which results in an epistemic relation). In my analysis I have set an extra criterion to disambiguate such cases: the relation is labeled *speech act* if the context explicitly mentions the person who is addressed by the speaker, as is the case in (10): "Etsijtes sweetheart". If there is no specific person who is addressed, the relation is classified as *epistemic* as in (11), in which *man*, 'one', cannot be regarded as a specific person. The rationale behind this choice is that there is no use in performing a speech act like giving an order or a piece of advice if the speech act does not have a specific addressee who may act out this order or piece of advice.

4.4 Qualitative and quantitative analyses

In order to test my hypotheses, I make use of both qualitative and quantitative methods. As Gerritsen (1987: 12) points out, these methods are complementary to each other. In the qualitative method each fragment is analyzed separately and valued on its own merits. This is in line with Weerman (1988: 292), who argues that frequency does not play a role in determining the grammaticality of a certain construction; there is no reason why a grammatical clause should occur frequently. The qualitative method can therefore be used to determine which interpretations or syntactic properties are grammatical during a certain time span and whether they occur frequently or not.

Regardless, a quantitative method still remains useful. As Givón (1987) proposes, we need to devise operationalized discourse measurements in order to test the empirical validity of our hypotheses (cf. also Ramsay 1987: 384). Such a quantitative approach contrasts with many works on text-linguistic properties that are based almost exclusively on intuitive observations or present a minimum of empirical evidence. My quantitative analysis is comprised of charting the distribution of different uses and analyzing these distributions statistically. The quantitative method can be applied to investigate whether significant shifts in the distribution patterns occur, thus preventing the researcher from jumping to conclusions.

In order to obtain reliable statistical results it is important to work with a representative sample across the periods. However, the samples of the three periods under investigation differ in an important way, since different numbers of words were needed to gather 50 connective fragments per period. This is why the quantitative data are statistically tested with logit analyses, which take this kind of variation into account (see Appendix B for some explanatory remarks). Alpha decision level is set at .05.

The results for the diachronic development of *want* and *omdat* are presented in Chapter 5 (closure) and 6 (domains of use) and the results for *dus* and *daarom* in Chapter 7. Chapter 8 focuses on the diachronic development of all four connectives in terms of subjectification.

APPENDICES TO CHAPTER 4

Appendix A – Primary sources for the diachronic corpus study

Overview of the content of Appendix A:

A – 1 *List of abbreviations* – a list of abbreviations that are used to refer to the digital sources mentioned in the description of individual texts.

A – 2 *Primary sources 13th century* – a list of all the non-rhyming and rhyming texts that were used to construct a random sample of pages from the 13th century.

A – 3 *Primary sources 16th century* – idem for the 16th century.

A – 4 *Primary sources 20th century* – idem for the 20th century.

A – 1. List of abbreviations

	Meaning	Digital source
cg1	Corpus Gysseling I – Ambtelijke Bescheiden 'Corpus Judicial Texts'	CD-rom *Middelnederlands*
cg2	Corpus Gysseling II – Literaire Handschriften 'Corpus Literary Texts'	CD-rom *Middelnederlands*
rijm	Corpus Rijmteksten 'Corpus Rhyme Texts'	CD-rom *Middelnederlands*
pro	Corpus Proza 'Corpus Prose'	CD-rom *Middelnederlands*
abc		CD-rom *Klassieke literatuur*
cos	text available on the website of the *Project Laurens Jansz. Coster*	http://cf.hum.uva.nl/dsp/ljc/
dur	text available on the website *De Tachtigjarige Oorlog* 'The Dutch revolt', Leiden University	http://dutchrevolt.leidenuniv.nl/

References:

- CD-rom *Middelnederlands* 'Middle Dutch' (1998). Den Haag/Antwerpen: Sdu.
- CD-rom *Klassieke literatuur: Nederlandse letterkunde van de Middeleeuwen tot en met de Tachtigers* 'Classical literature: Dutch literature from the Middle Ages to the Eightiers' (1999). Utrecht: Het Spectrum.
- Oostendorp, M. van (ed.). *Project Laurens Jansz. Coster: Klassieke Nederlandstalige literatuur in elektronische edities* 'Project Laurens Jansz. Coster: Classical Dutch literature in electronic editions'. Available on: http://cf.hum.uva.nl/dsp/ljc/.

A – 2. Primary sources 13th century

The oldest texts were taken mainly from the standard corpus for this period, the *Corpus Gysseling* (available on the *CD-rom Middelnederlands* 'Middle Dutch'). This corpus consists of legal agreements (charters, deeds, and similar judicial documents) and some literary texts (e.g. hagiographies).

I – Non-rhyming texts 13th century[15]

Text	Source	Date	Author	# words
Ambtelijke Bescheiden (longer than one page)	cg1	1240-1300	various authors	284081
Nederrijns Moraalboek	cg2	1270-1290	nn	36265
Nederbergse geneeskundige recepten	cg2	mid 13th century	nn	481
Noordlimburgse gezondheidsregels	cg2	shortly after 1253	nn	674

II – Rhyming texts 13th century

Text	Source	Date	Author	# words
Der naturen bloeme (Detm. handschrift)	cg2	1275-1300	J. van Maerlant	366
Limburgse Aiol (Aiol 1)	cg2	1220-1240	nn	4404
Esopet	abc	1215	nn	9240
Alexiuslegende	cg2	1280-1300	nn	342
De boec van Catone (Enaamse codex)	cg2	1290-1300	M. van Torhout (?)	2417
De boec van seden (Enaamse codex)	cg2	1290-1300	M. van Torhout (?)	2983
Boec vander biechten (Enaamse codex)	cg2	1290-1300	M. van Torhout	722
Boeve van Hamtone	cg2	1260-1270	nn	676
Der naturen bloeme: Munchense fragm.	cg2	1275-1300	J. van Maerlant	4866
Leven van sente Kerstine	cg2	1275-1300	Broeder Geraert	13311
Sente Lutgard	cg2	1263-1280	W. van Affligem	125274
Sinte Lutgard	cg2	1275-1300	Broeder Geraert	25266
Minnedichten uit Ter Doest	cg2	1290-1300	nn	227
Nevelingenlied, Brabantse vertaling	cg2	1260-1280	nn	1248
Ongeïdentificeerd fragment 1	rijm	1200-1250	nn	657
Parthonopeus van Bloys (Berlin)	rijm	1290-1310	nn	906
Parthonopeus van Bloys (Köln)	rijm	1290-1310	nn	571
Parthonopeus van Bloys (Brussels)[16]	rijm	1350-1400	nn	50024
Reynaert 1 fragment E	cg2	1275-1300	Willem	1689
Reynaert 1 fragment G	cg2	1260-1280	Willem	435

[15] One other text from the 13th century (the *Plantenglossarium* 'Plant glossary') was not useful for a connective study and still another (*Episch fragment uit de IJselstreek* 'Epical fragment from the IJssel region') was not available on the CD-rom *Middelnederlands* 'Middle Dutch' despite its reference in the index.

[16] As the date indicates, this fragment of *Parthonopeus van Bloys* stems from the 14th century. It has still been selected for the 13th-century sample, because the text shows an overlap with the fragment from Berlin, which does stem from the 13th century.

II – Rhyming texts 13th century (continued)

Text	Source	Date	Author	# words
Rijmbijbel (Scholastica)	cg2	1275-1300	J. v. Maerlant	186547
Roman van Perchevael	cg2	1275-1300	nn	4418
Van sente Caterinen (Enaamse codex)	cg2	1290-1300	M. v.Torhout (?)	758
Van sente Eustace (Enaamse codex)	cg2	1290-1300	M. v. Torhout (?)	2260
Van sente Marien Egyptiake (En. codex)	cg2	1290-1300	M. v. Torhout (?)	4148
Sint Servaes legende (Enaamse codex)	cg2	1190-1210	H. v. Veldeke	1824
Van sente Waernere (Enaamse codex)	cg2	1290-1300	M. v. Torhout (?)	1475
Tristant	cg2	1240-1260	nn	768
Van den bere Wisselau	cg2	1280-1300	nn	3002
Van der sielen ende van den lichgame (Enaamse codex)	cg2	1290-1300	M. v. Torhout (?)	727
Van onser vrouwen gheslachte (En. codex)	cg2	1290-1300	M. v. Torhout (?)	369
Van onser vrouwen lof (Enaamse codex)	cg2	1290-1300	M. v. Torhout (?)	392
Wrake van Ragisel	cg2	1260-1280	nn	5228
Floris ende Blanchefloer	abc	1260	Assenede	24894
Vanden vos Reynaerde	abc	1260	Willem	18916
Gedichten	cos	13th century	Hadewijch	18835

A – 3. Primary sources 16th century

Most texts from the 16th century were taken from CD-roms: the *CD-rom Middelnederlands* 'Middle Dutch' and the *CD-rom Klassieke literatuur* 'Classical literature'. The remaining texts come from websites with historical texts.

I – Non-rhyming texts 16th century

Text	Source	Date	Author	# words
(Dyalogus tusschen coninck) Salomon ende Marcolphus	pro	1501	nn	8097
Euangelien vanden spinrocke	pro	1510-1530	nn	11926
Exempel van een soudaensdochter	pro	1500-1520	nn	2145
(Historie van den) vier heemskinderen	pro	1508	nn	75451
Historie van den wonderlicken Merlijn	pro	1530-1550	nn	4361
Historie van Hughe van Bordeus	pro	1530-1550	nn	33236
Historie van Jan van Beverley	pro	1543	nn	5803
Historie van Malegijs	pro	1556	nn	105126
Historie van Margarieta van Lymborch	pro	1516	nn	98402
Droefliken strijt van Roncevale (non-rhyming parts)	pro	1510-1530	nn	10208
Historie van de borchgravinne van Vergi (non-rhyming parts)	pro	1550-1570	nn	6904
Leven van Sinte Clara	pro	1500-1520	nn	24192
Ulenspieghel	abc	1519	nn	22893
Boeventucht	abc	1587	D.V. Coornhert	4814
Loterijspel	abc	1596	Jan van Hout	10866
Super universas	dr	1559	De Lange (translator)	3413
Het verbond der edelen	dur	1566	several authors	1399
Smeekschrift der edelen	dur	1566	several authors	1704
Willem van Oranje roept op tot verzet	dur	1572	W. van Nassau	787
Holland en Zeeland dragen de hoge… (Dordrecht)	dur	1575	P. Buys	2268
Pacificatie van Gent	dur	1576	several authors	2872
Satisfactie van Schoonhoven	dur	1577	W. van Nassau	874
Satisfactie van Haarlem	dur	1577	W. van Nassau et al.	1785
Satisfactie van Heusden	dur	1577	W. van Nassau et al.	930
Satisfactie van Amsterdam	dur	1578	several authors	3444
Religievrede te Antwerpen	dur	1578	W. van Nassau et al.	1221
Unie van Utrecht	dur	1579	Lamzweerde	3489
Filips II doet Willem van Oranje in de ban	dur	1580	Filips II	5925
De Staten van Friesland verbieden…	dur	1580	several authors	1028
Holland en Zeeland dragen de Hoge… (Den Haag)	dur	1581	C. de Rechtere	2247
Plakkaat van Verlatinghe	dur	1581	J. van Asseliers	4883
De Staten-Generaal te Den Haag richten…	dur	1591	C. Aerssen	623
Reductie van de stad Groningen	dur	1594	several authors	1319

II – Rhyming texts 16th century

Text	Source	Date	Author	# words
Elckerlijc	abc	1500	P. Dorlandus (?)	6070
Mariken van Nieumeghen	abc	1501-1515	nn	10480
Geuzenliedboek	cos	1581	nn	5815
Spel van sinnen	cos	1597	J. Prins	7977
Droefliken strijt van Roncevale (rhyming parts)	pro	1510-1530	nn	7824
Historie van de borchgravinne van Vergi (rhyming parts)	pro	1550-1570	nn	6004
Historie van Jan van Beverley	pro	1543	nn	5803
Antwerps liedboek (Schoon liedekens boeck)	rijm	1544	nn	65629
Spiegel der jongers, Der kinderen spieghel	rijm	1510-1520 (?)	L. Goetman	3199
Devoot ende profitelyck boecxken	rijm	1539	nn	83935
Eerste muziekboeksken van Tielman Susato	rijm	1551	nn	2047
Historie van Gaver Capeel	rijm	1500-1520	nn	1341
Jan Splinters testament	rijm	1508 (?)	nn	1352
Suverlijc boecxken	rijm	1508	nn	10120
Tweede muziekboeksken van Tielman Susato	rijm	1551	nn	1475
Camp van der doot	cos	1503	J. Pertcheval (translator)	18466

A – 4. Primary sources 20th century

For the twentieth century I used non-literary newspaper fragments and a corpus of more or less literary narrative texts.

I – Non-rhyming / non-literary texts 20th century

The newspaper fragments were taken from the *Meppeler Courant* (1995-edition) available in the *38 Miljoen Woorden Corpus 1996* '38 million words corpus 1996' from the *Instituut voor Nederlandse Lexicologie* 'Institute of Dutch Lexicology'. See Kruyt & Dutilh (1997) for more information on this corpus.

II – Literary / narrative texts 20th century

The narrative corpus has kindly been given to me by Mirna Pit. It consists of the first 100 pages of 21 literary books (7 literary, 7 youth and 7 popular novels; see also Pit 2003).

<u>Literary novels</u>

Dis, A. van (1994). *Indische duinen*. Amsterdam: Meulenhoff, 5-135.
Dorrestein, R. (1996). *Verborgen gebreken*. Amsterdam: Contact, 9-115.
Hart, M. 't (1993). *Het woelen der gehele wereld*. Amsterdam: Arbeiderspers, 11-105.
Hermans, W.F. (1992). *Au pair*. Amsterdam: De Bezige Bij, 5-101.
Palmen, C. (1995). *De vriendschap*. Amsterdam: Prometheus, 9-115.
Winter, L. de (1993). *Hoffman's honger*. Amsterdam: De Bezige Bij, 5-101.
Zwagerman, J. (1991). *Vals licht*. Amsterdam: Arbeiderspers, 7-100.

<u>Popular novels</u>

Gils, A. van (1993). *De laatste wens*. Helmond: Westfriesland, 7-99.
Saris, L. (1993). *De gouden handjes*. Helmond: Westfriesland, 7-87.
Schuitemaker-Commandeur, T.C. (1990). *Vervuld van verlangen*. Helmond: Westfriesland, 7-91.
Veen, S. van der (1993). *Het land aan de horizon*. Ede: Hardeman, 7-101.
Veenhof, J.G. (1994). *In de schaduw bloeien de rozen*. Ede: Hardeman, 7-101.
Wageningen, G. van (1992). *Je weet toch waarom*. Ede: Zomer & Keuning, 5-89.
Wages, M. (1996). *Mooie woorden*. Amsterdam: Zomer & Keuning, 5-88.

<u>Youth novels</u>

Akker, W. van den (1992). *Het raadsel van de ketting*. Amsterdam: Contact, 5-87.
Beckman, Th. (1994). *De doge-ring van Venetië*. Rotterdam: Lemniscaat, 7-99.
Groen, E. de (1992). *Jeans voor een matrjosjka*. Tilburg: Elzenga, 8-105.
Hartman, E. (1993). *De voorspelling*. Rotterdam: Lemniscaat, 5-101.
Marijn, J. (1991). *Het verbroken zegel*. Houten: Lannoo, Van Holkema & Warendorf, 7-97.
Rood, L. (1990). *Erin de Enige*. Amsterdam: Leopold, 7-95.
Terlouw, J. (1989). *De kunstrijder*. Rotterdam: Lemniscaat, 5-95.

Appendix B – Explanatory remarks on the logit analyses

In a logit analysis, successive statistical analyses are performed applying different models. It is possible to add or leave out different factors or parameters, which results in different models. However, I am interested in the best model, which is the model that:

a. is the most sparse (i.e. the model with the fewest factors);
b. fits the data (i.e. the χ^2 of the model has a p-value > .05);
c. best describes the data (i.e. the model that has the lowest χ^2 and that cannot be improved significantly by adding a factor).

The χ^2 is a testing statistic that can be used to evaluate the difference between the observed frequencies and the model-based expected frequencies. If χ^2 is high, there is a large discrepancy between both, and hence the model fits the data poorly. If χ^2 is low (in relation to to the degrees of freedom), the model fits the data and the parameters for the individual variables can be interpreted meaningfully. In comparing different models we can compare the difference in χ^2 in relation to the difference in degrees of freedom (see Fienberg 1977).

For example, Table I below presents a logit analysis of the diachronic results for *omdat* (these will be discussed in Chapter 5 and 6).

Table I – Results logit analysis *omdat*

Logit (fr)	χ^2 model	df	p model	χ^2 factor	df	p factor
1. constant	280.69	35	< .001	-	-	-
+ 2. period (per)	275.54	33	< .001	5.20	2	< .1
+ 3. linearization (lin)	226.36	32	< .001	49.18	1	< .001
+ 4. word order (wo)	161.06	31	< .001	65.31	1	< .001
+ 5. domain (dom)	45.40	29	< .1	115.70	2	< .001
+ 6. period x word order	38.61	27	< .1	6.79	2	< .05
+ 7. linearization x domain	32.61	25	< .25	6.01	2	< .05
+ 8. per x lin + per x dom + lin x wo + wo x dom + per x lin x wo + per x lin x dom + per x wo x dom + lin x wo x dom + per x lin x wo x dom	0	0	1	32.61	25	< .25

The first model in Table I only takes a constant into account, predicting that the distribution of *omdat* over the different cells will be equal. This model does not fit the data (p < .001). The second model takes into account the period in which the connectives occur. However, period does not appear to influence the distribution of *omdat* (χ^2 (5) = 5.2; p < .1). The third model takes into account the constant, the factor 'period', and the factor 'linearization'. This model is significantly better than both the first and the second model; the drop in χ^2 is, as compared to the two previous models, significant. From this model it can also be concluded that the factor 'linearization' plays a significant role (see the column 'χ^2 factor': the linearization of the *omdat*-clauses is needed to account for the overall distribution pattern of *omdat*). However, the third model still does not fit the data. Therefore, in the fourth model, the 'word order' is introduced as an explanatory variable. This significantly increases the fit of the model. Hence, word order influences the distribution of *omdat*. Nevertheless, the model with effects of linearization and word order still does not fit the data. Therefore, effects of 'domain', and the interaction effects of 'period x word order' and 'linearization x domain' are introduced in the model (model 5, 6, and 7 respectively). This results in a model that fits the data (χ^2 (25) = 32.6; p < .25). Introducing even more factors does not result in significant

improvements of the model. For example, model 8, a saturated model that takes into account all the independent factors and all their interactions, is not significantly better than model 7.

Logit analyses have several advantages over chi-square analyses. First of all, whereas chi-square analyses are limited to cross tables with two variables, logit analyses enable us to investigate more than two variables (and interactions between them) at a time. Thus, one logit analysis gives more information than separate chi-square analyses.

Another important advantage is that logit analyses can handle samples divergent in terms of their connective frequency. For example, in order to obtain 50 occurrences of *want* from the corpus of the 16th century, I used a smaller sample of words than for 50 occurrences from the 13th-century corpus. In other words, sample sizes associated with each period are not equal. As a result, fluctuations between possible samples are larger for periods with a higher connective frequency. In a logit analysis these sample differences are taken into account (logit (F) = log (F/(N-F); F = number of connectives; N = total number of words). This is one reason why a logit analysis has to be preferred over a loglinear one. Indeed, century, domain and linearization are seen as independent variables. In this respect, the analysis resembles a traditional analysis of variance, except that the dependent variable is dichotomous.

Furthermore, logit analyses give the opportunity to specify the precise nature of interaction effects between variables with more than two values. The different parameter estimates of the model can be evaluated by means of their standard errors (i.e. if | parameter estimate | $\geq$ 1.96 * s.e.; $p \leq 0.05$). For each factor, there are N-1 estimates, since the last one serves as a "reference" category. For example, Table II presents the parameter estimates that belong to model 7 of the logit analysis in Table I. Table II shows which parameters have significant z-scores; these are the scores $z < -1.965$ or $z > 1.965$ (significance level for two-sided tests). For example, all other things being equal, there is no difference between 13th and 16th century in the use of *omdat* (p = .76). However, compared to the 13th century, *omdat* occurs more often in the 20th century (p = .01).

Table II – Parameter estimates *omdat* for model 7

Parameter	Estimate	s.e.	z-score	p
constant	-8.35	0.18	-45.52	< .001
period: 16th century	0.07	0.23	0.31	0.76
period: 20th century	0.56	0.23	2.47	0.01
linearization: preposed	-1.22	0.23	-5.37	< .001
word order: V-ambiguous	-0.84	0.31	-2.74	0.01
domain: epistemic	-1.07	0.21	-4.43	< .001
domain: speech act	-3.22	0.55	-5.82	< .001
period x word order: 16th V-ambiguous	-0.95	0.51	-1.87	0.06
period x word order: 20th V-ambiguous	-1.26	0.55	-2.31	0.02
linearization x domain: preposed epistemic	-0.69	0.56	-1.23	0.22
linearization x domain: preposed speech act	1.46	0.76	1.92	0.06

In each of the diachronic chapters to follow, one or more logit analyses of the frequencies per connective will be presented. For each connective, the comparison of different models is presented in an appendix, in a table like Table I. In the main text of the relevant chapter, I will only discuss the significant parameter estimates and the theoretical implications of the model that fits best. This is a logical thing to do, since the most adequate model indicates best which factors should or should not be considered to account for the diachronic connective data.

Chapter 5

Word order and closure: a diachronic analysis of want *and* omdat

Chapter 3 (see section 3.5) introduced several theories concerning word orders within connective clauses and their text-linguistic counterparts. The current chapter takes up one of these form-function relations: it investigates the interaction between the text-linguistic characteristics of *want*- and *omdat*-clauses based on closure and their respective word order patterns. To this end the diachronic developments of these complementizers are studied.

> *"Het voorstel dat ik wil doen komt erop neer dat het syntactische verschil niets te maken heeft met de* inhoud *van de relatie tussen tekstsegmenten, maar wel met de structurele organisatie van een tekst."* (Verhagen 2001: 111)
> 'The proposal I want to make boils down to this: the syntactic difference [between V2 *want* and V-late *omdat*, JEV] has nothing to do with the content of the relation between text segments, but rather with the structural organization of a text.'

5.1 Introduction

This chapter investigates the interaction between the text-linguistic characteristics of *want*- and *omdat*-clauses based on closure and their respective word order patterns. To this end the diachronic developments of these complementizers are studied. Two questions will be answered for both *want* and *omdat*.

(1) <u>Research questions of this chapter:</u>
   a. Did any syntactic changes occur during the selected time span?
   b. Can the syntactic changes be related to the conceptual analysis based on closure?

My diachronic investigation of the syntactic properties of *want* and *omdat* is not the first study in this field. Syntactic characterizations of both connectives can be found in the comprehensive *Middelnederlands Woordenboek* (*MNW*) 'Middle Dutch dictionary' (Verwijs & Verdam 1885-1952) and *Het Woordenboek der Nederlandsche Taal* (*WNT*) 'The dictionary of the Dutch language' (De Vries et al. 1882-1998). The syntactic characteristics of *want* have also been discussed in greater detail by Burridge (1993) and Van Megen (2002). Why then a new investigation of the diachronic development of *want*? In my opinion Burridge and Van Megen shed insufficient light on the conceptual side of the diachronic development, although they both do make remarks on certain conceptual properties of *want*-clauses. For example, they only distinguish between 'weak' versus 'strong' causal use, without giving clear definitions of these terms (see Burridge 1993: 57; Van Megen 2002: 17). The novelty of this study, then, lies in relating the largely known syntactic properties to a more elaborate conceptual characterization of the *want*- and *omdat*-clauses based on closure (Chapter 5) and domains of use (Chapter 6).

Chapter 5 is organized as follows. Section 5.2 introduces the necessary theoretical background on word order and on closure. In this section I also formulate hypotheses for the current research. In section 5.3 the methodology of this study is accounted for, after which the results from the corpus study are presented (see 5.4 on *want* and 5.5 on *omdat*). Section 5.6 presents the conclusions as well as some discussion. The results on the diachronic developments of *want* and *omdat* show that Verhagen's theory about the interaction between word order and closure should be restricted to postposed adverbial clauses and that additional research on Modern Dutch data is needed to draw firm conclusions.

5.2 Word orders related to closure

In the current section I present De Haan's (2001) theory about the relation between word order of the connective clause on the one hand and its internal and external syntax on the other (see 5.2.1). De Haan's ideas complement Verhagen's (2001) theory about the interaction between word order and level of attachment (5.2.2). The combination of the two theories will be used to formulate hypotheses for the diachronic research on *want* and *omdat* (5.2.3).

5.2.1 De Haan (2001) on word orders

In this section I introduce De Haan's (2001) theory and show how it can be used to account for the syntactic distinction between modern Dutch *want* and *omdat*. De Haan tries to give an account of Frisian 'subordinate' clauses, which normally exhibit V-late, but sometimes show V2 despite the presence of a complementizer.[1] In the discussion below, I will focus on his discussion of Frisian clauses that are introduced by the complementizer *omdat* 'because'. In its V2-use (see (2)) Frisian *omdat* is equivalent to both Frisian *want* (cf. De Haan 2001: 29) and modern Dutch *want*. In its V-late use (see (3)) it parallels modern Dutch *omdat*.

(2) *Hy koe net komme* ***omdat*** *hy* <u>*moest*</u> *Teake helpe.* (De Haan 2001: 8)
He could not come because he must Teake help.

(3) *Hy koe net komme* ***omdat*** *hy Teake helpe* <u>*moest*</u>*.* (De Haan 2001: 8)
He could not come because he Teake help must.
'He could not come because he had to help Teake.'

De Haan argues that the clauses showing V2 should be regarded as structural root clauses (root CPs). Root CPs differ in two respects from their V-late counterparts, which are labeled non-root CPs. First of all, there is a difference in their internal syntax. "Root clauses are considered to be independent expressions, that is, they belong to sentence types (…) with a particular illocutionary force. Non-root clauses do not have an illocution of their own: they are part of the root and belong to the illocutionary type of the root" (De Haan 2001: 25). This difference can also be applied to the modern Dutch connectives *want* and *omdat*. The illocutionary force of *want*-clauses can be altered independently of the illocutionary force of the matrix clause (see the a-variants of (4)-(6)). As the b-variants show, *omdat* does not have this possibility; this connective can only occur in clauses of the (default) declarative type, unless the whole *omdat*-clause is part of a different speech act (e.g. interrogative, imperative) expressed by the matrix clause.[2, 3]

---

[1] See De Haan (2001: 3-5) and the references there for some remarks on the complementary distribution of fronted verbs and the presence of complementizers.

[2] See Verstraete (2000: 9) on the suitability of this test of changing the illocutionary force of either the second or the first segment.

(4) a. *Nathan bood Mirjam een snoepje aan,* ***want*** *ze was verdrietig.*
b. *Nathan bood Mirjam een snoepje aan,* ***omdat*** *ze verdrietig was.*
'Nathan offered Mirjam a candy, because she was sad.'
(5) a. *Nathan bood Mirjam een snoepje aan,* ***want*** *was ze niet verdrietig?*
b. **Nathan bood Mirjam een snoepje aan,* ***omdat*** *was ze niet verdrietig?*
'Nathan offered Mirjam a candy, because wasn't she sad?'
(6) a. *Bied Mirjam een snoepje aan,* ***want*** *ze is verdrietig!*
b. *#Bied Mirjam een snoepje aan,* ***omdat*** *ze verdrietig is!*
'Go and offer Mirjam a candy, because she is sad!'

The second difference between root and non-root CPs concerns their external syntax. Unlike their V-late counterparts root CPs are unintegrated with respect to their matrix clause, i.e. they do not occupy a structural position within that clause. As De Haan (2001: 22) formulates it, the V2 variants "differ externally from their subordinated counterparts in that the former occur outside the clause in which they function syntactically". There are several properties of the Frisian V2 *omdat*-clauses confirming that they are unintegrated with respect to their matrix clause (cf. De Haan 2001: 16-19). One of these concerns the scope of negation. Negative sentences with a causal adverbial clause allow for ambiguity depending on whether the causal clause is inside or outside the scope of negation. Such sentences are expected not to display ambiguity when the causal clause exhibits V2, because in that case it is an unintegrated root CP and hence outside the scope of a negative phrase in the matrix clause (cf. De Haan 2001: 18). This difference also shows up in the two uses of Frisian *omdat*. The V-late example in (7) is ambiguous. In a first interpretation, the causal clause is outside the scope of negation, meaning 'the reason for his not-coming is the bad weather'. In a second interpretation, the causal clause is inside the scope of negation, meaning 'the reason for his coming is not that the weather was bad'. The V2-example (8) can only receive the first interpretation.

(7) *Hy komt net [***omdat*** it min waar <u>wie</u>].* (De Haan 2001: 18)
He comes not because it bad weather was.
(8) *Hy komt net [***omdat*** it <u>wie</u> min waar].* (De Haan 2001: 18)
He comes not because it was bad weather.
'He does not come because the weather was bad.'

The same scopal difference can be found in the use of the modern Dutch equivalents of Frisian *omdat* (see (9) and (10)). A clause combination containing the modern Dutch subordinator *omdat* allows for both interpretations, whereas a clause combination containing *want* can only receive the first interpretation (cf. De Vries 1971: 417-418; Van Belle 1989: 437-439; Degand 1996: 163-164; Pit 2003: 18).

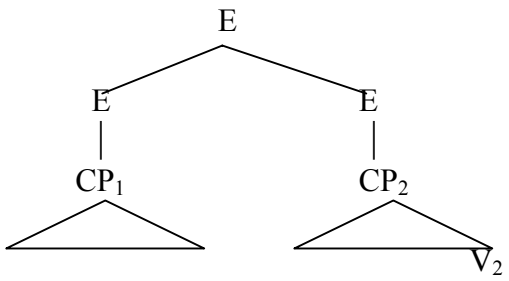
(9) *Hij komt niet [***omdat*** het slecht weer <u>was</u>].*
He comes not because it bad weather was.
(10) *Hij komt niet [***want*** het <u>was</u> slecht weer].*
He comes not because it was bad weather.

---

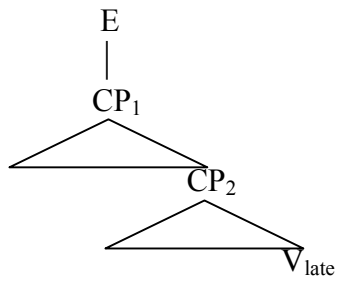
[3] The utterance in (6)b is grammatical if the *omdat*-clause is regarded as a part of the imperative. In that case, the *omdat*-clause provides a reason for giving the candy and not a justification of the speech act.

Schematically, the difference in external syntax can be presented as in (11), a slightly adapted version of De Haan's (2001: 22) representation. The a-variant shows the structure of a clause combination with a root CP, whereas (11)b shows the structure of a clause combination with a non-root clause. In both cases CP_1 represents the matrix clause and CP_2 the connective clause. Following Banfield (1973), De Haan uses the presence versus absence of an E-node above the CP to reflect the difference in internal structure between root and non-root clauses: the V-late CP_2 in (11)b lacks such an E-node, indicating that this clause does not have an illocution of its own. The V2 CP_2 in (11)a, on the other hand, does have such an E-node of its own.

(11) a.



b.



In terms of De Haan (2001), then, the V2-complementizer *want* can be said to connect two root CPs, both with their own illocutionary force. The subordinator *omdat*, on the other hand, connects a root CP (with illocutionary force) and a non-root CP (without illocutionary force of its own). In his view, a modern Dutch *omdat*-clause should be regarded as a CP subordinated to the matrix verb.

5.2.2 Word order related to text structure

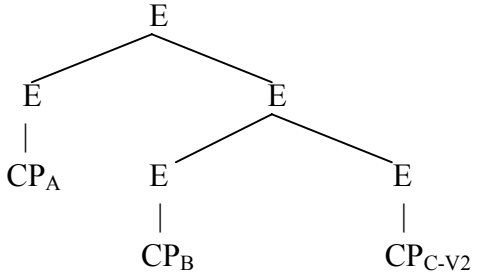
De Haan's (2001) ideas can be used to account for Verhagen's (2001) theory about the interaction between word order and level of attachment of the adverbial clause. As mentioned in Chapter 3 (see section 3.5.2), Verhagen's (2001: 111) proposal is that the syntactic difference between *want* and *omdat* has nothing to do with the content of the relation between two text segments (i.e. the specific domain they are used in), but that it relates to the structural organization of the text. He argues that in combinations of three clauses [A-B-C], a subordinating *omdat*-clause C, as in (12), can only be related to clause B. This results in late closure, represented as [A-[B-C]]. If the same information is connected by the coordinator *want*, as in (13), the C-clause can be interpreted as relating either to B, or to the clause combination A-B.

(12) [**A** *Getuigeverklaring X mag niet in de beschouwingen betrokken worden.*]
'Testimony X may not be taken into consideration.'
[**B** *Verder moeten de stukken Y en Z uit het dossier verwijderd worden,*]
'Furthermore, documents Y and Z have to be removed from the dossier,'
[**C** ***omdat*** *vroegere veroordelingen in de lopende zaak geen rol* <u>*mogen*</u> *spelen.*]
because earlier convictions in the current case no role may play
'because earlier convictions may not play a role in the current case.'

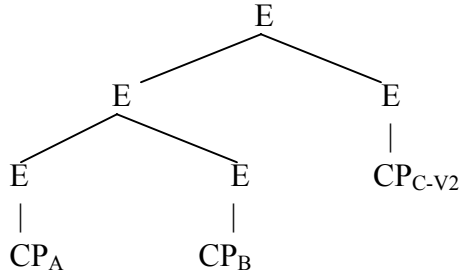
(13) [**A** *Getuigeverklaring X mag niet in de beschouwingen betrokken worden.*]
[**B** *Verder moeten de stukken Y en Z uit het dossier verwijderd worden.*]
[**C** ***Want*** *vroegere veroordelingen* <u>*mogen*</u> *in de lopende zaak geen rol spelen.*]
Because earlier convictions may in the current case no role play.
'Because earlier convictions may not play a role in the current case.'

The difference in behavior of V2 and V-late clauses based on closure can be related to De Haan's syntactic analysis based on root and non-root clauses. V2 clauses have an E-node of their own. This allows them to either be attached to the previous clause, resulting in late-closure (as represented in (14)a), or to a combination of two previous clauses (see (14)b).

(14) a.

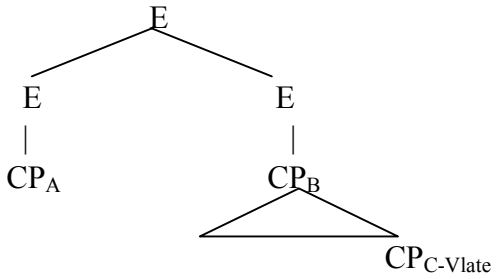


b.



V-late clauses, on the other hand, do not have this latter possibility; in fact, they even do not have the structural properties represented in (14)a. Lacking an E-node themselves, these non-root clauses cannot be attached to another clause independent of their matrix clause (see (15)).

(15)



So, De Haan's theory can be used to account for Verhagen's claim about the interaction between word order and attachment level of a clause. The latter claim has been formulated within a synchronic analysis of *want* and *omdat*. Verhagen's line of reasoning is convincing, but he only supports it with exemplary examples. The diachronic corpus study in this chapter allows me to put his claims to the test.

5.2.3 Hypotheses for the diachronic study

In the current section Verhagen's proposal about the interaction between word order and closure will be used to formulate hypotheses about the diachronic development of *want* and *omdat*. These hypotheses build on observations in the *MNW*, the *WNT* and the work of Burridge (1993) and Van Megen (2002) that – in contrast to Modern Dutch coordinating *want* – Medieval *want* could be combined both with V2 and with V-late.

Verhagen's claim about the interaction between word order and attachment level of a clause has been formulated within a synchronic analysis of *want* and *omdat*. According to him a clause marked with V-late can only receive a late-closure reading, whereas V2 does not impose such a restriction on the attachment level of the adverbial clause. From his line of reasoning, two hypotheses can be inferred for the diachronic analysis of the two complementizers.

(16) <u>Hypotheses for the diachronic study:</u>

a. The Middle Dutch *want*-fragments showing V-late are all cases of late closure.
b. All of the *omdat*-fragments showing V-late are cases of late closure.

In other words, the V-late fragments should always show up in the hierarchical text structure [A-[B-C]] and not in the text structure [[A-B]-C]).

5.3 Methodology

This section introduces the methodology specific to the diachronic study of *want* and *omdat*, which is performed in order to shed light on the tenability of the hypotheses in section 5.2.3. Using the general methodology mentioned in the previous chapter, the following selection of connective fragments was arrived at, both for *want* and *omdat* (see Table 5.1).

Table 5.1. Number and nature of connective fragments selected for diachronic analysis

Period	**# of non-rhyming and/or non-literary fragments**	**# of rhyming and/or literary fragments**	**Total**
13th century	25	25	50
16th century	25	25	50
20th century	25	25	50
Total	75	75	150

All 150 fragments per connective have been subject to two syntactic analyses (5.3.1) and one conceptual analysis (5.3.2). The syntactic analysis includes examination of both the word order patterns within the connective clause and the linearization of the connective clauses as a whole. These analyses can be used to determine the categorical status of the connective (coordinator or subordinator). The conceptual analysis of each adverbial clause involves a characterization based on its textlinguistic level of attachment.

5.3.1 Syntactic analyses

The first syntactic analysis of the fragments containing *want* or *omdat* involves the word order patterns within the connective clauses. All clauses – except one without a finite verb – were classified into three types according to the positioning of the finite verb. The category of *V2* (also referred to as *coordinate use*) includes utterances in which verb second can clearly be established. In *V2*-fragments it can be shown that the finite verb does not occupy a position within the verb phrase (VP), but a position within a higher functional projection of the connective clause. The finite verb occupies a 'higher' position if it appears before the subject (in the case of inversion, as in (17)) or if some other linguistic element (e.g. a sentential adverb, a negation adverb, a direct or indirect object) occurs between the finite verb and the remaining (parts of the) verb(s). In (17) the inversion of the finite verb *is* 'is' and the subject *dinen tijt* 'your time' is taken as an indication of V2. This inversion is triggered by the topicalization of *seer cort* 'very short'. In (18) the finite verb *was* 'was' appears in second position, after the subject *hij* 'he'. The presence of other linguistic elements (the direct object *de tijd* 'the time' and the adverb *volkomen* 'completely') between the finite verb and the participle *vergeten* 'forgotten' is taken as further evidence in favor of a V2-analysis.

(17) *Ghi moet dat vale peert beriden / **want** seer cort <u>is</u> dinen tijt*
(Devoot ende profitelyck boecxken, 1539)
You must the dun horse ride / because very short is your time
'You have to ride the dun horse / because your time is very short.'

(18) *Mary zou minder goed te spreken zijn* ***want*** *hij* *was* *de tijd volkomen vergeten (…).*
(Vervuld van verlangen, 1990)
Mary would less well to speak be because he was the time completely forgotten
'Mary would be in a bad mood because he had completely forgotten the time'

The second category is labeled *V-late* (or *subordinate use*) and includes all clauses that unambiguously do not show verb second. In *V-late*-fragments it can be shown that the finite verb does occupy a position within the verb phrase. This is the case if the finite verb occurs after the other (parts of the) verb(s), as in (19), or if other linguistic elements occur between the subject and the finite verb, such as *pays* 'peace' in (20).

(19) *Als dit coninc Karel sach was hi droevich* ***om dat*** *hem Reynout ontreden* *was.*
(Historie van den vier heemskinderen, 1508)
'When king Charles saw this he became sad because Reynout rode away from him.'
(20) *(…) si droeghen die tijdinghe in die stadt.*
daer grote blijschap was. ***om dat*** *si pays hebben* *souden.*
(Historie van Margrieta van Lymborch, 1516)
'(…) they brought the message into town,
where great joy arose, because they would have peace.'

As the following two examples illustrate, it is not always possible to distinguish between coordinate and subordinate use of the connectives at hand. This use of *want* or *omdat*, in which neither the presence nor the absence of V2 can be established, is labeled *V-ambiguous* (or *ambiguous use*). For example, the connective clause in (21) only contains a subject and a finite verb; in other words: although the finite verb occupies the second linear position in the connective clause, there are no other linguistic elements which can indicate whether the finite verb occupies a hierarchical position within the VP or within a higher functional projection. Ambiguity may also arise from extrapositioning of certain linguistic elements within the connective clause. Especially in Middle Dutch, extrapositioning is a frequent phenomenon, which differs both quantitatively and qualitatively from extrapositioning in Modern Dutch. In (22) and (23), the linear position of the finite verb again hints at V2. However, in both cases this linear order does not necessarily imply V2, since it can also result from right-extrapositioning of the complement clause (as in (22)) or – as is the case in (23) – of the indirect object *van hem* 'from him'.

(21) *men las doen opten seluen sondach / tes conuents messe als men plach /*
dees euwangelie ***want*** *si* *behoerde* (Sinte Lutgard, 1275-1300)
'Then, on that same Sunday, at the convents' Mass one read – as usual – from this gospel because it was her turn'
(22) *Menne vergat sijns om de spise. / De rouers mercten dese wise.*
Ende waer so was een huvs besloten. / Liepsi vp met hare roten.
Om dat *si* *waenden* *dat menre at.* (Rijmbijbel, 1275-1300)
'People forgot their properties because of the food. The robbers noticed this habit. And where a house was empty they all went together, because they thought that people were eating.'

(23) *Sems gheslachte hadde asia. / Cam egypten ende affrica.*
Japhet frigien ende europen. / Die tve gheslachten laten wi lopen.
Ende tellen mest van sem. / ***Vm dat*** *marie* *cam* *van hem.* (Rijmbijbel, 1275-1300)
'Shem's family had Asia. Ham Egypt and Africa. Japheth Frigia and Europe. Those two families we let go and tell most about Shem, because Mary descends from him.'

The second syntactic analysis involves an analysis of the linearization of the connective clauses as a whole. An analysis based on the positioning of the connectives themselves is disregarded here. As complementizers, *want* and *omdat* occupy a fixed position at the beginning of (or according to traditional analyses: even outside) the clause. In order to perform this linearization analysis, the clauses related by the connective were isolated from their context. Being so-called backward causals, both *want* and *omdat* always head the antecedent-clause of the two causally related clauses. This antecedent-clause can appear in a *preposed* position, that is, before the sentence functioning as the consequent (as in (24)). It can also take an intermediate or *intraposed* position, that is, between the different constituents of the consequent-clause, as in (25). The antecedent-clause can furthermore occur after the consequent-clause, resulting in a *postposed* position of the connective at stake. This is the case in (26).

(24) [$_{S1}$ ***Omdat*** *de kamer niet verlaten is*], [$_{S2}$ *schrikken ze allebei behoorlijk.*]
'**Because** the room is not deserted, they are both pretty startled.'
(25) [$_{S1}$ *Allebei schrikken ze* [$_{S2}$ ***omdat*** *de kamer niet verlaten is*] *behoorlijk.*]
'The both are – because the room is not deserted – pretty startled.'
(26) [$_{S1}$ *Allebei schrikken ze behoorlijk*], [$_{S2}$ ***omdat*** *de kamer niet verlaten is.*]
'They are both pretty startled, **because** the room is not deserted.'
(Het land aan de horizon, 1993)

Both word order and linearization can be used to derive the co- or subordinating nature of the connectives at hand. Coordinators show V2 and do not occur in the preposed position, whereas subordinators show V-late and can occur in all three positions.

5.3.2 Closure analysis

Each of the 150 connective fragments was also subject to an analysis based on closure. During this analysis I determined whether the connective clause should be linked to the immediately preceding clause (resulting in late closure: [A-[B-C]]) or to a combination of two preceding clauses ([[A-B]-C]).[4] In order to determine the hierarchical organization of the text structure, a criterion is needed to determine which units should be regarded as discourse segments. In line with Mann & Thompson (1988) and Sanders & Van Wijk (1996) I treat grammatical clauses as the basic discourse segments. A more precise definition of my *clause criterion* is given in (27).

[4] For ease of exposition, I refer only to the structural variants with the connective clause in postposed position. However, this analysis also applies to clause combinations with the connective clause in preposed position.

(27) Clause criterion:
  (i) Each clause is a segment, and a clause is a structure headed by a verb.
  (ii) Exceptions to (i) are: a) restrictive relative clauses b) clausal subjects, c) clausal complements, and d) restrictive adverbial clauses.

This definition shows that not all types of grammatical clauses are regarded as discourse segments. The first three exceptions in (ii) are mentioned in Mann & Thompson (1988: 248) and Sanders & Van Wijk (1996: 109); the fourth exception is taken from Pander Maat (1994: 33-36) and Schilperoord & Verhagen (1998: 158-161). The connections these exceptional clauses entertain with their matrix structures are not located on the level of discourse, but on the level of grammatical structure. Schilperoord & Verhagen (1998) argue there is good reason to exclude the four clause types in (ii) as separate discourse segments. They explain the exceptional status of restrictive relative clauses, restrictive adverbial clauses and the two types of complement clauses by reference to the notion of *conceptual dependency*: "one clause is conceptually dependent upon another clause, if its semantics cannot be conceptualized without essential reference to the conceptualization of another clause." In other words: "If a constituent of clause A is conceptually dependent on a clause B, B is an integral part of the conceptualization of A, and therefore not available as a separate discourse segment" (Schilperoord & Verhagen 1998: 150). Clause combinations with one of the four clause types in (ii) all exhibit this notion of conceptual dependency. For example, the adverbial B-clauses in (28) and (29) restrict the range of their A-clauses. They function as restrictive modifiers of the predicate and therefore should not be seen as separate discourse segments.

(28) [**A** *Vrouwen, zei Kat, wisten vanzelf wat ze doen moesten*] [**B** *als er een wild dier op hun pad sprong*], [**C** ***omdat*** *ze die verhalen zo vaak hadden gehoord.*] (Erin de Enige, 1990)
'Women, said Kat, automatically knew what to do if a wild animal crossed their path, because they had heard those stories so often.'

(29) *En* [**A** *Toon laat zich dat gemoeder maar aanleunen,*] [**B** *zolang het niet hinderlijk wordt.*] [**C** ***Want*** *als Marie al te zeer begint te drammen, trekt hij een bulldog-gezicht en gromt: "Mens, laat me de rest van de dag met rust! Je hebt vandaag al meer dan genoeg gedaan!"*] (De laatste wens, 1993)
'And Toon accepts that mothering, as long as it does not become too annoying. Because when Marie begins to nag too much, he pulls a bulldog face and grumbles: "Would you leave me alone for the rest of the day! You have done more than enough today!"'

In addition to the restrictive or complement clauses Sanders & Van Wijk (1996: 126) exclude yet another type of clause. In cases of contracted coordinate clauses, they consider the second conjunct as a separate discourse segment provided that only one major constituent is contracted. This implies that Sanders & Van Wijk exclude clauses in which more than one constituent is contracted. In my opinion, this approach is too strict. For example, their criterion would result in a wrong interpretation of (30). In this fragment, *omdat* in the D-clause marks a causal relation with the C-clause: asbestos is cheap because it can be extracted so easily. However, Sanders & Van Wijk would not regard the C-clause as a separate discourse segment, since both the subject *het* 'it' and the verb *is* 'is' are contracted. According to their analysis, the B- and C-clause must be regarded as one discourse segment. This analysis results in a wrong interpretation of the *omdat*-clause in D: the fact that asbestos is so

easy to extract cannot be regarded as an argument for the fact that it has important tensile strength.

(30) [**A** *Het* (= asbest) *isoleert uitstekend warmte en electriciteit,*] [**B** *is zeer trekvast*] *en* [**C** *vooral onwaarschijnlijk goedkoop,*] [**D** ***omdat*** *het zo gemakkelijk te winnen is.*]
(Degand 2001: 107)
'It (= asbestos) is a very good insulator from heat and electricity, it has important tensile strength and in particular [it is] incredibly cheap, because it is so easy to extract.'

In line with Pander Maat's criticism to the approach of Sanders & Van Wijk (see Pander Maat 1994: 272, footnote 11) my analysis does regard such clauses with more than one contracted element as separate discourse segments.

Using the clause criterion in (27) I determined per fragment which clauses were the A-, B- and C-clause respectively. I then established whether the combination [[A-B]-C] was possible at all, given the content of A (the first of the three clauses) and C (the connective clause). For instance, in (31) there is no other alternative than a late-closure [A-[B-C]] interpretation, because the propositional content of the connective clause cannot be linked to the propositional content of the A-clause in a sensible causal way. What Spinoza formulates in his book, can only be regarded as a reason to go on reading right away (i.e. the B-clause).

(31) [**A** *Hij vulde het glas met wodka*] *en* [**B** *las meteen verder*] [**C** ***omdat*** *Spinoza drie leefregels formuleerde die de lezer konden helpen bij het verkrijgen van een verbeterd verstand*]. (Hoffman's honger, 1993)
'He filled his glass with vodka and went on reading immediately because Spinoza formulated three regimens that could help the reader in obtaining an improved mind.'

For those fragments in which the propositional content of the three clauses did not exclude a [[A-B]-C] combination, I determined whether this was indeed the intended reading (given the context of the fragment) or if a late-closure interpretation should be derived instead.

5.4 Results of the diachronic analysis of *want*

In the current section and in section 5.5, the qualitative and quantitative results on *want* and *omdat* are discussed respectively. An overview of the quantitative data and statistical results per connective is given in Appendix C (*want*) and Appendix D (*omdat*). In the following sections, I will only discuss the results and the theoretical implications of the most adequate model in the logit-analysis. The significance of specific factors will only be mentioned if they are relevant for the discussion.

The current section provides the results of the diachronic analysis of *want*. First, an answer is given to the question: Did any syntactic changes occur during the selected time span? Then, the conceptual analysis based on closure is presented, as well as the interaction between closure and word order (5.4.2).

5.4.1 Syntactic analysis of *want*

At the syntactic level, modern Dutch *want* functions like English 'for', although it has nothing of the same stylistic formality. Like 'for', *want* in the 20th century can be described as a coordinator; as (32) and (33) illustrate, it shows word order patterns typical of any main

clause.[5] Similar word orders appear in 13th- and 16th-century sentences headed by *want* (see (34) and (35)).

(32) *De verbouwing heeft ook haar weerslag op het aantal tentoonstellingen,* ***want*** *die* <u>*zullen*</u> *in dit jaar en het volgende jaar iets in aantal afnemen.* (MC, 1995)
'The renovation also has its repercussions on the number of exhibitions, because those will slightly decrease in number in this and the following year.'

(33) *Er moet hem een behoorlijk bedrag nagelaten zijn,* ***want*** *zijn ouders* <u>*waren*</u> *redelijk welgesteld.* (In de schaduw bloeien de rozen, 1994)
'A substantial amount must have been left to him, because his parents were fairly well-off.'

(34) *Heues du luttel wagte dat du dar v*^{o}*mbe nit vrek en siis. mar milde.* ***Want*** *man* <u>*mag*</u> *wel milde wesen van klenen gu*^{o}*de* (NM, 1270-1290)
'If you possess little, be aware that you are not mean because of that, but generous. Because one is allowed to be generous with little.'

(35) *Beyaert dede mede grote moert* ***want*** *het* <u>*dede*</u> *menigen ridder de sadel rumen met biten ende slaen.* (Historie van den vier heemskinderen, 1508)
'Beyaert (= a horse) also did terrible things because it unhorsed many knights with biting and kicking.'

But medieval *want* could just as easily be combined with word orders more typical of subordination. For example, the finite verb *was* 'was' in (36) is preceded by the subject *et* 'it' and a temporal preposition; in (37) the finite verb *hadde* 'had' is preceded by both the subject and a direct object. In these cases, *want* syntactically behaves more like Modern Dutch subordinating *omdat* 'because'.

(36) *Doen dit ghedaen was, soo ghingen si ter maeltijt (****want****et byder noenen* <u>*was*</u>*)* (Historie van Malegijs, 1556)
'When this was done, they went to have a meal (because it was almost noon)'

(37) *Ende* ***want*** *sij geen teykenen en* <u>*hadde*</u> *om haer pater noster aen te spreken,* ***soe*** *hadde sij een vergaderinge van steenkens; daer plach sij Gode haer ghebedekens aen te tellen.* (Leven van Sinte Clara, 1500-1520)
'And because she didn't have a rosary to say an Our Father, she used a collection of pebbles to say her prayers to God.'

Example (37) also shows other characteristics indicating a subordinating nature of *want*: in (37), *want* is preceded by the conjunction *ende* (which generally is taken to be impossible with coordinators, cf. Greenbaum 1969). This example also shows a double marking of the coherence relation: the correlative or resumptive element *soe* in the second of the two combined clauses strengthens the causal relation marked with *want* in the first clause (see Burridge (1993: 54) for similar examples).[6] No wonder, then, that the *MNW* describes *want* as a conjunction, not specifying whether it is subordinating or coordinating.[7]

---

[5] More precisely, *want* is a restricted coordinator, since it cannot conjoin subordinate clauses and cannot show ellipsis of the subject (see Van Dijk 1979; Van Belle 1989; Burridge 1993: 53; Van der Heijden 1999).

[6] In combined clauses showing correlation two corresponding items are used to express the coherence relation; one is present in the first clause of the combined clauses, the other in the second (König &

From a Modern Dutch perspective it is remarkable that medieval *want* could head subordinating clauses at all, because in normal Modern Dutch language use this is not acceptable anymore. As Figure 5.1 shows, this intuition is confirmed by the disappearance of V-late occurrences in the 20th century ($z = -2.20$; $p = .03$). Together with a significant decrease in the number of ambiguous word orders ($z = -3.11$; $p < .001$), this results in a distribution over word orders that is not stable across ages ($\chi^2 (4) = 21.7$; $p < .001$).

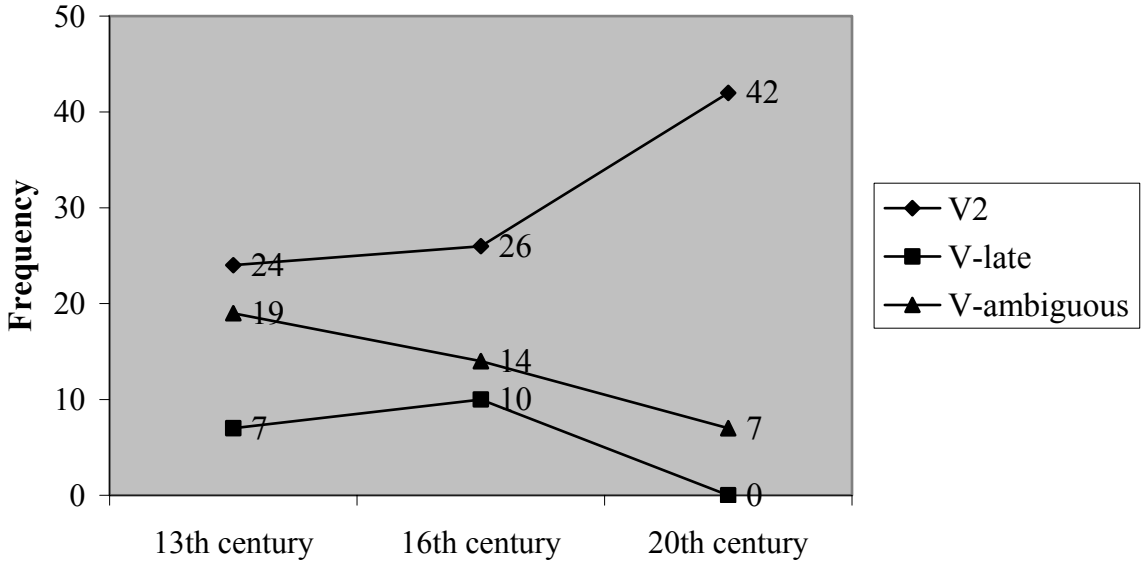


Figure 5.1. *Want*: Word order frequencies in three periods[8]

The observation that V-late *want* is still present in the 16th century is in sharp contrast to claims in the *MNW* that this use has disappeared by that time. However, it is in line with Van Megen's (2002) claim that the subordinating use continues to occur for a longer period than traditionally assumed. Still, the distribution presented in Figure 5.1 clearly shows that the subordinating use was not the most prominent of 13th-century *want*; an observation that goes against Van Megen's conclusion based on utterances in the *MNW* (see Van Megen 2002: 25).

A critical reader might argue that the V-late occurrences can be explained by reference to deviating word orders because of rhyme demands, since half of the fragments stem from rhyme texts. This explanation can be ruled out, since Figure 5.2 reveals that half of the V-late fragments appear in non-rhyming texts.

Van der Auwera 1988 use the term 'resumption' for the same phenomenon). Correlative connectives are usually comprised of a conjunction occurring in the subordinate clause and a linking adverb or conjunct in the matrix clause, which has the effect of echoing or emphasizing the relationship indicated by the subordinator (Burridge 1993: 21). Examples are [*omdat* 'because' S1, *daarom* 'therefore' S2] or – as in this case – [*want* S1, *so* S2].

[7] The *MNW* also mentions adverbial and prepositional use. These uses, however, are not attested to in my corpus. Burridge (1993: 55) even questions the classification of *want* as an adverb. Her extended study on word order in Middle Dutch gives no single example of *want* in combination with subject-verb inversion.

[8] In this analysis one 20th-century fragment without a verb has been left out of consideration.

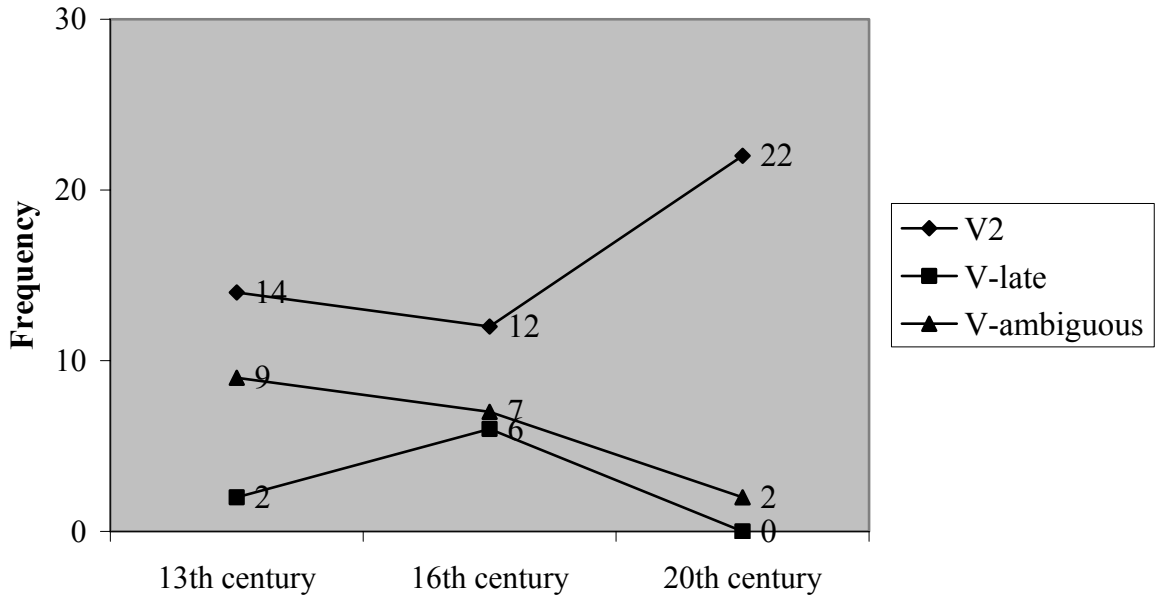


Figure 5.2. *Want*: Word order frequencies in three periods – only non-rhyming texts

The subordinating nature of several *want*-fragments is also supported by their analysis based on linearization. In Modern Dutch the *want*-clause – like other clauses headed by a coordinator – cannot be moved to precede the clause to which it is conjoined. This ungrammaticality accounts for the fact that none of the 20th-century *want*-fragments is preposed (see Table 5.2).

Table 5.2. *Want*: Linearization frequencies in three periods

	Preposed	**Intraposed**	**Postposed**	**Total**
13th century	0	1	49	50
16th century	3	0	47	50
20th century	0	2	48	50
Total	3	3	144	150

However, Medieval *want* was different. The examples in (34), (35), and (36) already showed that Medieval *want*-clauses can occur in the postposed position. As the following preposed examples illustrate, the Middle Dutch *want*-clause is not sequentially fixed.

(38) *God men verwacht v ende louet / Ende v betaemt geloeft te sijn*
[S1 ***Want*** *ghi tghebet verhooren* <u>*moghet*</u>] [S2 *Coempt alle vleesch voer v gheloefte doen*]
(Devoot ende profitelyck boecxken, 1539)
'God, one expects and praises You / And You ought to be praised
Because You can answer the prayers / All flesh comes to take the vows for You'

(39) *(…)* ***want*** *hi ons so scandelic* <u>*heeft*</u> *verraden, so en sel ic hem gheen genade doen*
(Historie van den vier heemskinderen, 1508)
'(…) because he has betrayed us so scandalously, I will not show mercy to him'

Again, the preposed occurrences cannot be explained by reference to rhyme demands, since two of the three preposed fragments are taken from non-rhyming texts.

The preposed *want*-fragments support the subordinating status of *want* in yet another way: they exclude an alternative account for the occurrence of V-late in *want*-clauses. In the older versions of Middle Dutch, certain main clauses could show V-late, despite the absence of a subordinating complementizer (cf. Van Gestel, Nijen Twilhaar, Rinkel & Weerman 1992: 102-104). It might be argued that the *want*-clauses showing V-late should also be regarded as this type of main clause, which implies that *want* in these cases still functions as a coordinator. However, this analysis is not possible when the *want*-clause is preposed, since clauses introduced by a coordinator cannot be preposed in general. The preposed fragments support the analysis that *want* sometimes triggers V-late itself, implying that it can also be regarded as a subordinator.

From the data and the discussion above, it can be concluded that medieval *want* indeed seems to hover between coordinator and subordinator, and that its clauses show word order and linearization characteristics of both of them. The subordinating use is maintained at least until the 17th century (cf. Van Megen 2002: 25).

5.4.2 Closure analysis of *want*

What does the conceptual picture of *want* look like and how does it relate to the word order properties mentioned in the previous section? The conceptual analysis based on closure reveals that there are only four fragments in which a combination of three clauses is possible at all. The other fragments do not allow for a three-clause combination (neither [[A-B]-C] nor [A-[B-C]]) because of the propositional content of the A- and C-clause. In all of the four three-clause combinations, an interpretation with the combination [[A-B]-C] is certainly possible. Two of these fragments show a preference for this interpretation (see (40) and (41)).[9]

(40) *Hi sende minen vader hier.* [**A** *Niet dat hi soude ontheren.*] [**B** *Mar v dinc ten besten keeren.*] [**C** ***Want*** *war hi comen als viand. / Tfolc tonterne ende v land. Hi adde commen dese pord storen.*] (Rijmbijbel, 1275-1300)
'He sent my father here. / Not to dishonor you / But to turn your case for the better. Because had he come as an enemy / he would have destroyed this place, and dishonor the people and the country.'

(41) [**A** *als ghy tot Montalbaen comt . suldi terstont totten coninc gaen*] [B *ende seggen hem dat hi gheringe vliet,*] [**C** ***want*** *ist dat hem Reinout vint, hi sal hem doen hangen . of quader doot doen sterven*] (Historie van den vier heemskinderen, 1508)
'When you come to Montalbaen, you must go to the king directly and tell him to run away immediately, because if Reinout finds him, he will hang him or give him some other evil death.'

The interaction between word order and the conceptual analysis based on closure is in line with the diachronic hypotheses inferred from Verhagen's theory. None of the *want*-examples with V-late allow for an interpretation other than late closure (in fact, none of these examples allows for a three-clause combination at all). The fragments in which the [[A-B]-C] combination is possible do not exhibit V-late.

---

[9] Both C-clauses in (40) and (41) are complex clauses at the grammatical level. Since the preposed conditional clauses within these C-clauses are of the restrictive type, these grammatical clause complexes form one segment at the discourse level (see the clause criterion in section 5.3.2).

5.5 Results of the diachronic analysis of *omdat*

This section provides the diachronic analyses of *omdat* based on syntactic properties (5.5.1) and closure (5.5.2). More details on the statistical analysis are presented in Appendix D.

5.5.1 Syntactic analysis of *omdat*

A first observation at the syntactic level is that *omdat* in the 13th century did not always appear as *omdat*. Other variants were *om-dat-dat* and *daar-om-dat*, as the following examples show.

(42) *Wine sin in diese werelt niet al ene vomb vons*
Mar ***vomb dat. dat*** *manlik anderen helpen sal* (NM, 1270-1290)
'We are in this world not only for ourselves, but because men should help each other.'

(43) *Pinse in dir herten di*^{e} *du*^{o}*gde inde di*^{e} *quatheit di*^{e} *gesciin mu*^{o}*gen.*
dar v^{o}***mbe dat*** *du di*^{e} *du*^{o}*gede mu*^{o}*ges maten inde di*^{e} *quaitheit gedogen* (NM, 1270-1290)
'Think about the virtue and evil that may happen, so that you will assume virtue and resist evil'

The frequencies of the three appearances of *omdat* can be found in Table 5.3. In the syntactic and conceptual analyses discussed below, the distinction between these forms is not taken into account, since they all express a causal relation.

Table 5.3. *Omdat*: Appearance in three periods

	Om-dat	**Om-dat-dat**	**Daar-om-dat**	**Total**
13th century	36	9	5	50
16th century	50	0	0	50
20th century	50	0	0	50
Total	136	9	5	150

Both the occurrence of *om-dat-dat* and the appearance of *daar-om-dat* support the claim that *om-dat* in the 13th century cannot be considered as one lexical unit. This is in line with the explanation for the genesis of *omdat* mentioned in the literature. For example, Weerman (1989: 194-196) argues that *omdat* arose via a process of reanalysis out of the combination of the preposition *om* and the relativum *dat* (an elliptical variant of *om die reden dat* 'for the reason that'). He claims that the presence of the subordinating complementizer *dat* was optional in Middle Dutch, but that this possibility to leave out the complementizer gradually disappeared. A lexical filler for the complementizer position could be obtained by reinterpreting the two-word combination *om+dat* as a one-word complementizer *omdat*.[10]

[10] A second way to obtain a lexical complementizer is by inserting *dat* in the complementizer position. Within this structural configuration *om* remains a preposition within the matrix clause. Weerman (1989: 195-196) states that this second analysis is equally tenable for modern Dutch *omdat*. In this case, the preposition *om* obligatorily selects the complementizer *dat* at the head of a finite subclause (see (i)). This analysis can also be applied to a preposition like *na* 'after', but not to modern Dutch *voor* 'before' and *tot* 'till' (see (ii)-(iv); all examples taken from Weerman 1989: 196).

(i) *We denken* ***om-dat/*om*** *we bestaan.*
'We think for-that/*for we exist.'

(ii) *We denken* ***na-dat/*na*** *we bestaan.*
'We think after-that/*after we exist.'

Both a reanalysis of constituency and a reanalysis of the word category were needed to transform the combination of *om* plus *dat* into the complementizer *omdat*.[11, 12]

The *daar-om-dat*-fragments support this analysis of the genesis of *omdat*. In this combination, *daar* refers cataphorically to the content of the clause introduced by *dat*. Since *daar* can combine with a preposition, but not with a complementizer, it can only occur in combination with *om* if this word is regarded as a (prepositional) lexical item independent of *dat*. The fact that the combined form *daar-om-dat* does not occur anymore in the 16th century leads to the conclusion that the reanalysis into the complementizer *omdat* must have taken place before that period.

The *om-dat-dat*-fragments also support the genetic analysis of *omdat*. In the 13th-century combination *om-dat-dat* (see (42)), *om* functions as a preposition within the matrix clause, with the second *dat* as a relativum that heads the subordinate clause. The intermediate *dat* in this combination is the singular form of a neuter demonstrative pronoun with deictic force; in (42) it is used "to point to" the subordinate clause "men should help each other".[13]

A second syntactic observation is that the connective *omdat* hardly occurs with the V2 word order. Figure 5.3 shows the only two V2-fragments in the corpus stem from the 13th century. Statistical analysis (see Appendix D) reveals that the distribution over word orders is not stable across ages (χ^2 (2) = 6.7; $p < .05$); in the 20th century *omdat* occurs less often with the ambiguous word order than in the 13th century ($z = -2.31$; $p = .02$).[14]

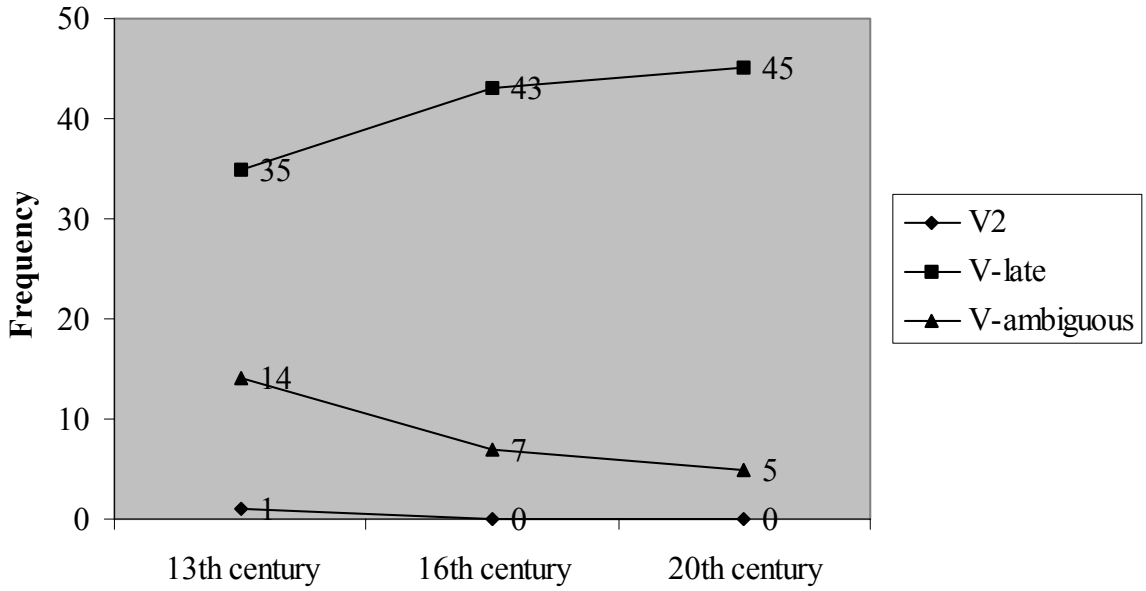


Figure 5.3. *Omdat*: Word order frequencies in three periods

(iii) *We denken* ***voor-dat/voor*** *we bestaan.*
'We think before-that/before we exist.'

(iv) *We denken* ***tot-dat/tot*** *we bestaan.*
'We think till-that/till we exist.'

Unlike *om* and *na*, the words *voor* and *tot* are able to function as complementizers themselves.

[11] See Harris & Campbell (1995: 287) for further remarks on different types of reanalysis.

[12] Dubinsky & Williams (1995) also relate the rise of certain complementizers to a diachronic process of reanalysis: several English temporal words (e.g. *after*, *before*, *while*) originally only functioned as prepositions. Nowadays they can also be used as complementizers introducing finite subclauses.

[13] Weerman (1989: 199-203) relates the doubling of *dat* to the rise of the relativum *dat* out of the pronoun *dat*.

[14] In this analysis, the V2 fragment is grouped together with the ambiguous fragments.

The fact that there is only one *omdat*-fragment showing V2 (see (44)) makes it interesting to perform a more detailed analysis into the nature of this fragment.

(44) *Die onghehuwede selen hier comen,*
Om dat *si* *hebben* *daer vernomen,*
Hoe mijn huwe ghesproken es; (Parthonopeus van Bloys, ± 1350)
'The unmarried will come here, because they have been told there how my marriage is celebrated.'

An alternative explanation for the V2 word order in this fragment may come from the syntactic process of Verb Projection Raising (cf. Haegeman & Van Riemsdijk 1986). This process, in which the infinitive is moved together with its complements, was possible in Middle Dutch, and it is still a common process in Modern Flemish (cf. Blom 2002: 15, Haegeman 1994). Application of this syntactic operation to the constructed V-late clause in (45)a illustrates that this operation can result in V2 (see (45)b).[15]

(45) a. *Nathan en Mirjam trokken hun jas aan, omdat ze naar buiten gaan wilden.*
b. *Nathan en Mirjam trokken hun jas aan, omdat ze wilden naar buiten gaan.*
'Nathan and Mirjam put on their coats, because they wanted to go out.'

The fragment in (44), however, cannot be analyzed as a case of Verb Projection Raising. Haegeman (1994: 511) mentions that the range of verbs triggering "Verb Projection Raising in their nonfinite complements in West Flemish is restricted": it includes perception verbs, causative verbs, modal verbs, and aspectual verbs. The auxiliary *hebben* 'have' in (44) does not belong to one of these categories. Hence, the remarkable V2 word order in this fragment asks for yet another explanation.

This explanation may come from the fact that the fragment in (44) is taken from a rhyme text. The fact that 'coordinating' *omdat* does not occur in non-rhyming texts (but only in a rhyming one), seems to indicate that the combination V2-*omdat* is not productive in the 13th century. This V2-occurrence can be explained, then, by reference to a deviating word order because of rhyme demands.

The subordinating nature of *omdat* is supported by the distribution based on linearization: *omdat*-clauses frequently occur in preposed position. Statistical analysis of the distribution in Table 5.4 shows that *omdat* occurs more frequently in a postposed than in a preposed position (χ^2 (1) = 49.2; p < .001).[16] Over the years, this distribution remains relatively constant.

Table 5.4. *Omdat*: Linearization frequencies in three periods

	Preposed	**Intraposed**	**Postposed**	**Total**
13th century	15	0	35	50
16th century	10	1	39	50
20th century	7	0	43	50
Total	32	1	117	150

[15] Note that both variants in (45) are acceptable in Modern Flemish, but that in Modern Dutch only the a-variant is grammatical.
[16] In this analysis, the intra- and postposed fragments are taken together, because of the very low number of intraposed fragments.

Given the diachronic data on the word order and linearization properties of *omdat*, in the remainder of this chapter I will treat *omdat* as a subordinating connective.

5.5.2 Closure analysis of *omdat*

In addition to the syntactic analysis, I also performed a conceptual analysis based on closure. The majority of the *omdat*-fragments did not allow for a three-clause combination because the propositional content of the C-clause could not be related to the propositional content of the A-clause in a sensible way. The closure analysis could be applied to 11 *omdat*-fragments (see Table 5.5). The label [A-[B-C]] represents three-clause combinations with a late-closure interpretation. The label [[A-B]-C] represents fragments in which the *omdat*-clause (C) combines to the clause-complex [A-B] as a whole. Although these labels may suggest otherwise, Table 5.5 includes both postposed and preposed *omdat*-clauses that appear in combination with two other clauses.[17]

Table 5.5. *Omdat*: Frequencies based on closure

	[A-[B-C]]	[[A-B]-C]
13th century	3	4
16th century	3	0
20th century	1	0
Total	7	4

Example (46) illustrates that V-late *omdat* can indeed trigger a late-closure [A-[B-C]] interpretation. Remarkably, however, it also allows for the combination [[A-B]-C]. There are two such fragments with a preposed *omdat*-clause (one of which is presented in (47)), and one fragment with a postposed *omdat*-clause (see (48)).

(46) [**A** *Zij* (= Célestine) *knikte hem met een fijn glimlachje toe,] en* [**B** *gaf niet meer uitleg,*] [**C** ***omdat*** *Sarah al teruggelopen was en naast haar stond.*] (Het verbroken zegel, 1991)
'She (= Célestine) nodded to him with a subtle smile, and did not explain anymore, because Sarah had returned already and stood next to her.'

(47) *Mar* [**C** ***v^{o}mb dat*** *hi* (= de leeuw) *naturlik ku^{o}nheit an hu^{o}me heuet.*] [**B** *so scaimt hi v^{o}me angst te hebene.*] [**A** *Jnde lopt den man v^{o}p. als hastelik als di^{e} man dru^{o}p siit.*]
(NM, 1270-1290)
'But because he (= the lion) has a natural courage in him, he is ashamed of having fear and attacks the man as soon as he looks at him.'

(48) *Andermans misdat di^{e} bistu sculdig gerne te vergeuene. Mar* [**A** *dine misdait di^{e} du misdu^{o}s di^{e} ne sols du nit vergeuen.*] *Mar* [**B** *du^{o} sols altos gedinchen.*] [**C** ***v^{o}mb dat****. dat du di tebat huden su^{o}ls in misdat te vallen.*] (NM, 1270-1290)
'Other people's crimes you have to forgive willingly. But you should not forgive your own crime, but you should always remember (it), so that you will prevent yourself from falling into crime.'

An analysis of the interaction between closure and word order reveals that Verhagen's hypothesis is not borne out. Contrary to his hypothesis, the two preposed *omdat*-fragments without a late-closure reading both exhibit V-late. This also holds for the postposed fragment

[17] A more adequate representation of fragments with a preposed *omdat*-clause (C) would be [[C-B]-A] and [C-[B-A]] respectively. For ease of reference I only use the labels mentioned in Table 5.5.

with the interpretation [[A-B]-C] given in (48)). There is one other postposed example (see (49)) in which *omdat* does not trigger late closure. However, this fragment cannot be regarded as a counterexample, since it shows an ambiguous word order (the particle *dod* 'dead' as well as the complement *Gods propheten* 'God's prophets' are extraposed).

(49) [**A** *Die* (= Abdias) *hadde hemelike bestolen. / .C. propheten in tveen holen.*] *Ende* [**B** *gaf hem burne ende brod.*] [**C** ***Om dat*** *iesabel sloech dod. Gods propheten dar soese bevinc.*]
(Rijmbijbel, 1275-1300)
'Abdias had secretly hidden / thousand prophets in two caves / and gave them spring water and bread / because Isabel killed God's prophets when she found them.'

The closure analyses of *omdat*, then, provide counterevidence to Verhagen's hypothesis.

5.6 Conclusions and discussion

In section 5.2.3 two hypotheses concerning the interaction between word order and closure were formulated. These hypotheses will be taken up for further discussion in section 5.6.1; section 5.6.2 presents a discussion on the decrease in ambiguous word orders.

5.6.1 On closure

The hypotheses about the interaction between word order and closure are repeated in (50).

(50) <u>Hypotheses for the diachronic study:</u>
a. The Middle Dutch *want*-fragments showing V-late are all cases of late closure.
b. All of the *omdat*-fragments showing V-late are cases of late closure.

Firstly, it was expected that the Middle Dutch subordinating *want*-fragments are all cases of late-closure. This prediction is borne out: all the [[A-B]-C]-combinations show V2, whereas the V-late *want*-fragments have no other alternative than late closure. In none of the V-late fragments can the propositional content of the connective clause be linked to the propositional content of the A-clause in a sensible way.

Secondly, *omdat* should always show up in late-closure fragments (with the hierarchical structure [A-[B-C]]). This prediction is not supported by the *omdat*-fragments showing V-late. Several V-late fragments indeed force a late-closure reading, but there are also three V-late *omdat*-fragments that exhibit an [[A-B]-C]-combination. These three fragments from the 13th century present counterevidence to the hypothesis in (50)b.

The closure analyses in this chapter provide a first test of the hypothesis put forward by Verhagen (2001). The data from the 16th and 20th centuries are in line with his theory, but the *omdat*-results from the 13th century indicate that his claim about the interaction between word order and closure needs to be modified. The fact that preposed *omdat*-clauses do not always receive a late-closure interpretation indicates that his hypothesis should be restricted to postposed V-late clauses. Excluding preposed V-late clauses from the interaction hypothesis seems a plausible option: if a language user wants to produce a three-clause combination with a preposed connective clause, he cannot select the V2 word order. Since connective clauses showing V2 are not allowed in preposed position, he needs to use the V-late word order.

If the interaction hypothesis is indeed restricted to postposed connective clauses, only one subordinating *omdat*-fragment remains as counterevidence to Verhagen's hypothesis. This 13th-century fragment is interpreted using 20th-century intuitions, and hence provides a very weak basis for rejecting the modified hypothesis concerning the present-day interaction

between word order and level of attachment. Further research into three-clause combinations in Modern Dutch is imperative, then, to find support in favor of or against the modified version of Verhagen's hypothesis.

5.6.2 On the decrease in ambiguous word orders

From the historical survey presented in this chapter it became clear that both *want* and *omdat* show a significant decrease in ambiguous word orders. The higher number of ambiguous fragments in the 13th and 16th centuries does not come as a surprise, because it is already known from the literature that Middle Dutch clauses have a more flexible structural nature (cf. Burridge 1993: 222; Van Gestel et al. 1992: 113). Middle Dutch shows grammatical possibilities that Modern Dutch does not have (to such an extent). One of these is the possibility to extrapose all kinds of objects and other constituents. An example is given in (51), in which the direct object *dat soete kindekijn* 'that sweet child' is extraposed.

(51) *daer om moesten sijt besteruen / veel kinderen dats wel aenschijn*
om dat *hi* <u>*wilde*</u> *bederuen / dat soete kindekijn* (Suverlijc boecxken, 1508)
'That's why many children had to die / that has become apparent
because he wanted to ruin / that sweet child'

In Modern Dutch only complement clauses, so-called 'heavy' noun phrases and certain preposition phrases can be extraposed (see the *ANS*, Haesereyn et al. 1997: 1364-1387). So, in Modern Dutch only sentences like (52) and (53) are classified as *ambiguous*. The ambiguity in (52) results from the absence of constituents other than the subject and the verbs; the ambiguity in (53) results from the right-extraposition of the complement clause (and the absence of other constituents). The fact that *want* can be replaced by modern Dutch *omdat* in these examples supports the analysis that the word orders in these clauses are really ambiguous.

(52) *De Chinees gaf mij nieuwtjes voor mijn vader mee,* ***want*** *onze radio* <u>*was*</u> *verzegeld*
(Indische duinen, 1994)
'The Chinese man gave me pieces of news for my father, because our radio was confiscated.'

(53) *Zijn ouders maken zich niet ongerust,* ***want*** *ze* <u>*weten*</u> *dat hij een goed ruiter is, prima in staat voor zichzelf te zorgen.* (Het land aan de horizon, 1993)
'His parents did not worry, because they knew that he was a good horseman, perfectly able to take care of himself.'

It can be concluded that the decrease in ambiguous fragments is not related to a change in the structural properties of the connectives *want* and *omdat* themselves, but to a change in the extraposition possibilities of Dutch clauses in general. In the following chapter, the syntactic changes in the use of *want* and *omdat* mentioned in the current chapter will be related to an alternative conceptual analysis: one based on domains of use.

APPENDICES TO CHAPTER 5

Appendix C – Logit analysis of the diachronic development of *want*

I – Remarks

- Different numbers of words were needed to select 50 fragments per period: 34253 words for the 13th century, 14790 words for the 16th century, and 75799 words for the 20th century.
- As the total number of *want*-fragments indicates (see II-Data), three cases have been left out of the statistic analyses. These involve one *want*-fragment without a verb (20th century, postposed, content) and two *want*-fragments with the meaning 'whenever' (16th century, postposed, content).
- The factor *Linearization* has not been taken into account in the logit analysis because of the low number of preposed *want*-clauses. Furthermore, the three intraposed *want*-fragments (one in the 13th century and two in the 20th century) have been placed under the postposed category.

II – Data: Distribution of *want* in three periods

		Content		Epistemic		Speech act		Total
		Pre	Post	Pre	Post	Pre	Post	
13th century	V2	0	3	0	19	0	2	24
	V-late	0	5	0	2	0	0	7
	V-ambiguous	0	7	0	9	0	3	19
16th century	V2	0	5	0	13	0	8	26
	V-late	3	3	0	1	0	1	8
	V-ambiguous	0	4	0	5	0	5	14
20th century	V2	0	15	0	25	0	2	42
	V-late	0	0	0	0	0	0	0
	V-ambiguous	0	2	0	5	0	0	7
Total		3	44	0	79	0	21	147

III – Results logit analysis *want*

Logit (fr)	χ^2 model	df	p model	χ^2 factor	df	p factor
1. constant	214.56	26	< .001	-	-	-
+ 2. period	154.93	24	< .001	59.63	2	< .001
+ 3. domain	119.38	22	< .001	35.55	2	< .001
+ 4. word order	56.18	20	< .001	63.20	2	< .001
+ 5. period x domain	42.29	16	< .001	13.88	4	< .01
+ 6. period x word order	20.60	12	< .1	21.70	4	< .001
+ 7. domain x word order	4.40	8	< .9	16.20	4	< .005

IV – Parameter estimates *want* for model 7

Parameter	Estimate	s.e.	z-score	p
constant	-8.97	0.40	-22.46	< .001
period: 16th century	1.02	0.49	2.09	0.04
period: 20th century	0.33	0.46	0.73	0.47
domain: epistemic	1.38	0.43	3.23	0.001
domain: speech act	-0.64	0.63	-1.02	0.31
word order: V-late	0.15	0.55	0.27	0.79
word order: V-ambiguous	0.25	0.46	0.55	0.58
period x domain: 16th epistemic	-0.47	0.51	-0.93	0.35
period x domain: 16th speech act	1.06	0.66	1.59	0.12
period x domain: 20th epistemic	-0.69	0.50	-1.40	0.16
period x domain: 20th speech act	-1.37	0.92	-1.49	0.14
period x word order: 16th V-late	-0.04	0.65	-0.06	0.95
period x word order: 16th V-ambiguous	-0.52	0.48	-1.08	0.28
period x word order: 20th V-late	-4.12	1.87	-2.20	0.03
period x word order: 20th V-ambiguous	-1.63	0.52	-3.11	0.002
domain x word order: epistemic V-late	-2.42	0.73	-3.32	< .001
domain x word order: epistemic V-ambiguous	-0.72	0.47	-1.54	0.12
domain x word order: speech act V-late	-2.26	1.06	-2.13	0.03
domain x word order: speech act V-ambiguous	-0.16	0.62	-0.25	0.80

Appendix D – Logit analysis of the diachronic development of *omdat*

I – Remarks

- Different numbers of words were needed to select 50 fragments per period: 83400 words for the 13th century, 95505 words for the 16th century, and 61574 words for the 20th century.
- The category V-ambiguous includes one V2 fragment (13th century, postposed, epistemic).
- The postposed category includes one intraposed *omdat*-fragment (16th century, epistemic).

II – Data: Distribution of *omdat* in three periods

		Content		Epistemic		Speech act		Total
		Pre	Post	Pre	Post	Pre	Post	
13th century	V-late	9	19	2	3	0	2	35
	V-ambiguous	1	6	1	5	2	0	15
16th century	V-late	7	29	0	5	1	1	43
	V-ambiguous	2	3	1	1	0	0	7
20th century	V-late	6	23	0	15	1	0	45
	V-ambiguous	0	5	0	0	0	0	5
Total		25	85	4	29	4	3	150

III – Results logit analysis *omdat*

Logit (fr)	χ^2 model	df	p model	χ^2 factor	df	p factor
1. constant	280.69	35	< .001	-	-	-
+ 2. period	275.54	33	< .001	5.20	2	< .1
+ 3. linearization	226.36	32	< .001	49.18	1	< .001
+ 4. word order	161.06	31	< .001	65.31	1	< .001
+ 5. domain	45.40	29	< .1	115.70	2	< .001
+ 6. period x word order	38.61	27	< .1	6.79	2	< .05
+ 7. linearization x domain	32.61	25	< .25	6.01	2	< .05

IV – Parameter estimates *omdat* for model 7

Parameter	Estimate	s.e.	z-score	p
constant	-8.35	0.18	-45.52	< .001
period: 16th century	0.07	0.23	0.31	0.76
period: 20th century	0.56	0.23	2.47	0.01
linearization: preposed	-1.22	0.23	-5.37	< .001
word order: V-ambiguous	-0.84	0.31	-2.74	0.01
domain: epistemic	-1.07	0.21	-4.43	< .001
domain: speech act	-3.22	0.55	-5.82	< .001
period x word order: 16th V-ambiguous	-0.95	0.51	-1.87	0.06
period x word order: 20th V-ambiguous	-1.26	0.55	-2.31	0.02
linearization x domain: preposed epistemic	-0.69	0.56	-1.23	0.22
linearization x domain: preposed speech act	1.46	0.76	1.92	0.06

Chapter 6

Word order and domains: a diachronic analysis of want *and* omdat

Chapter 5 presented the diachronic results for the word order properties of *want* and *omdat*. The current chapter relates these word order changes to an analysis of the two connectives based on domains of use. The interaction between word order and domains in the current chapter, and the interaction between word order and closure in Chapter 5 can be seen as alternative theories that do not necessarily exclude each other.

> *"(…) there is a close relationship between the choice of particular word order and the discourse-pragmatic function of the clause."* (Günthner 1996: 336)

6.1 Introduction

Section 3.5.1 (Chapter 3) provided evidence in favor of an interaction between domains of use and specific syntactic properties of the adverbial clause. The German data presented by Günthner (1993, 1996) and Keller (1995) indicated a link with word order. In addition, Haegeman's proposal could be characterized as a successful attempt to show that there is a link between source of coherence on the one hand and the internal and external syntax of the adverbial clause on the other.

The German data seem promising, but they do not necessarily imply that the interaction between word order and domains of use holds for the Dutch language as well. In fact, Verhagen (2001) even explicitly rejects this interaction for *want* and *omdat*. In the current chapter, then, I investigate the interaction between the text-linguistic characteristics of *want* and *omdat* based on domains and the word order patterns within the clauses they head. To this end, the diachronic developments of these complementizers are studied. Two questions will be answered for both *want* and *omdat*.

(1) <u>Research questions of this chapter:</u>
   a. Did any conceptual changes occur during the selected time span?
   b. Can these conceptual changes be related to the syntactic changes discussed in the previous chapter?

In section 6.2 the interaction hypothesis is worked out in more detail. This section will also introduce the hypotheses for the diachronic research. In section 6.3 (on *want*) and 6.4 (on *omdat*) I present the results from the corpus study. Finally, section 6.5 contains the conclusions, as well as a discussion of the main findings. The diachronic data in this chapter are in line with the hypothesis that there is a partial interaction between word order and domains. Further domain restrictions in the use of connectives seem to be triggered by connective-specific properties.

6.2 Word orders related to domains of use

In this section I will argue for an indirect relation between word order, specifically the positioning of the finite verb, and domains of use. My claim is that the word order properties of adverbial clauses can be related to different syntactic configurations, which in turn can be related to different interpretation based on domains. The three steps in my line of reasoning can be summarized as in (2).

(2) Line of reasoning:
a. word order interacts with syntactic configuration
b. syntactic configuration interacts with domains of use
c. word order interacts with domains of use

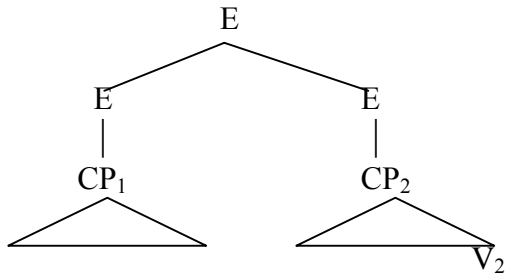
The theoretical basis for these three steps will be discussed in section 6.2.1 to 6.2.3 respectively. These sections also show how the relevant theories can be applied to the Dutch connectives *want* and *omdat*. The interaction claim in (2)c was put to the test on the basis of a diachronic corpus study on *want* and *omdat*. The hypotheses for this diachronic research are presented in section 6.2.4.

6.2.1 Word orders and their syntactic configurations

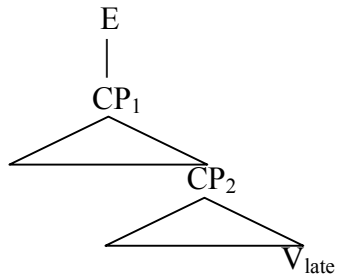
My first claim is that different word orders relate to different syntactic configurations. I will argue for four syntactic configurations; one of these combines with V2 and the other three with V-late. Two of the syntactic configurations can be found in De Haan (2001), Haegeman (2003) and Verstraete (2000); the other two syntactic constructions are based solely on Verstraete (2000).

De Haan (2001) distinguishes between two syntactic configurations.[1] He explicitly argues for a relation between word order of the connective clause on the one hand and its internal and external syntax on the other. His idea is that clauses showing V2 should be regarded as root clauses, which are unintegrated with respect to their matrix clause and which have an illocutionary force of their own. V-late clauses should be seen as non-root clauses integrated into their matrix clause and without an independent illocutionary force (De Haan 2001: 25). Schematically, these syntactic configurations can be presented as in (3).

(3) a.



b.



Applying the syntactic configurations in (3) to *want* and *omdat* results in the following claims. The V2-complementizer *want* can be said to connect two root CPs, both with their own illocutionary force (as represented schematically in (3)a). The subordinator *omdat*, on the other hand, connects a root CP with illocutionary force and a non-root CP without an illocutionary force of its own. In this view, a modern Dutch *omdat*-clause should be regarded as a CP subordinated to the matrix verb (as in (3)b). It can be shown that *want* and *omdat*

[1] See section 5.2.1 in the previous chapter for a more elaborate discussion of his theory.

indeed differ in their syntactic configuration by altering the illocutionary force of the clauses they head. The illocutionary force of *want*-clauses can be altered independently of the illocutionary force of the matrix clause (see the a-variants of (4)-(6), repeated here from section 5.2.1). As the b-variants show, *omdat* does not have this possibility; this connective can only occur in clauses of the default declarative type, unless the whole *omdat*-clause is part of a different speech act (e.g. interrogative, imperative) expressed by the matrix clause.

(4) a. *Nathan bood Mirjam een snoepje aan,* ***want*** *ze was verdrietig.*
b. *Nathan bood Mirjam een snoepje aan,* ***omdat*** *ze verdrietig was.*
'Nathan offered Mirjam a candy, because she was sad.'
(5) a. *Nathan bood Mirjam een snoepje aan,* ***want*** *was ze niet verdrietig?*
b. **Nathan bood Mirjam een snoepje aan,* ***omdat*** *was ze niet verdrietig?*
'Nathan offered Mirjam a candy, because wasn't she sad?'
(6) a. *Bied Mirjam een snoepje aan,* ***want*** *ze is verdrietig!*
b. *#Bied Mirjam een snoepje aan,* ***omdat*** *ze verdrietig is!*
'Go and offer Mirjam a candy, because she is sad!'

In addition, by applying the syntactic operations of clefting and *why* questioning to *want* and *omdat*, it can be shown that the two connectives differ in terms of their scopal properties (see (7)-(9)).[2, 3] Both syntactic operations impose an embedded construction on the example to be tested, demanding the connective clause function within the scope of the matrix clause (cf. Verstraete 1998: 194). Pit (2003: 19) argues that *why* questioning can easily be applied to *omdat*-clauses like (7)a (i.e. without changing meaning), whereas this operation is problematic for their counterparts marked with *want* (compare the a- and b-example in (8)). In addition, Van Belle (1989: 437) shows that clefting is only possible for *omdat* (cf. also Degand 1996: 162). This is illustrated in (9), the clefted versions of the utterances in (7). The fact that the application of the two operations to *want*-clauses results in unacceptable utterances supports the idea that *want*-clauses cannot occur within the scope of another clause and should be seen as an instantiation of the syntactic configuration in (3)a.[4, 5]

---

[2] Note that this property is not independent of the previous one. If both conjuncts have their own illocutionary value, this partial independence of the conjuncts naturally prevents one conjunct from falling within the scope of the interpersonal resources of the other (Verstraete 2000: 15).
[3] These are the Dutch equivalents of English examples mentioned in Verstraete (1998: 183).
[4] Note that the ungrammaticality of the *want*-examples in (8)b and (9)b cannot be put down to the conceptual properties of *want.* The connective *want* can and frequently does occur in semantic relations. This implies that the ungrammaticality of these examples has to be attributed to the syntactic properties of the *want*-clause.
[5] Verstraete (1998: 191) argues that two other syntactic operations can be applied to show the difference in scope: shift in polarity and shift in illocutionary force (see also Rutherford 1970: 98, 108). As De Vries (1971: 417-418) and Van Belle (1989: 437-439) have shown, these operations can be applied to *omdat*-clauses like (i)a without changing the original interpretation, but not to their counterparts marked with *want* (compare the a- and b-versions in example (ii) and (iii)). The brackets in (ii)a and (ii)b indicate that a wide-scope reading is intended: the negative phrase in the matrix clause has scope over the entire clause combination, indicating that John's anger must have been caused by something other than his brother damaging his car. This meaning can be obtained in (ii)a, but not in (ii)b, which implies that the *want*-clause must be outside the scope of the matrix clause. A similar line of reasoning holds for (iii): the a-version has a wide-scope reading in which the causal relation between John's anger and his brother damaging his car is questioned (there might be another reason for John's anger). This wide-scope reading is not available in (iii)b, which indicates that *want*

(7) a. *Jan is ziek,* ***omdat*** *hij te lang buiten in de regen* *is* *gebleven.*
b. *Jan is ziek,* ***want*** *hij* *is* *te lang buiten in de regen gebleven.*
'John is ill, because he has been out in the rain too long.'

(8) a. *Waarom is Jan ziek?* ***Omdat*** *hij te lang buiten in de regen* *is* *gebleven.*
b. *Waarom is Jan ziek? #****Want*** *hij* *is* *te lang buiten in de regen gebleven.*[6]
'Why is John ill? Because he has been out in the rain too long.'

(9) a. *Het is* ***omdat*** *Jan te lang buiten in de regen* *is* *gebleven dat hij ziek is.*
b. **Het is* ***want*** *Jan* *is* *te lang buiten in de regen gebleven dat hij ziek is.*
'It is because John has been out in the rain too long that he is ill.'

De Haan uses two syntactic properties to distinguish between the syntactic configuration in (3): the presence versus absence of illocutionary force in the connective clause, and the presence versus absence of a relation of scope between the connective clause and the matrix clause. The same defining properties as well as highly similar syntactic configurations can be found in the work of Haegeman (2003) and Verstraete (2000).[7] The only difference is that Haegeman and Verstraete do not mention a word order component, which is the logical result of the fact that they study English. The two syntactic configurations in (3) can be equated with Haegeman's distinction between *peripheral* versus *central* adverbial clauses (cf. Haegeman 2003: 330-336 on the presence of a functional projection encoding illocutionary force, and Haegeman 2003: 320-327 on the relation of scope). De Haan's syntactic configurations can also be inferred from Verstraete's (2000) work, which is formulated within a functional

---

must function outside the scope of the clause to which it is linked. Both the negated and the questioned *omdat*- and *want*-variants of (i) may receive a narrow-scope reading (see the c- and d-examples in (ii) and (iii)). For example, in (iii)c and d, only John's anger is questioned instead of the relation between John's anger and its possible cause. However, in these narrow-scope interpretations, the original content relation is either shifted to an epistemic relation (see (ii)) or to a speech-act relation (see (iii)).

(i) a. *Jan is [boos* ***omdat*** *zijn broer zijn auto beschadigde].*
b. *Jan is [boos* ***want*** *zijn broer beschadigde zijn auto].*
'John is angry because his brother damaged his car.'

(ii) a. *Jan is [niet boos* ***omdat*** *zijn broer zijn auto beschadigde].*
b. *#Jan is [niet boos* ***want*** *zijn broer beschadigde zijn auto].*
c. *Jan is [niet boos]* ***omdat*** *zijn broer zijn auto beschadigde.*
d. *Jan is [niet boos]* ***want*** *zijn broer beschadigde zijn auto.*
'John is not angry because his brother damaged his car.'

(iii) a. *Is Jan [boos* ***omdat*** *zijn broer zijn auto beschadigde]?*
b. *#Is Jan [boos* ***want*** *zijn broer zijn auto beschadigde]?*
c. *Is Jan [boos]* ***omdat*** *zijn broer beschadigde zijn auto?*
d. *Is Jan [boos]* ***want*** *zijn broer beschadigde zijn auto?*
'Is John angry because his brother damaged his car?'

The syntactic scope tests in (i)-(iii) show that *omdat* can occur within the scope of the matrix clause, while Dutch *want*-clauses cannot occur in such configurations.

[6] Note that this utterance is only ungrammatical in the interpretation in which the *want*-clause functions as an answer to the why-question, i.e. when it receives the same content-interpretation as the *omdat*-clause in (8)a. The *want*-utterance is only acceptable if it is regarded as the reason for asking the question, i.e. as a marker of a relation in the speech-act domain. The grammaticality of this second interpretation can be illustrated by the *want*-clause in (i), in which the content-interpretation is ruled out beforehand (hearing John cough cannot be the reason for his being ill).

(i) *Waarom is Jan ziek?* ***Want*** *ik hoorde hem vanochtend hoesten.*
'Why is John ill? Because I heard him coughing this morning.'

[7] In recent work, Verhagen (2005) argues for a semantic explanation of these data in terms of mental spaces.

instead of a generative framework. The configuration in (3)a can be equated with Verstraete's *free coordination construction*. It is characteristic of connective clauses in free coordination that they have an illocutionary force of their own and that they function outside the scope of their matrix clause (Verstraete 2000: 9). De Haan's configuration in (3)b can be equated with Verstraete's *bound subordination construction*. Connective clauses in bound subordination lack an illocutionary force of their own and function within the scope of their matrix clause (Verstraete 2000: 14). Applying Verstraete's construction terminology to the two Dutch connectives shows that *want* can occur in the free coordination construction, but not in the bound subordination construction, whereas the reverse holds for *omdat*.

The discussion so far implies a one-to-one mapping between word order and syntactic configuration: V2 combines with free coordination (represented in (3)a), whereas V-late is characteristic of bound subordination (see (3)b). However, Verstraete (2000) argues for two additional syntactic configurations (see Table 6.1), which implies that the mapping between word order and syntactic configuration cannot be one-to-one.

Table 6.1. Construction types according to Verstraete (2000)

Construction type	**Illocutionary force of its own?**	**Modal value of its own?**	**Outside the scope of the matrix clause?**
Free coordination	+	+	+
Bound coordination	-	+	+
Free subordination	-	-	+
Bound subordination	-	-	-

In the remainder of this section I will give some more explanation of Verstraete's *bound coordination construction* and *free subordination construction*, and show that *omdat* can occur in these constructions as well.[8] From the discussion which follows, I will conclude that V2 is part of the *free coordination construction*, whereas V-late is characteristic of the other three syntactic constructions, at least in a Germanic language like Dutch.

In addition to the two parameters discussed earlier – the presence of illocutionary force and the relation of scope – Verstraete mentions a third parameter. This parameter concerns the presence versus absence of a separate modal value in the connective clause. The presence of modal value is partially dependent on the presence of illocutionary force: presence of illocutionary force in a conjunct automatically implies the presence of modality, but absence of illocutionary force does not automatically imply absence of modality (Verstraete 2000: 12). For example, a connective clause headed by *since*, which – according to Verstraete – cannot have an illocutionary force of its own, allows modal verbs (such as *may* in (10)) other than the default simple indicative mood. The Dutch equivalents with *want* and *omdat* in (11) and (12) illustrate that both connectives can introduce clauses with a modal value of their own.

(10) The meetings of the Central Preparatory Commission of the Council <u>will</u> be more important than the sessions of the Council itself, **since** greater frankness and sincerity <u>may</u> be expected in them. (Verstraete 2000: 13)

[8] The choice of these labels does not imply that Verstraete accepts the traditional categorization associated with them (cf. Verstraete 2000: 17). They are only used as shorthand labels for the four construction types (but see Verstraete 2000: 43-48 for a comparison of his classification and the traditional categorization based on coordination and subordination).

(11) *De bijeenkomsten van X zullen belangrijker zijn dan de sessies van Y zelf,* ***want*** *je zou daarin een grotere openheid en oprechtheid mogen verwachten.*
(12) *De bijeenkomsten van X zullen belangrijker zijn dan de sessies van Y zelf,* ***omdat*** *je daarin een grotere openheid en oprechtheid zou mogen verwachten.*
'The meetings of X will be more important than the session of Y itself, because you may expect a greater frankness and sincerity in them.'

In the case of *want*, this modal value is present because of the presence of illocutionary force (i.e. it is a case of free coordination); in the case of *omdat*, the modal value is present despite the absence of such an illocutionary layer. In Verstraete's terms, the *omdat*-clause in (12) must be regarded as an instantiation of the *bound coordination construction*.[9]

The combination of the three parameters (presence versus absence of illocutionary layer, modal value and relation of scope) generates a fourth syntactic configuration: the *free subordination construction*. As Table 6.1 indicates, connective clauses within such constructions neither have an illocutionary force nor a modal value of their own, and they occur outside the scope of their matrix clause. The fact that *omdat*-clauses can be preposed (see (13)a) shows that they can appear in free subordination; *want*-clauses cannot be preposed and therefore cannot be regarded as instantiations of free subordination (see (13)b).[10]

(13) a. ***Omdat*** *Jan te lang buiten in de regen is gebleven, is hij ziek.*
b. ****Want*** *Jan is te lang buiten in de regen gebleven, is hij ziek.*
'Because John has been out in the rain too long, he is ill.'

The claims in this section can be summarized as follows. De Haan's root clauses showing V2 can be equated with Verstraete's adverbial clauses occurring in free coordination, since both clause types are the only ones with a separate illocutionary layer. De Haan's V-late non-root clauses, on the other hand, can appear in three different types: bound coordination (with a modal value of its own), free subordination (without a modal value, but outside the scope of the matrix clause), and bound subordination (without a modal value, and within the scope of its matrix clause). So, the proposed mapping between V2 and syntactic configuration is one-to-one, whereas the mapping between V-late and syntactic configuration is one-to-three. These claims are supported by Modern Dutch intuitions on the use of *want* and *omdat*.

6.2.2 Syntactic configurations and their domains of use

The second step in my line of reasoning is that different syntactic configurations not only relate to different word orders, but also to different interpretations based on domains of use or source of coherence. Haegeman's (2003) and Verstraete's (1998, 2000) theories contain such an interaction between the syntactic properties of adverbial clauses and their text-linguistic interpretation based on source of coherence.[11] In the current section I present some of their

[9] An additional argument for the claim that *omdat*-clauses may occur in bound coordination is that they sometimes have a separate intonation contour (see Verstraete 2000: 18 for remarks on the distinction between integrated and non-integrated intonation contours as a discriminative factor).
[10] See Verstraete (2000: 18) on the suitability of this test in distinguishing between bound and free subordination.
[11] Verstraete's (1998) ideas are highly similar to Haegeman's, despite the fact that Haegeman works within a generatively oriented framework and Verstraete within the framework of Functional Grammar. Table A presents an overview of their respective terms, as well as the terminology used in Sanders et al. (1992). Conceptual or syntactic terms that are mentioned in the same column refer to the

arguments to show that bound subordination constructions are restricted to semantic or content use, whereas the other constructions are not restricted in the text-linguistic interpretations they may receive.

In Verstraete's (2000) terminology, the semantic adverbial clause in (14) is a case of bound subordination. Its connective clause is part of, or applies to the instantiated type of the matrix clause (i.e. the state of affairs expressed by the lexical verb and its argument); this embedded connective clause does not have grounding resources (i.e. a modal and illocutionary value) of its own (nor a separate intonation contour, see Verstraete 1998: 190).[12] The *because*-clause of the pragmatic utterance in (15) applies to the grounding resources of the matrix clause (i.e. the speech event, its setting, and its participants); it functions as a separate clause with its own modal and illocutionary layers.[13] The connective clause in (15) can therefore be said to occur in a free coordination construction.

(14) John is ill **because** he has been out in the rain too long.
(15) John is ill, **because** I heard him coughing this morning.

A first argument for the claim that bound subordination constructions are restricted to a semantic interpretation, is that only examples like (14) can undergo so-called 'factive nominalization' without meaning change (Verstraete 1998: 198). Factive nominalization is a syntactic operation that removes the grounding resources of a clause. Applying this operation to the pragmatic example in (15) leads to a meaning change in (17). As (16) shows, the same operation is not problematic for an embedded clause like (14), in which a separate ground is already absent before applying the operation (cf. also Rutherford 1970: 98).

(16) John's being ill because he has been out in the rain too long [did not come as a surprise].
(17) #John's being ill because I heard him coughing this morning [did not come as a surprise].

A further argument for the text-linguistic difference between bound subordination constructions and the other construction types comes from the application of two other syntactic operations: clefting and *why* questioning. The semantic adverbial in (14) allows for these operations without a change in interpretation (see (18) and (19)). In contrast, these operations are not possible with pragmatic use without affecting the interpretation of the relation; they shift it to the semantic level (cf. Verstraete 1998: 194-195; see also Rutherford 1970: 105-107; Sanders et al. 1992: 9). Both clefting and *why* questioning are operations that "impose an embedded construction on the example to be tested" (Verstraete 1998: 194); they

---

same notion. For example, Haegeman's syntactic distinction between *central* and *peripheral* clauses can be equated with Verstraete's distinction between *embedded* and *non-embedded* clauses. Haegeman even uses the same terminology when she states that event adverbials are "more deeply embedded" (Haegeman 2001: 1).

Table A. Comparison of the conceptual and syntactic terminology

Author	**Conceptual terminology**		**Syntactic terminology**	
Sanders et al. 1992	semantic	pragmatic	embedded	not embedded
Haegeman 2003	event	discourse	central	peripheral
Verstraete 1998	external	internal	embedded	non-embedded

[12] See also Couper-Kuhlen 1996 on the relation between intonation and use in domains.
[13] Both examples and the syntactic variants in (16)-(23) are taken from Verstraete (1998: 183).

demand the connective clause to function as an adjunct within the instantiated type of the matrix clause. As embedded clauses, they are part of the instantiated type of the main clause, and therefore also fall within the scope of its grounding resources. This makes a pragmatic reading impossible (see (20) and (21))[14]: "a conjunctive relation cannot at the same time be within the scope of grounding and apply to this same grounding" (Verstraete 1998: 193). As (22) and (23) illustrate, the insertion of a clause like *I know* is needed to recover the original reading of the clefted and questioned alternatives.

(18) It is because John has been out in the rain that he is ill.
(19) Why is John ill? Because he has been in the rain too long.
(20) #It is because I heard John coughing this morning that he is ill.
(21) #Why is John ill? Because I heard him coughing this morning.
(22) It is because I heard John coughing this morning that I know he is ill.
(23) Why/how do you know John is ill? Because I heard him coughing this morning.

From the syntactic tests above, Haegeman infers a strict one-to-one mapping between source of coherence and syntactic structure: her *peripheral* clauses (i.e. free coordination constructions) are pragmatic per definition, whereas her *central* clauses (i.e. bound subordination constructions) are always semantic.[15] Verstraete explicitly rejects such a one-to-one mapping. He convincingly argues that although embedded clauses (i.e. bound subordination constructions) always receive a semantic interpretation, non-embedded clauses can be used to express both semantic and pragmatic relations. This implies that semantically used adverbial clauses can occur either in bound subordination constructions or in one of the other three construction types.

Haegeman's idea of a one-to-one mapping between source of coherence and syntactic structure predicts that adverbials should either receive a semantic or a pragmatic reading. This theory, with a strict division of labor, is incorrect, since it cannot account for adverbial clauses that are ambiguous in terms of source of coherence. Verstraete's idea of more than one syntactic possibility for semantic adverbials is more attractive in this respect, since it leaves open the possibility of ambiguity in the domains interpretation of adverbial clauses (as discussed by Sanders 1997).

What can be concluded about the relation between syntactic configuration and interpretation based on domains of use? Verstraete (2000) claims that the bound subordination type is restricted to the content domain, without being more specific about the domains interpretations of the other construction types. The syntactic tests presented in the current section reveal that the bound subordination construction is indeed restricted to semantic use, whereas the free coordination construction is not restricted in this respect. Since I do not know of other factors restricting the domain interpretations of the remaining two construction types (bound coordination and free subordination), I will assume that these construction types can be used to express all three types of domains. To conclude this section, it can be stated that there is a link between source of coherence and the syntactic properties of the connective

[14] Note again that the hash sign does not necessarily signal ungrammaticality or incoherence; it merely indicates that the intended pragmatic meaning is not preserved in the "hashed" connective clause.

[15] Strictly speaking, this division of labor does not even follow from Haegeman's line of reasoning: her arguments and syntactic tests (see Haegeman 2003: 320-324) support her claim that adverbial clauses in central position cannot receive a pragmatic reading, but this does not imply that the reverse holds as well. In other words, her articles do *not* prove her claim that adverbial clauses in peripheral position can only receive a pragmatic reading and no semantic reading.

clause, but that this link is not one-to-one. Bound subordination implies semantic use (i.e. use in the content domain); the other construction types allow for both semantic and pragmatic interpretations (i.e. use in the content, epistemic and speech-act domain).

6.2.3 Domains of use and their word orders

In section 6.2.1 I put forward the claim that word order interacts with syntactic configuration. In section 6.2.2 I presented arguments for the claim that syntactic configurations differ in the allowed interpretations based on domains of use. Combining these claims from the two previous sections results in the claim that word order interacts with domains of use. An overview of the interactions between syntactic structure, word order and interpretation based on domains of use is presented in Table 6.2.[16]

Table 6.2. Interactions according to a combination of Verstraete (2000) and De Haan (2001)

Construction type	Word order	Interpretation based on domains
Free coordination	V2	content, epistemic, speech act
Bound coordination	V-late	content, epistemic, speech act
Free subordination	V-late	content, epistemic, speech act
Bound subordination	V-late	content

How do these claims relate to Dutch *want* and *omdat* and their respective word orders? The causal connective *want* exhibits the V2 word order. Therefore, De Haan would analyze *want*-clauses as root clauses with an illocutionary force of their own, 'coordinated' (paratactically) to their matrix clause. In Verstraete's terms, *want*-clauses occur in free coordination, which implies that they can receive all three domain interpretations. V-late *omdat*-clauses, on the other hand, are regarded as non-root clauses without an illocutionary force of their own; in Verstraete's terminology *omdat*-clauses are excluded from the free coordination construction. *Omdat*-clauses can occur in bound subordination constructions, which are restricted to the content domain, but also in free subordination or bound coordination constructions with other domain interpretations. This division of labor on the basis of the word order properties of *want* and *omdat* is in line with the Modern Dutch intuitions on the use of these connectives presented in section 6.2.1.

From Table 6.2 it might be concluded that the word order characteristics of a connective completely determine the construction type in which this connective can or cannot occur. This is indeed true for V2 connectives, which always occur in free coordination constructions. However, a connective's V-late nature only implies that this connective cannot occur in free coordination. It is a connective-specific characteristic of each V-late connective in which of the other three construction types it can or cannot occur. Section 6.2.1 showed that the causal connective *omdat* can occur in all three construction types. Pit (2003) convincingly argues that another Dutch causal connective, *aangezien* 'since', cannot function within the scope of its matrix clause, despite the fact that it does exhibit V-late. In Verstraete's terms: *aangezien*

[16] Note that this table still cannot account for two important differences between *want* and *omdat*. The connective *want* can be used not only to mark content relations, but also for speech act relations and both deductive and abductive epistemic relations. The connective *omdat* can only be used to mark content relations and deductive epistemic relations (cf. Degand 1996, 2001). Further research is needed to investigate whether these differences ought to be related to differences in the internal and external syntax of the connective clauses or not.

cannot occur in the bound subordination construction, although it can occur in free subordination and bound coordination.

The combined theory presented in Table 6.2 can also be used to account for the facts on German *weil* (presented in section 3.5.1). It appears that German *weil* in its V-late version could only occur in the bound subordination construction, which explains for its restriction to content interpretations. In spoken German, however, *weil* has adopted the function as well as the word order of *denn*, which nowadays is regarded as belonging to the literary language. In Verstraete's terminology: V2 *weil* can occur in free coordination, which results in an expansion of its meanings to the epistemic and speech-act domains.

It can be concluded that the proposed combination of the ideas in Verstraete (1998, 2000), Haegeman (2003) and De Haan (2001) seems promising. In the remainder of this chapter, I will put this theory about the interaction between word order and domains to the test on the basis of a diachronic corpus study.

6.2.4 Hypotheses and methodology for the diachronic study

The previous section introduced a proposal concerning the interaction between syntactic properties (including word order) and text-linguistic properties of adverbial clauses. This proposal was applied to the Dutch connectives *want* and *omdat*, and partially tested using synchronic data on the use of these connectives. In the current section this proposal will be used to formulate hypotheses about the diachronic development of *want* and *omdat*. These hypotheses build on the syntactic observations in the previous chapter (see section 5.4.1) that the Modern Dutch coordinator *want* in earlier periods could also be combined with V-late.

A first hypothesis concerns the interaction between V-late and use in the content domain. Given that the bound subordination construction can only combine with V-late and that it can only receive a content interpretation, my hypothesis is that the V-late word order more often expresses a content relation than an epistemic or a speech-act relation.

The second hypothesis concerns the use of *omdat* and is related to the first hypothesis. The V-late connective *omdat* differs from *want* (and *aangezien*) in that it can occur in the bound subordination construction, which forces a content interpretation. My hypothesis is that the overall profile of *omdat* is more content-oriented than that of *want*. Only in that case does *omdat* exploit its abilities, and guarantee its right to exist.

A third, more provisional hypothesis concerns the use of *want*. If word order is indeed related to construction type, Middle Dutch *want* should be able to occur in constructions other than free coordination. The syntactic analyses in the previous chapter already revealed that this V-late connective could occur in preposed clauses, or in Verstraete's free subordination constructions. If Middle Dutch V-late *want* could also occur in bound subordination constructions, this could also imply that Middle Dutch V-late *want* was more content-oriented than Middle Dutch V2 *want*. However, Middle Dutch language users did not have to select V-late *want* more often if they wanted to mark a content-relation, because they could also opt for V-late *omdat*. In this view, the V-late nature of Middle Dutch *want* can be seen as a sufficient condition for more often expressing content-relations, but – given the presence of another V-late alternative – not as a necessary condition. This third hypothesis will therefore be put forward as a weak hypothesis. An overview of the hypotheses is given in (24).

(24) Hypotheses for the diachronic research in this chapter:

a. The V-late word order has a preference for use in the content domain.
b. The overall profile of V-late *omdat* is more content-oriented than that of *want*.
c. Middle Dutch V-late *want* is more content-oriented than Middle Dutch V2 *want*.

The same samples of connective fragments from the previous chapter (see section 5.3) were used for the diachronic study in which these hypotheses were put to the test. Per connective, each of the 150 fragments has been subject to a conceptual analysis based on domains (see section 4.3 in Chapter 4 for an operationalization). The diachronic results concerning the syntactic properties of the *want*- and *omdat*-clauses mentioned in the previous chapter were used to investigate the interaction with conceptual developments based on domains of use.

6.3 Results of the diachronic analysis of *want*

The current section presents the results of the corpus study on *want*. In section 6.3.1 the conceptual analyses will be discussed, which will then (section 6.3.2) be related to the word order properties discussed in the previous chapter. Statistical details can be found in Appendix C (see Appendices to Chapter 5).

6.3.1 Conceptual analysis of *want*

What does the conceptual picture of *want* look like and how does it relate to the syntactic properties mentioned in the previous chapter? As early as the 13th century – the earliest period under investigation – *want* is mainly used as a connective signaling causal coherence relations. The only non-causal use in my corpus appears in the 16th century, where *want* appears as a temporal connective, meaning 'whenever', 'until'.[17] In the 20th century this temporal use, exemplified by (25) and (26), has disappeared. These two temporal *want*-fragments are left out in the detailed conceptual analyses based on domains.

(25) *Haer docht dat sij die salicheit harer zielen vant,* ***want*** *als sij heymelicke steden conste vinden die tot beden bequaem waren* (Leven van Sinte Clara, 1500-1520)
'She thought she would obtain salvation of her soul, when she could find a secret place that was appropriate to pray.'

(26) *Cleyn was v siluer v gout*
Want *ghi den cost ghinct halen.* (Devoot ende profitelyck boecxken, 1539)
'Little was your silver, your gold, when you went to get your food.'

During the three periods under investigation, *want* can occur both in the content domain (see (27)) and in the epistemic and speech-act domain (examples (28) and (29) respectively).

(27) *(…) nochtans had hi* (= Hughe) *die meeste sorghe voor Claramonde sijn lief niet wetende waer die veruaren was* [$_{S1}$ *so badt hi god den here dat hi se doch wilde behoeden vander doet*] [$_{S2}$ ***want*** *hi gad grote sorghe dat die sarasinen haer souden brengen tot haren oom coninc yuorijn ende hi dacht, comtse in haer ooms handen so salse ymmer moeten steruen.*] (Historie van Hughe van Bordeus, 1530-1550)
'(…) still he (= Hugh) worried most about his love Claramonde. Not knowing where she was, he prayed to God the Lord that He would keep her from death, because he had tremendous worries that the Arabs would bring her to her uncle, king Yorin, and he thought, "If she falls into the hands of her uncle she will surely have to die." '

[17] According to the examples in the *MNW* this temporal use already existed in the 13th century. Other meanings of *want* mentioned in the *MNW*, such as the expression of purpose ('so that'), are not attested in my corpus.

(28) *Er moet hem een behoorlijk bedrag nagelaten zijn,* ***want*** *zijn ouders* <u>*waren*</u> *redelijk welgesteld.* (In de schaduw bloeien de rozen, 1994)
'A substantial amount must have been left to him, because his parents were fairly well-off.'

(29) *Spiet seyde. Ic wil gaen versoecken dyen reuse met mijnder metalen colve. Die ridder seyde. Laet staen die woorden* ***want*** *hy doode uwer .xxv. wel (…)*
(Historie van Malegijs, 1556)
'Spiet said, "I want to try and kill that giant with my metal club." The knight said, "Stop with those words because he already killed your 25 [men]" (…)'

Figure 6.1 makes clear that this causal connective has a preference for the epistemic domain. Statistical analysis (see Appendix C) reveals that the distribution over the domains is not constant through the ages (χ^2 (4) = 13.9; p < .01). A separate, more detailed domains analysis of *want* (see Chapter 8 and Evers-Vermeul & Stukker 2003) shows that this interaction effect is restricted to the speech-act domain (χ^2 (1) = 12.2; p < .001): during the 16th century *want* is used to motivate speech acts with more frequency than in either of the other periods.[18]

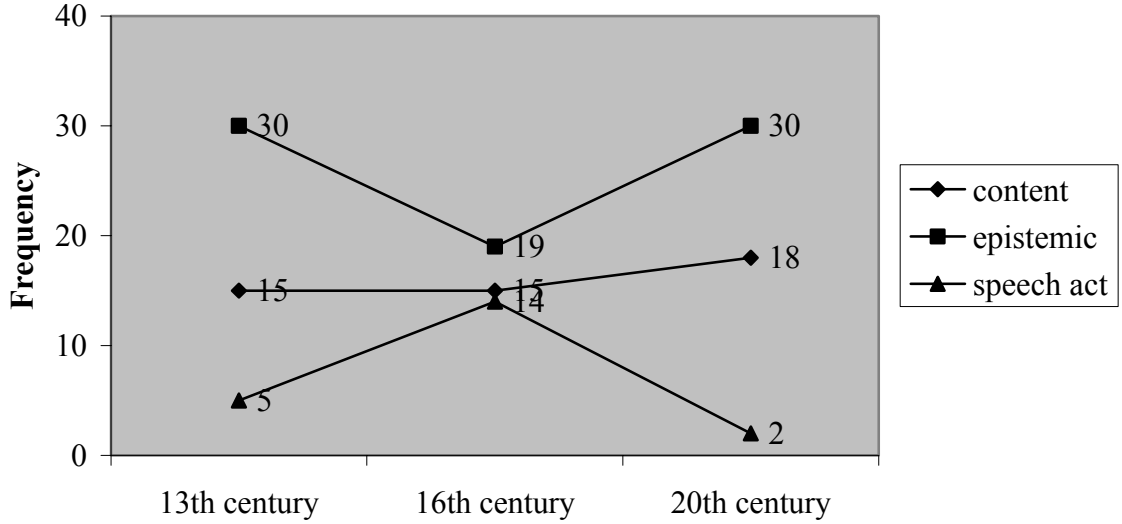


Figure 6.1. *Want*: Domain frequencies in three periods

At the conceptual level, it can be concluded that apart from the disappearance of the temporal use of *want* after 16th century, there are no lasting changes in its use as causal connective.

6.3.2 Interaction between conceptual and syntactic properties of *want*

In this section, I focus on the interaction between word order and interpretation based on domains. Because of the low number of preposed *want*-fragments, a statistical testing of the interaction between the factors 'linearization' and 'domains' would be unreliable.

From the syntactic data in the previous chapter (see section 5.4.1) it can be concluded that medieval *want* seems to hover between coordinator and subordinator, and that its clauses

[18] The domains analysis mentioned in Evers-Vermeul & Stukker (2003) takes into account the difference between non-volitional and volitional use of *want*. This subdivision within the content domain has not been applied in the current study. It was not necessary from a theoretical point of view and the very low number of non-volitional *want*-fragments would have made a statistical analysis of the interaction between domains of use and word order unreliable.

show word order and linearization characteristics of both of them. Table 6.3 shows the distribution of the word order patterns over the different domains of use.

Table 6.3. *Want*: Word order versus domains (all periods taken together)

	V2	**V-late**	**V-ambiguous**	**Total**
Content	23	11	13	47
Epistemic	57	3	19	79
Speech act	12	1	8	21
Total	92	15	40	147

Statistical analysis reveals that there is an interaction between word order and domains (χ^2 (4) = 16.2; $p < .005$): V-late appears less often in the epistemic domain ($z = -3.32$; $p < .001$) and in the speech-act domain ($z = -2.13$; $p = .03$). This implies that the subordinating word order is mainly used to express content relations, a picture that becomes even more vivid if the two fragments (disregarded in the statistical analysis) in which *want* means *when* are taken into account. These *when*-fragments can also be classified as V-late relations in the content domain. This interaction is in agreement with my first hypothesis that V-late should have a preference for content use. The fact that only the V-late fragments show this preference for content use in combination with the fact that V-late *want* only appears in the 13th and the 16th centuries, leads to the conclusion that the third hypothesis is also on the right track. Middle Dutch V-late *want* was indeed more content-oriented than Middle Dutch V2 *want*. This preference also becomes apparent from Table 6.4, which only shows the medieval fragments (disregarding the ambiguous word orders).

Table 6.4. Medieval *want*: Word order versus domains (ambiguous word orders disregarded)

	V2	**V-late**	**Total**
Content	8	11	19
Epistemic	32	3	35
Speech act	10	1	11
Total	50	15	65

Statistical analysis of the data in Table 6.3 also reveals that the interaction between word order and domains is stable across the three periods under investigation. This is remarkable: the disappearance of V-late has no equivalent change at the conceptual level (see also Figure 6.1); for example, the disappearance of subordinating *want* does not lead to an increase in epistemic use or a decrease in content use.

6.4 Results of the diachronic analysis of *omdat*

This section provides the diachronic analyses of *omdat* based on its conceptual properties (6.4.1) and the interaction between its syntactic and conceptual properties (6.4.2). Statistical details can be found in Appendix D (see Appendices to Chapter 5).

6.4.1 Conceptual analysis of *omdat*

In the 13th century, *omdat* has already made its appearance as causal connective, which can be used in all three domains. Example (30) illustrates the content use of *omdat*; the examples in (31) and (32) exemplify its epistemic and speech-act use respectively.

(30) *ende* ***om dat*** *haer lof ghemeert / daer af worde so seid sijt voerd /*
dat si van har hoerde dees woerd. (Sinte Lutgard, 1275-1300)
'And in order that her praise would be increased because of that, she spread the word that she heard from her.'

(31) *We zouden gezellig met z'n vieren uitgaan om het ijs te breken. Jij hebt daar kennelijk geen zin in. Nou,* [$_{S1}$ *dan is er hier kennelijk één te veel,*] [$_{S2}$ ***omdat*** *jij de sfeer al een tijdlang zit te verknoeien.*] (De gouden handjes, 1993)
'The four of us were going to go out enjoyably to break the ice. Obviously, you don't feel like that. Well, apparently there is one person too many here, because you have been ruining the atmosphere for some time now.'

(32) *Doe ombod hem aldus god.*
[$_{S1}$ ***Om dat****tu (= Jerobiam) daets mijn ghebod.*]
[$_{S2}$ *Dijn gheslachte*] *weet dat wel.*
[$_{S2\text{-continued}}$ *Sullen sijn coninghe in ysrahel.*
Al tote dat comt ten vierden lede.] (Rijmbijbel, 1275-1300)
'Then God said to him:
Because you did my command
Your family – know that well – will be kings of Israel until the fourth line.'

Especially the speech-act use of *omdat* is remarkable given the observations in the literature that Modern Dutch *omdat* cannot mark speech-act relations at all. My corpus even contains a 20th-century fragment that marks a speech-act relation (see (33)).

(33) *Peggy Ann,* [$_{S1}$ ***omdat*** *het nog vroeg is en ik het niet leuk vind de hele avond hier te blijven hangen,*] [$_{S2}$ *mag jij het zeggen.*] (De gouden handjes, 1993)
'Peggy Ann, since it is still early and I don't like the idea of staying here all evening, you decide.'

Figure 6.2 introduces the distribution of *omdat* over the three domains. Statistical analysis of the data in this table shows that *omdat* has a preference for content-relations (χ^2 (2) = 115.7; p < .001). This domains profile is very constant over the years.

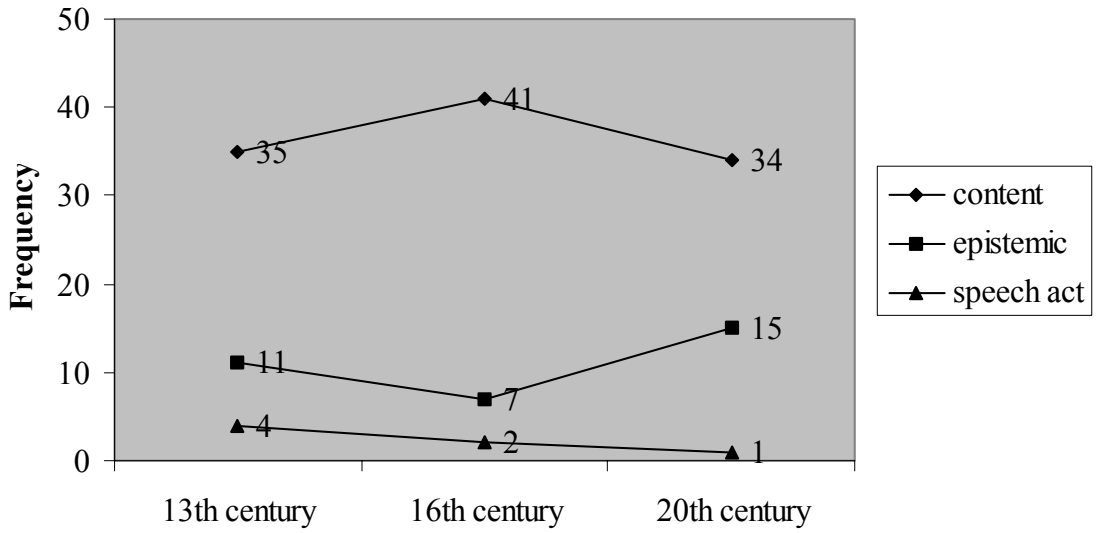


Figure 6.2. *Omdat*: Domain frequencies in three periods

A remarkable characteristic of *omdat* in the 13th and 16th centuries is that it could also be used for relations in which speakers of modern Dutch would use the connective *opdat* 'in order to' (even though the connective *opdat* already existed during these periods). In sum, there are twenty occurrences (11 in the 13th century and 9 in the 16th century) in which *omdat* marks such a finalistic causal relation between an antecedent and a desired, but not yet realized consequence.[19] An example is given in (34).

(34) *Doe ginc Lodewijc tegen Adelairt over sitten ende men brocht dair een scaeck bort dat costelic was ende subtijl van werc: men brocht dat constelic spel* ***om dat*** *overmits de subtijlheit vant werck Adelaert sijn spel versien soude.*
(Historie van den vier heemskinderen, 1508)
'Then Lodewijk sat across from Adelaart and someone brought them a chessboard that was expensive and well-finished. Someone brought the skillful set so that Adelaart would overlook his play because of the subtlety of the work.'

Although the specific type of causality is different in finalistic fragments like (34), *omdat* still marks a causal relation. Therefore, these finalistic fragments were included in the analyses. To be sure that the finalistic use does not disturb the domains picture of 'normal' *omdat*, I performed a more detailed analysis of the *omdat*-fragments from the 13th and 16th centuries.[20] Table 6.5 shows the distribution patterns of the two uses.

Table 6.5. Medieval *omdat*: Domain comparison of 'normal' and finalistic use

	'Normal' use	**Finalistic use**	**Total**
Content	62	14	76
Epistemic	16	2	18
Speech act	2	4	6
Total	80	20	100

This table reveals that the domains-profile of *omdat* in its finalistic use differs from the profile in its 'normal' use. Compared to the 'normal' use of *omdat*, the finalistic use shows a relatively high number of speech act relations (2 out of 80 = 2.5% vs. 4 out of 20 = 20%).[21] An example of this speech-act use is given in (35).

(35) *Ic heb hem hooren seggen dat u vader hem versende in Palerne,*
om dat*men hem dooden soude* (Historie van Malegijs, 1556)
'I heard him say that your father sent him to Palerne, so that they could kill him.'

This finding that finalistic *omdat* shows a relatively higher number of speech-act relations might be considered a reason to exclude the finalistic fragments from the statistical analysis. Disregarding the finalistic fragments does *not* affect the variables within the model in a

[19] This finalistic use has already been noticed by a.o. Bouman (1918: 115), De Rooij (1982: 340) and Heersche (1991: 73).
[20] The data from the 20th century were excluded from this analysis, since *omdat*-fragments from the 20th century cannot receive a finalistic interpretation.
[21] See also Chapter 8 for yet another difference between the two: within the content domain the finalistic use is restricted to volitional relations, whereas 'normal' *omdat* can mark both volitional and non-volitional relations.

significant way (compare Table III in Appendix D with Table III in Appendix E). However, disregarding the finalistic fragments *does* result in one significant change in the parameter estimates of the logit model of *omdat* (compare the final line of Table IV in Appendix D with the one in Appendix E): the finalistic *omdat*-fragments seem to blur the picture of the interaction between linearization and domains of use. It shows that normal *omdat*-fragments have a preference for speech acts with the *omdat*-clause in preposed position ($z = 2.07$; $p = .04$), a picture that could not be obtained from the analysis that included the finalistic fragments.

Through the ages, then, *omdat* is remarkably constant in its preference for content relations. Over the years, *omdat* has lost its ability to express finalistic causal relations (which nowadays are marked with *opdat* or *zodat* 'in order to/so that').

6.4.2 Interaction between conceptual and syntactic properties of *omdat*

The previous chapter (see section 5.5.1) revealed that *omdat* has an overwhelming preference for V-late in all three periods. Is there an interaction between these syntactic properties and the conceptual characteristics of *omdat* mentioned above? Table 6.6 shows the distribution of the domains over the different word order patterns.

Table 6.6. *Omdat*: Word order versus domains (all periods taken together)

	V2	**V-late**	**V-ambiguous**	**Total**
Content	0	93	17	110
Epistemic	1	25	7	33
Speech act	0	5	2	7
Total	1	123	26	150

The distribution in Table 6.6 does not result in any significant interactions. *Omdat*'s overwhelming preference for content relations (see section 6.4.1) applies both to its V-late fragments and its ambiguous fragments. Given the observation that *want* favors the epistemic domain, this finding supports my second hypothesis that *omdat* should be more content-oriented than *want*.

A significant interaction can however be found between linearization and domains (χ^2 (2) = 8.7; $p < .025$). As the distribution in Table 6.7 indicates, *omdat*-clauses that motivate speech acts more often occur in preposed position than *omdat*-clauses marking content or epistemic relations ($z = 2.03$; $p = .04$).[22]

Table 6.7. *Omdat*: Linearization versus domains (all periods taken together)

	Preposed	**Postposed**	**Total**
Content	25	85	110
Epistemic	4	29	33
Speech act	4	3	7
Total	33	117	150

From a modern Dutch perspective, it is remarkable that speech-act *omdat*-fragments occur in a postposed position at all, since this combination is ungrammatical in modern Dutch. The occurrence of these speech-act *omdat*-clauses in postposed position can be explained on the

[22] In this interaction analysis, the intra- and postposed use of *omdat* have been grouped together.

basis of the finalistic use of *omdat*: all three postposed fragments concern finalistic causal relations like (36).

(36) *Hoort, wat ic u noch leeren sal, mijn schoone minne,*
Om dat *ghi die nighermancie sout laten varen.* (Mariken van Nieuwmeghen, 1515)
'Hear what more I have to teach you, my pretty love, so that you will leave the necromancy.'

Looking at the interaction between conceptual and syntactic properties, it can be concluded that V-late *omdat* has an overwhelming preference for use in the content domain. This supports my second hypothesis. A further interaction can be observed between linearization and domains: speech-act relations relatively often have the *omdat*-clause in a preposed position.

6.5 Conclusion and discussion

In conclusion I can say that there is clear evidence that early on in the Middle Dutch period, the present-day complementizers *want* and *omdat* made their appearance as causal connectives. The general development of *want* can be depicted as in Table 6.8.

Table 6.8. *Want*: General profile in three periods

	13th century	**16th century**	**20th century**
Coherence relation	temporal causal	temporal causal	causal
Preferred domain	epistemic domain	epistemic domain	epistemic domain
Word order	V2 and V-late	V2 and V-late	V2

The general development of *omdat* is shown in Table 6.9.

Table 6.9. *Omdat*: General profile in three periods

	13th century	**16th century**	**20th century**
Coherence relation	causal finalistic use	causal finalistic use	causal no finalistic use
Preferred domain	content domain	content domain	content domain
Word order	V-late	V-late	V-late

In the remainder of this section, three topics will be discussed, the interaction between word order and domains (6.5.1), the disappearance of V-late *want* (6.5.2), and the speech-act use of *omdat* (6.5.3).

6.5.1 On the interaction between word order and domains

How do the conceptual and syntactic changes of *want* and *omdat* relate to the hypotheses in (37) (repeated from section 6.2.4)?

(37) <u>Hypotheses for the diachronic research in this chapter:</u>
a. The V-late word order has a preference for use in the content domain.
b. The overall profile of V-late *omdat* is more content-oriented than that of *want*.
c. Middle Dutch V-late *want* is more content-oriented than Middle Dutch V2 *want*.

The hypothesis in (37)a concerns the interaction between word order and domains of use. In section 6.2.3 I claimed that the V-late word order is the only one that can be used in the bound subordination construction, in which only a content interpretation is possible. This hypothesis is borne out: both the V-late fragments of *want* and the V-late fragments of *omdat* have a preference for use in the content domain.

The findings on *want* and *omdat* are in line with the hypothesis in (37)b: *want* has a preference for use in the epistemic domain, whereas *omdat* has a preference for the content domain. Hence, it can be concluded that *omdat* indeed "exploits" its syntactic abilities, guaranteeing its right to exist.

The hypothesis in (37)c, which concerns the interaction between domains and word order in the use of *want*, is also confirmed. The medieval fragments with the V-late word order indeed show a preference for the content domain, a preference that is not shared by the V2 fragments from the same period. The disappearance of the V-late word order of *want* does not result in a decrease in content use. Apparently, the syntactic possibilities of *want* do not influence its overall conceptual profile, which is remarkably stable. This does not imply that the proposed mapping between V2 and free coordination is wrong. As noted before, Middle Dutch language users did not have to select V-late *want* more often if they wanted to mark a content relation, because they could also opt for V-late *omdat*. In this view, the V-late nature of Middle Dutch *want* can be seen as a sufficient condition for more often expressing content relations, but, given the presence of a V-late alternative with a preference for content relations, not as a necessary condition.

Given these results, what can be concluded on the interaction between word order, syntactic configuration and domains of use? It appears that the mapping between V2 and free coordination on the one hand, and V-late and the other three construction types on the other is on the right track. These combinations result in a partial interaction between word order and domains: only V-late connectives in the bound subordination construction are restricted to the content domain; V2 clauses in free coordination as well as V-late clauses in free subordination and bound coordination are not restricted in this respect. Further domain restrictions in the use of connectives seem to be triggered by connective-specific properties. Before I address that topic (see section 6.5.3), I will discussion the disappearance of V-late *want*.

6.5.2 On the disappearance of V-late *want*

The results for the diachronic development of *want* revealed that this connective has lost its ability to combine with a V-late word order (as well as the ability to occur in preposed position). This diachronic change can be seen as a case of *structural scope expansion* in the sense of Tabor & Traugott (1998): the connective *want* has disposed of its occurrence in embedded CPs in which the *want*-clause functions within the scope of the matrix clause. Nowadays, this connective only occurs outside the scope of the matrix clause, or in De Haan' terms, it occurs as a root clause with a functional E-projection of its own.

The question arises why *want* has changed. Although I do not yet have any concrete evidence for this claim, I would like to suggest that the connectives *want* and *omdat* originally stem from different dialects and that they became competitors. From this competition perspective it is not remarkable that the V-late use of *want* has disappeared. Given that the V-late occurrences of *want* were mainly used to express content relations, V-late *want* can indeed be regarded as a direct competitor of the connective *omdat*. It is likely that Dutch language users came to prefer this latter connective to express V-late content relations, since *omdat* typically occurs in V-late constructions with a content interpretation. In the 20th

century, the disappearance of the V-late use of *want* has resulted in a nice division of labor between *want* and *omdat*: V2 *want* is mainly used to express epistemic relations, whereas V-late *omdat* typically occurs in the content domain.

6.5.3 On speech act *omdat*

From a modern Dutch perspective, it is remarkable that speech-act *omdat*-fragments are attested to in all three periods under investigation, since relevant literature (e.g. Degand 2001; Pit 2003) claims that Modern Dutch *omdat* cannot occur in speech-act relations whatsoever. This literature only provides evidence in which postposed *want*-clauses with a speech-act interpretation are compared to postposed *omdat*-clauses. The data in my corpus confirm this analysis: as I mentioned earlier, the occurrence of speech-act *omdat*-clauses in postposed position can be explained on the basis of the finalistic use of *omdat*: all three postposed speech-act fragments concern finalistic causal relations.

However, the preposed *omdat*-clauses expressing a speech-act relation are not restricted to the finalistic use of *omdat*. Even in the 20th century there is one preposed *omdat*-fragment with a speech-act interpretation. It can be concluded, then, that the Dutch connectives *want* and *omdat* are in complementary distribution: V2 *want*-clauses mark speech-act relations in which the antecedent follows the consequent, whereas V-late *omdat*-clauses mark speech-act relations in which the antecedent precedes the consequent. Such preposed *omdat*-clauses seem to create the grounds on which the speech act should be interpreted (compare the remarks on the function of preposed clauses in section 3.3). This grounding function can only be fulfilled by a V-late connective, since coordinators cannot occur in preposed clauses for grammatical reasons.

It can be concluded from my analyses that the expression of causal speech-act relations is not restricted to connectives with V2 word order. V-late *omdat*-clauses in preposed position can be used to express speech-act relations (although this is not a frequent phenomenon). In addition, V-late *omdat*-clauses expressing a finalistic relation can be used to express speech-act relations, both in preposition and in postposition. It is important to notice that the difference between 'normal' and finalistic *omdat* in postposition cannot be attributed to the construction types in which they can or cannot occur, since both can be used in all construction types except free coordination. It appears that in addition to Verstraete's constructions and their respective word orders, there are connective-specific properties that impose extra restrictions on the precise domain possibilities. Such extra connective-specific restrictions are also needed to account for the difference between *want* and *omdat* in the area of epistemic relations (cf. also footnote 15). In this respect the mapping between free coordination and V2 on the one hand, and V-late and the other three construction types on the other hand cannot fully account for the conceptual properties of the connectives at hand. Further research is needed, in order to investigate whether the remaining conceptual differences should be related to differences in the internal and external syntax of the connective clauses, or that they should indeed be classified as connective-specific restrictions.

APPENDIX TO CHAPTER 6

Appendix E – Revised logit analysis of the diachronic development of *omdat*

I – Remarks

- Different numbers of words were needed to select 50 fragments per period: 83400 words for the 13th century, 95505 words for the 16th century, and 61574 words for the 20th century.
- This revised version of the analyses in Appendix D excludes the finalistic use of *omdat*.

II – Data: Distribution of *omdat* in three periods

		Content		Epistemic		Speech act		Total
		Pre	Post	Pre	Post	Pre	Post	
13th century	V-late	8	14	2	2	0	0	26
	V-ambiguous	0	6	1	4	2	0	13
16th century	V-late	6	24	0	5	0	0	35
	V-ambiguous	2	2	1	1	0	0	6
20th century	V-late	6	23	0	15	1	0	45
	V-ambiguous	0	5	0	0	0	0	5
Total		22	74	4	27	3	0	130

III – Results logit analysis *omdat*

Logit (fr)	χ^2 model	df	p model	χ^2 factor	df	p factor
1. constant	267.39	35	< .001	-	-	-
+ 2. period	256.69	33	< .001	10.70	2	< .001
+ 3. linearization	215.30	32	< .001	41.40	1	< .001
+ 4. word order	160.53	31	< .001	54.77	1	< .001
+ 5. domain	49.59	29	< .025	110.90	2	< .001
+ 6. period x word order	42.00	27	< .05	7.61	2	< .025
+ 7. linearization x domain	33.33	25	< .25	8.65	2	< .025

IV – Parameter estimates *omdat* for model 7

Parameter	Estimate	s.e.	z-score	p
constant	-8.64	0.21	-41.45	< .001
period: 16th century	0.16	0.26	0.63	0.53
period: 20th century	0.85	0.25	3.46	< .001
linearization: preposed	-1.20	0.24	-4.98	< .001
word order: V-ambiguous	-0.69	0.34	-2.03	0.04
domain: epistemic	-1.01	0.22	-4.47	< .001
domain: speech act	-4.82	1.29	-3.72	< .001
period x word order: 16th V-ambiguous	-1.05	0.55	-1.91	0.06
period x word order: 20th V-ambiguous	-1.42	0.56	-2.52	0.01
linearization x domain: preposed epistemic	-0.64	0.57	-1.11	0.27
linearization x domain: preposed speech act	2.94	1.42	2.07	0.04

Chapter 7

Diachronic developments of dus *and* daarom

The two previous chapters treated the diachronic development of two complementizers, which exhibit a fixed position at the beginning of the connective clause. The current chapter investigates the diachrony of two connectives mainly used as sentence adverbials: *dus* and *daarom*. It is characteristic of such adverbs that they can be inserted into various syntactic slots, which makes them interesting objects in the search for interactions between positioning and conceptual properties of these connectives.

> *"As has long been noted, the position of an adverb is correlated with difference of meaning (…)."* (Traugott & Dasher 2002: 158)

7.1 Introduction[1]

This chapter focuses on two Dutch adverbial connectives: *dus* 'so' and *daarom* 'that's why'. A characteristic of sentential adverbs is their ability to be inserted into various syntactic slots, as is illustrated for *dus* in (1) and (2).

(1) *Ik heb het altijd naar mijn zin gehad,* ***dus*** *wil ik graag opnieuw beginnen.* (MC, 1995)
'I was always content, so I would gladly start again.'

(2) *Het ging erom wie het mooiste, het beste produkt maakte. Dat was ik* ***dus****.* (MC, 1995)
'The point was who could make the most beautiful, the best product. And that was me.'

At first sight, the different positioning possibilities might seem arbitrary. However, Traugott & Dasher (2002: 158) claim they are not: "As has long been noted, the position of an adverb is correlated with difference of meaning (e.g. Greenbaum 1969, Jackendoff 1972, McConnell-Ginet 1982, Ernst 1984, Quirk et al. 1985, Cinque 1999, to mention only a few)." The hypothesis underlying their claim is attractive: a multifunctional word exploits its syntactic possibilities in order to differentiate between its semantic aspects.

For the two Dutch adverbials treated here, it is not clear yet what this relation between positioning and meaning looks like. Given the developmental approach in this thesis I will investigate their form-function interactions from a diachronic perspective. In this chapter, the following three questions will be answered for both *dus* and *daarom*.

(3) <u>Research questions of this chapter:</u>
a. Did any conceptual changes occur during the selected time span?
b. Did any syntactic changes occur during the selected time span?
c. Can these syntactic changes be related to the conceptual changes?

[1] I would like to thank Ninke Stukker for making her diachronic corpus of *dus*- and *daarom*-fragments available to me. See Stukker (2005) for additional conceptual analyses of *dus* and *daarom*.

To be more precise, I am looking into the relation between the positioning of *dus* and *daarom* and the different (connective and non-connective) meanings they can express. Furthermore, I investigate the possibility of an interaction between positioning of the connective and its use in different domains. The theoretical background on position-meaning interactions will be discussed in section 7.2, in which I also formulate hypotheses for the current research. In section 7.3 the sample and methodology of this study are accounted for, after which the results from the corpus study are presented (see 7.4 on *dus* and 7.5 on *daarom*) and discussed (7.6). The underlying methodological assumptions in this chapter are that "a detailed semantic analysis presupposes a detailed syntactic analysis and that syntax reveals much more about semantics than one generally expects. In other words, the meaning of an expression arises from its position in different networks of syntactic configurations" (Smessaert & Beeken 1995: 376).

7.2 Positioning related to function

The analyses of *jedenfalls* 'anyway' and *toch* 'still/after all' in Chapter 3 (see section 3.2) illustrate a distributional pattern in which positioning singles out the connective function from other (in particular: discourse marker) text-linguistic functions of these words. A similar claim is found in Ariel's (1988, 1999) work. Ariel focuses on words that may function both as connectives and as a specific type of discourse marker, namely as a marker signaling that the presented information is already accessible to the reader or hearer. According to Ariel (1988: 570-571), syntactic factors as well as intonation distinguish the discourse marker use from other uses these multifunctional words may have.

Ariel's claim about the interaction between positioning of a word and its respective meanings seems promising. Because she (1988: 570) claims that accessibility markers can be found in a variety of languages (mentioning examples from languages like German, Russian, Swedish, Polish, Hungarian, and Armenian), it seems attractive to investigate whether the Dutch language also shows such a syntactic division of labor between the connective and discourse marker function of words. In this chapter, the interaction between function and positioning is investigated for the Dutch adverbials *dus* and *daarom*.

How can the findings in the literature be translated into hypotheses about the interaction between positioning and function in the use of the Dutch causals *dus* and *daarom*? To start with, these words may show a very general interaction. To formulate this general hypothesis in a diachronic way: if *dus* and *daarom* express more than one function during a certain time span, these functions have different positioning preferences. A more specific variant of this hypothesis is that once *dus* and *daarom* come to serve as connectives, different positionings are related to usage in different domains.

For one of the two connectives, *dus*, it seems attractive to investigate whether an interaction with accessibility can be found.[2] The word *dus* is often regarded as the Dutch equivalent of English *so*. As Schiffrin (1987) has observed, *so* can be used to mark a causal coherence relation, but also as an instruction to the reader or hearer "to recover a conclusion (…) which has already been presented, or which is otherwise mutually known because of e.g. just-presented reasons and/or support" (p. 223). An indication that a similar analysis is possible for Dutch *dus* as well, is the often-heard intuition mentioned in Pander Maat & Sanders (1996: 205) that "*dus* makes the relation between premise and act exhibit a kind of

[2] An interaction between positioning and accessibility seems less likely for *daarom*. Because it has retained its original anaphoric meaning, in that it refers back to information mentioned in the preceding discourse, it always marks antecedents that are accessible to the reader or hearer.

'obvious' relation". If *dus* indeed functions as the Dutch equivalent of English *so*, it is likely that its use as an accessibility marker differs in its positioning from its use as a connective.

(4) Hypotheses for the diachronic study:
   a. If *dus* and *daarom* express more than one function during a certain time span, these functions have different positioning preferences.
   b. Once *dus* and *daarom* come to serve as connectives, different positionings are related to usage in different domains.
   c. Once *dus* comes to serve as discourse marker signaling accessibility, its use as an accessibility marker differs in positioning preferences from its use as connective.

To investigate these hypotheses, I have looked at diachronic changes in the text-linguistic functions and the positioning of *dus* and *daarom*. The methodology of this study is the topic of the next section.

7.3 Methodology

For my study of the diachronic development of *dus* and *daarom* I used the samples of texts that were introduced in Chapter 4 (see section 4.2). The procedure for selecting connectives discussed there resulted in the division over text and periods mentioned in Table 7.1.

Table 7.1. Number and nature of connective fragments selected for diachronic analysis

Period	**# of non-rhyming and/or non-literary fragments**	**# of rhyming and/or literary fragments**	**Total**
13th century	25	25	50
16th century	25	25	50
20th century	25	25	50
Total	75	75	150

All 150 fragments per connective were subject to one conceptual analysis and two syntactic analyses. The conceptual analysis involves an analysis based on domains. The syntactic analyses include an examination of the categorical status of *dus* and *daarom* and their positioning within the clauses they appear in. As Chapter 2 already showed, these two syntactic properties are not completely independent of each other. The 20th-century *dus*-fragments were subject to an additional conceptual analysis based on accessibility.

In order to perform the conceptual analyses, for each fragment I selected a large context and made a translation if necessary. Per fragment, I determined whether *dus* and *daarom* were used as causal connectives by looking whether an implication relation ($p \rightarrow q$) between the *dus*- or *daarom*-clause and the preceding clause could be derived. Certain fragments showed ambiguity between the function of causal connective and another conceptual function (e.g. of an anaphor referring to the previous clause as a whole). For example, in (5) it is possible to interpret *daer om* as having an anaphoric adverb meaning (which in modern Dutch would be signaled by *daarover*). In this case S2 should be translated as "he (Reinout) was mad about that", where "that" refers to the whole proposition in S1 and not just to a constituent of it. In the causal interpretation, *daer om* expresses a cause-effect relationship between the entire proposition in S1 and the following S2. In that case *daer om* takes S1 as the anaphoric ground for "his anger".

(5) [$_{S1}$ *ic heb tegen Reynout een camp an ghenomen om dat hi mi huden die verradenis dede. ende nam mi Ywijn;*] [$_{S2}$ ***daer om*** *was hi toernich*]
(Historie van den vier heemskinderen, 1508)
'I have opened a combat with Reinout because he betrayed me today and he took Ywijn away from me; that's why he was angry / he was angry about that'

For ease of statistical analysis, I decided to classify ambiguous fragments like (5) as connective fragments. In the discussion of the results, however, I will make some remarks on the number of ambiguous fragments.

Each of the connective fragments was subject to the paraphrase test (see section 4.3 in Chapter 4) to determine the relevant domain of use (i.e. *content*, *epistemic*, or *speech act*). For an additional conceptual analysis of the 20th-century *dus*-fragments I marked the information in the *dus*-clause as 'accessible' if any of the following criteria was satisfied: a) the information was overtly or anaphorically available from the discourse, b) it was available by inference from the previous discourse, or c) it was pragmatically available by either social or world knowledge (cf. Prideaux 1993: 58). All other information was treated as new.

In the first syntactic analysis I looked at the positioning of *dus* and *daarom* in a linear way, distinguishing between clause-initial and clause-medial use (see also Chapter 2, section 2.3.1).[3] Fragments were labeled 'clause-initial' if *dus* or *daarom* occurred at the head of the clause (as in (6)). This category includes fragments in which the connectives are preceded by a coordinator like *en* 'and' (see (7)). Fragments were labeled 'clause-medial' if the connective occurred after the finite verb (in the case of main clauses) or after the subject (in subordinate clauses). An example of clause-medial use in a subordinate clause is given in (8).

(6) *O, oom George, Peggy Ann heeft nog iets vergeten, maar ze durfde niet meer te bellen,* ***dus*** *doe ik het maar.* (De gouden handjes, 1993)
'Oh, uncle George, Peggy Ann forgot something, but she did not dare to call, so now I am doing it.'

(7) *Buiten werd zowaar luid meegezongen en* ***daarom*** *durfde ik wat harder te spelen.*
(Het woelen der gehele wereld, 1993)
'Outside someone was even loudly singing along and that's why I dared to play somewhat louder.'

(8) *Als die Fransoysen die coninc* ***dus*** *hoirden roepen (…)*
(Historie van den vier heemskinderen, 1508)
'When the French heard the king call this way (…)'

In the second syntactic analysis I took a hierarchical approach to the positioning of *dus* and *daarom*. This analysis involved a classification based on categorical status, with a distinction between *adverbial*, *complementizer* and *ambiguous use*. All fragments in which *dus* or *daarom* occurred in clause-medial position were considered as adverbial fragments. In addition, the adverbial use contains fragments in which *dus* and *daarom* appear in clause-initial topic-position and trigger inversion of the finite verb and the subject. In the clause-initial example in (6), the adverb *dus* triggers inversion of the subject *ik* 'I' and the verb *doe* 'do'. In an example like (7), the adverbial status of *daarom* is supported by the fact the connective is preceded by a coordinator.

[3] The clause-final position has been disregarded here: the extraposition possibilities of the Middle Dutch language make it very hard to discriminate between clause-medial and clause-final positions.

The category of complementizer use (which only applies to *dus*) includes fragments in which *dus* appears at the head of the clause, and in which *dus* does not trigger inversion of the subject and the finite verb (see (9)).

(9) *Patrick woonde in een zijstraat van de laan van Nieuw Guinea,* ***dus*** *Hendriks* <u>*kwam*</u> *er vlak langs.* (De kunstrijder, 1989)
'Patrick lived in a road off the New Guinea avenue, so Hendriks came right past it.'

The category of *ambiguous use* contains fragments in which *dus* and *daarom* also head the clause in which they appear, but in which the absence or presence of inversion cannot be established because the subject is absent. This is only the case in imperatives like (10). Here, the finite verb *toont* 'show' can occupy the first position within the imperative clause (so-called V1), which implies that *dus* should be interpreted as a complementizer. Alternatively, the finite verb can be in second position (V2), with *dus* as a topicalized adverb in front of it. Because the subject is absent, no decision between these syntactic analyses can be made. Therefore, this fragment is labeled *ambiguous*. Note that this second option is only available for Middle Dutch adverbs (cf. Van Gestel 1992: 46-47); in Modern Dutch imperatives, the finite verb obligatorily occupies the first position.

(10) *En nu die sommige out en cout / sy hebben niet waer mee*
dat sij haer hongerige buyck / sullen versaeden
dus <u>*toont*</u> *aan haer u lieft* (Spel van sinnen, 1597)
'And now some old and cold ones, they don't have anything with which they can satisfy their hungry stomachs, so show your love to them.'

The results of these conceptual and syntactic analyses are discussed in section 7.4 (on *dus*) and 7.5 (on *daarom*). Statistical details on these analyses are provided in Appendices F and G.

7.4 Results of the diachronic analysis of *dus*

This section provides the results of the diachronic analysis of *dus*. In section 7.4.1 I answer the question: Did any conceptual changes occur during the selected time span? Then the syntactic changes are presented (7.4.2), after which I turn to the hypotheses by looking at the interaction between the conceptual and syntactic properties of *dus* (7.4.3-7.4.5). The main findings and conclusions concerning *dus* are summarized in section 7.4.6.

7.4.1 Conceptual analysis of *dus*

The word *dus* exhibits a variety of both connective and non-connective functions, which I have subdivided into three main groups: *anaphor*, *discourse marker*, and *connective*. As Figure 7.1 depicts, the meaning of the word *dus* has changed drastically over the years (χ^2 (4) = 64.5; $p < .001$, see Appendix F-2). From a word mainly used as an anaphoric expression, it changed into a word mainly used as a causal connective. This change does not necessarily imply that the connective use developed out of the anaphoric use, but that the decrease in ambiguous fragments can be seen as evidence in line with such a conceptual development. Six of the seven connective fragments from the 13th century are ambiguous between an anaphoric and a connective reading. In contrast, only five of the 21 connective fragments from the 16th century and none of the fragments from the 20th century show such an ambiguity.

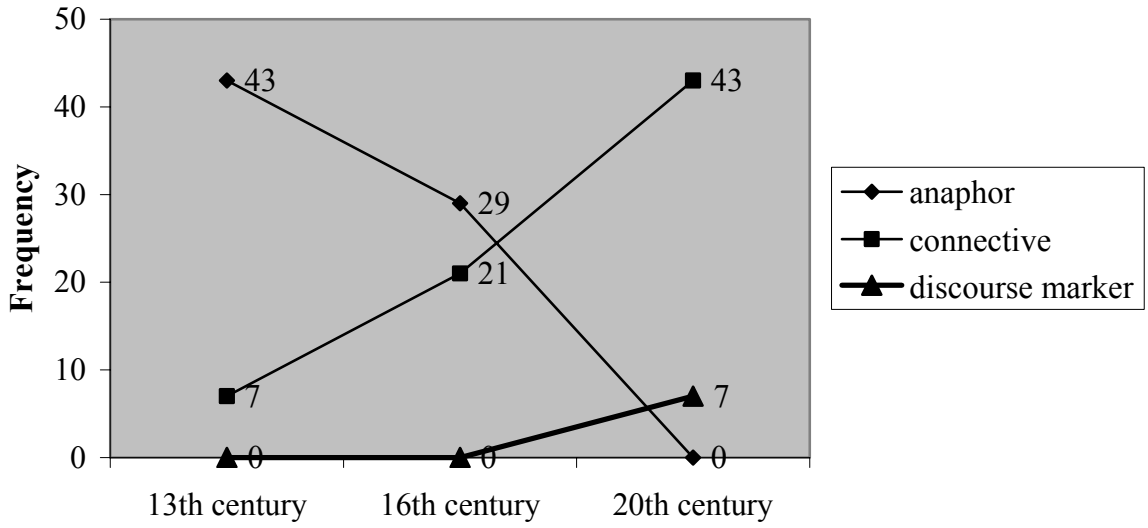


Figure 7.1. *Dus*: Conceptual functions in three periods

Below, the diachronic development of the three different uses will be discussed in more detail. In the 13th and 16th centuries, *dus* could function as an anaphoric expression, meaning 'thus' or 'this way'. As Figure 7.1 shows, this use has disappeared in the 20th century, which results in a significant change over time ($z = -3.69$; $p < .001$, see Appendix F-2). Formulating the disappearance of this use more precisely, I can state that *dus* has lost three subtypes of anaphoric use. First of all, in 20 medieval Dutch fragments *dus* can be translated as 'as follows', referring cataphorically to a part of an upcoming clause. This cataphoric use, illustrated in (11) and (12), only occurred during the 13th century.

(11) [text about a tree] (Der naturen bloeme, 1275-1300)
ende dese es ***dus*** *ghenaturt dattet groene .i. iaer ghedurt*
'and its nature is such that the green remains for one year'

(12) *Die gigant wie lesent* ***dus****. Hi was lanc .x. cubitus.* (Rijmbijbel, 1275-1300)
'The giant, we read as follows: He was ten ells tall.'

A second anaphoric translation of *dus* is 'thus/in that way'. As (13) and (14) exemplify, this *dus* refers back to information mentioned in the previous discourse, often the previous clause.

(13) *Ende alle de ghemeyn soudeniers quamen in Kaerls heyr daer si minnelijc ontfangen ende versolt werden. Dewile dat dit* ***dus*** *geschiede so quam die goede hertoge van Monbaes (…)* (Droefliken strijt van Roncevale, 1510-1530)
'And all the mercenaries came to Karl's army together, where they were welcomed and paid. At the time this was happening "in this way" the duke of Monbaes came.'

(14) *"Vrouwe," seit hi, "ic heb u gelogen!*
Wi hebben enen vader ende enen moeder, / si es mijn suster ende ic haer broeder."
Dus *began hi in sine tale werren.* (Floris ende Blanchefloer, 1260)
'"Woman", he said, "I lied to you! We have one father and one mother, she is my sister and I am her brother." So he began in his gibberish.'

This second anaphoric use only occurred in the 13th and 16th centuries. Both the first and the second anaphoric use have been preserved in the modern Dutch word *aldus*. This fixed

combination developed out of the separate words *al* and *dus,* which in the 13th century already appeared in combination relatively frequently.[4] This combination, which also shows up in my sample from the 13th century, could serve both the cataphoric and the anaphoric function, as the following examples illustrate.

(15) *Want die wise man spriekt* ***aldus****. duo girigheit van di. inde angst sal die laten*
(Nederrijns Moraalboek, 1270-1290)
'For the wise man speaks as follows: put greed from you and fear will leave you.'
(16) *Slape luttel ende bade alle tíít vro.* ***aldus*** *holdes tu din herte.*
(Noordlimburgse gezondheidsregels, ± 1254)
'Sleep a little and always bathe cheerfully; in that way you protect your heart.'

A third anaphoric function of *dus* is its use as an intensifier or degree adverb, similar to English 'so/to such an extent'. This function, which only appeared in the 16th century, is illustrated in (17) and (18).[5]

(17) *Als Hughe* ***dus*** *deerlic claechde liggende gebonden aen handen ende voeten (…)*
(Historie van Hughe van Bordeus, 1530-1550)
'When Hugh complained so pitifully, lying bound hand and foot (…)'
(18) *(…) hebdi mijns* ***dus*** *lange gederft ghi sult noch wel een half ure ontbeyden.*
(Historie van Malegijs, 1556)
'(…) did you get along without me for such a long time, you can do without me another half hour.'

Apart from the three types of anaphoric use mentioned above, *dus* appears in another non-connective use. This discourse marker use, which is illustrated in (19) and (20), shows up in the 20th century. In these fragments *dus* neither expresses a causal relation, nor is it used as an anaphoric expression. Its function is best described as signaling information status: the information presented in the *dus*-clause is marked as already being familiar to the reader or listener. For example, in (19), taken from a newspaper interview with the winner of a cooking contest, *dus* signals the knowledge state of the speaker, who considers the information presented as accessible to the hearer (or reader).

(19) [Fragment from a newspaper interview with the winner of a cooking contest]
(…) het was de eerste keer dat hij aan een dergelijke wedstrijd deelnam. Maar zijn zachte bolletjes en het harde Wener vruchtengebak werden tijdens de voorronde van de Heerenveense school als beste beoordeeld. 'Het ging erom wie het mooiste, het beste produkt maakte. Dat was ik ***dus****.'* (Meppeler Courant, 1995)
'(…) it was the first time that he took part in a competition like that. But during the preliminary his soft rolls and Wiener fruit pie were judged the best. "The point was who could make the most beautiful, the best product. And that was me."

---

[4] In the population of 13th-century texts *dus* occurred 1734 times. Almost one third of these *dus*-fragments (554) concerned instances of the combination *al+dus*. To indicate the remarkable frequency of *al+dus*: the second most frequent medieval Dutch combination, *dus+gedane* (which resulted in modern Dutch *dusdanig(e)* 'such/to such an extent') only appears 61 times.

[5] As Michels (1949: 212) notices, the combination *dus lang(e)* even developed an idiomatic interpretation meaning *tot nu toe* 'up till now'. This interpretation is also possible in (18).

(20) [Context: this is a report on the opening of a swimming pool. The information that the opening was done by Hilde Zoer had been mentioned earlier in the text.]
Op het moment dat de eerste burger zich openlijk afvroeg wie de openingshandeling wilde verrichten, ging een groot aantal vingers de lucht in. Het werd ***dus*** *Hilde Zoer.* (Meppeler Courant, 1995)
'At the moment the mayor explicitly asked who wanted to do the opening ceremony, many people raised their hands. So Hilde Zoer became the person to do it.'

The category of discourse markers also includes fragments in which *dus* marks a paraphrase relation between the *dus*-clause and (part of) the previous clause. An example is given in (21).

(21) *Hoe komt een advertentie tot stand. De acquisitie zoekt adverteerders op, gaat winkeliers bij* [sic!] *langs kledingzaken, meubelzaken, groentezaken, supermarkten etc. Ze vragen de teksten, wat* ***dus*** *in de advertentie moet komen te staan.* (Meppeler Courant, 1995)
'How an advertisement comes about. The canvasser visits advertisers, calls on storekeepers at clothes stores, furniture firms, greengroceries, supermarkets etc. They ask for the texts, so for what should be placed in the advertisement.'

In addition to the two non-connective functions (anaphor and discourse marker), *dus* can also fulfill the function of a connective that marks causal coherence relations. The connective use hardly occurred in the 13th century, whereas it is the most prominent use in the 20th century. In the 16th century, *dus* can be characterized as a causal connective in 21 fragments; in the other 29 examples *dus* can only be interpreted as retaining its original manner adverb meaning. Examples (22) and (23) show fragments in which *dus* introduces an event as being causally linked to the occurrence of the event described in the previous clause.

(22) *Vergeet niet dat de paus in Avignon kortgeleden alle aanhangers van zijn rivaal in Rome in de ban heeft gedaan: de Romeinse paus, al zijn kardinalen, bisschoppen, abten en volgelingen. Daartoe behoren wij ook.* ***Dus*** *<u>mogen</u> we nog van geluk spreken dat de monniken hier ons uit dat noodweer hebben gehaald en in deze stinkende stal hebben opgesloten.* (De doge-ring van Venetië, 1994)
'Don't forget that recently in Avignon the pope has put all the followers of his rival under the ban: the Roman pope, all of his cardinals, bishops, abbots, and followers. We belong to that group as well. So we are lucky that the monks here took us out of that heavy weather and locked us in a stinking stable.'

(23) *Patrick woonde in een zijstraat van de laan van Nieuw Guinea,* ***dus*** *Hendriks <u>kwam</u> er vlak langs.* (De kunstrijder, 1989)
'Patrick lived in a road off the New Guinea avenue, so Hendriks came right past it.'

As a causal connective *dus* can occur in all three domains of use (see Table 7.2).

Table 7.2. *Dus*: Domains of use in three periods

	Content	**Epistemic**	**Speech act**	**Total**
13th century	2	5	0	7
16th century	9	3	9	21
20th century	8	34	1	43
Total	19	42	10	71

The distribution of *dus* over the different domains is not stable over time (χ^2 (4) = 31.1; p < .001); in the 20th century there is an increase in the use in the epistemic domain (z = 3.29; p = .001, see Appendix F-3).

Over the years, the meaning of the word *dus* has changed drastically. It has lost its original anaphoric meaning and a new use as discourse marker has appeared. Furthermore, the number of fragments in which *dus* occurs as a causal connective has increased.

7.4.2 Syntactic analysis of *dus*

A first syntactic analysis of *dus* involves its categorical status. In the adverbial example in (22), *dus* triggers inversion of the finite verb and the subject. This inversion cannot be observed in (23), which indicates that *dus* serves as a complementizer there. As Figure 7.2 shows, the adverbial use of *dus* is the most prominent use in all three periods under investigation, and the only use in the 13th century. In Figure 7.2, the ambiguous use and the complementizer use have been grouped together as non-adverbial use.

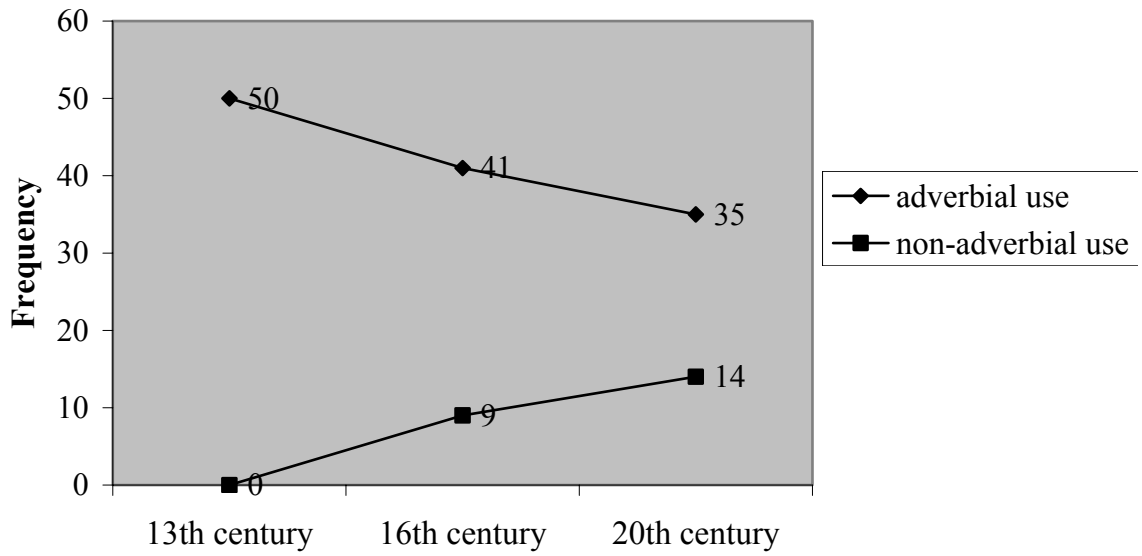


Figure 7.2. *Dus*: Categorical status in three periods[6]

The nine fragments of non-adverbial use in the 16th century all involve ambiguous cases, in which neither inversion nor the absence of inversion can clearly be established. This is the case in imperatives like (24), but the phenomenon also occurs with adverbs other than *dus*.

(24) *Wi moeten van dese werelt sceyden / Als onsen tijt hier comen sal*
Dus *<u>wil</u> hem elc daer toe bereyden / Oft anders coemter af misual*
(Devoot ende profitelyck boecxken, 1539)
'We have to depart from this world / When our time here comes
So everyone should prepare himself for that / Or else it will cause mischief'

The real complementizer use only appears in the 20th-century sample. This use, then, can be regarded as a relatively new syntactic phenomenon in *dus*-fragments.[7]

[6] In this figure, one 20th-century fragment has been disregarded, since it does not contain a verb.
[7] All fragments with *dus* as a complementizer show verb second, which indicates that *dus* must be regarded as a coordinator. In contrast, Zaalberg (1973: 146-147) observes that 18th-century *dus* could sometimes show up as a subordinator (showing V-late). His observation is not supported here.

My second syntactic analysis involves the positioning of *dus*. Both the complementizer and the ambiguous use of *dus* show a fixed position at the head of the second of the combined clauses. As an adverb, however, *dus* can occur in several positions (see the distribution of the adverbial fragments in Table 7.3). The example in (22) can be classified as clause-initial; the two discourse marker fragments in (19) and (20) are instances of clause-medial *dus*.

Table 7.3. *Dus*: Positioning of the adverbial fragments in three periods

	In S1	**In S2**		**Total**
		Clause-initial	**Clause-medial**	
13th century	20	23	7	50
16th century	0	16	26	42
20th century	0	8	27	35
Total	20	47	60	127

In the 13th century *dus* could occur in the first of the two combined clauses, instead of in the second, which is the norm for *dus* in modern Dutch. This "S1"-use is restricted to the cataphoric function of *dus*, of which (25) gives an example. In 14 of the 20 "S1"-instances *dus* occurs in combination with *al* (as in (26)). In line with the disappearance of the cataphoric use, this "S1"-positioning of *dus* does not occur in the samples of the 16th and 20th centuries.

(25) *Die gigant wie lesent dus. Hi was lanc .x. cubitus.* (Rijmbijbel, 1275-1300)
'The giant, we read as follows: He was ten ells tall.'

(26) *Want die wise man spriekt* ***aldus****. duo girigheit van di. inde angst sal die laten*
(Nederrijns Moraalboek, 1270-1290)
'For the wise man speaks as follows: put greed away from you and fear will leave you.'

To conclude, *dus* shows several changes at the syntactic level. First of all, it has gained the categorical status of complementizer in addition to its original adverbial status. Furthermore, it has lost its ability to occur in S1.

Having discussed the conceptual and syntactic characteristics of *dus* independently of each other, it is now time to turn to the interaction between them. Because of the diversity of conceptual uses, I performed three separate statistical analyses (see Appendix F). The first analysis takes into account all three functions of *dus* (anaphor, connective, discourse marker, see section 7.4.3); the second focuses on the domains of use within the connective function of *dus* (section 7.4.4); the third deals with the accessibility of the information in the *dus*-clause (section 7.4.5). Within each analysis I investigated whether the different conceptual uses show different distributions over categorical status and/or positioning.

7.4.3 Conceptual functions in relation to the syntactic properties of *dus*

To start, I have looked at the interaction between the three conceptual functions of *dus* and their respective syntactic properties (categorical status and positioning). The fragments with the combination *al+dus* (29 in sum) were disregarded in this statistical analysis, since this combination could only occur with an anaphoric meaning (whereas *dus* could also be used as a causal connective by this century).[8] Figure 7.3 shows the interaction with categorical status. Again, ambiguous and complementizer use are grouped together as non-adverbial use.

[8] These findings on *aldus* contrast with the description of *aldus* in the *WNT*. The *WNT* claims that *aldus* could also fulfill the role of causal connective.

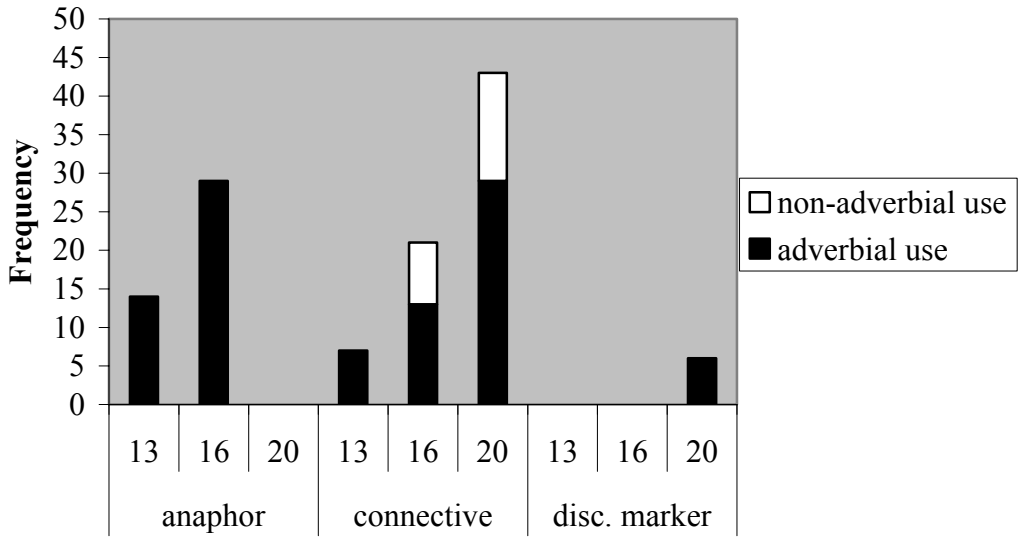


Figure 7.3. *Dus*: Interaction between functions and categorical status

The distribution over adverbial and non-adverbial usage is not the same for the three functions (χ^2 (2) = 24.0; p < .001, see Appendix F-1); the anaphoric use of *dus* is always adverbial (i.e. the non-adverbial use never occurs, z = -2.29; p = .02). The discourse markers in my sample have a preference for the adverbial use as well.[9, 10] The non-adverbial use is restricted to the connective function; only in the 20th century this non-adverbial use involves the use as a complementizer (see section 7.4.1). Categorical status, then, only gives a partial distinction between the three conceptual functions: the adverbial use is suitable for all three functions of *dus*, whereas the non-adverbial use is preserved for the connective function.

The three conceptual functions of *dus* interact with positioning as well (χ^2 (2) = 24.8; p < .001, see Appendix F-2). Figure 7.4 depicts the distribution of the three functions over the clause-initial and the clause-medial position within S2, the second of the combined clauses.[11, 12] It shows that the anaphoric use has a preference for the clause-medial position (z = 4.54; p < .001), and that the discourse marker use in this sample is even restricted to the clause-medial position.[13] Statistical analysis of the data in Figure 7.4 reveals that the distribution over the two positions changes over time (χ^2 (2) = 25.3; p < .001, see Appendix F-2): in the 16th and the 20th centuries *dus* occurs more often in the clause-medial position than in the 13th century. The increase in the 16th century can be related to the preference of anaphoric *dus* for

[9] At least in the fragments that have been taken into account in this statistic analysis; there is also one discourse marker fragment that has been disregarded here because it lacks a verb.

[10] With the rise of *dus* as discourse marker, a new syntactic use shows up as well. In the fragments in which *dus* marks a paraphrase, it often relates to a previously mentioned phrase (e.g. an NP), instead of a full clause. See (i) for an illustration (see example (21) for more context of this *dus*-fragment):

(i) *Ze vragen de teksten, wat* ***dus*** *in de advertentie moet komen te staan.* (MC, 1995)
'They ask for the texts, so for what should be placed in the advertisement.'

This indicates an extension of the syntactic possibilities of *dus* compared to the earlier periods.

[11] Another interaction has already been mentioned in section 7.4.2: the occurrence of *dus* in S1 is restricted to the cataphoric use of *dus*. This interaction is supported by the finding that – after the 13th century – the S1-use has disappeared, together with the possibility of cataphoric interpretations.

[12] Remember that the clause-initial position includes both complementizers and topicalized adverbs.

[13] Strictly speaking, there is no significant interaction effect between disourse marker use and positioning, which is probably due to the relatively low number of discourse marker fragments in my sample.

the clause-medial position. The increase in the 20th century can be related to the increase in the clause-medial use of connective *dus* (z = 2.26; p = .02, see Appendix F-4).

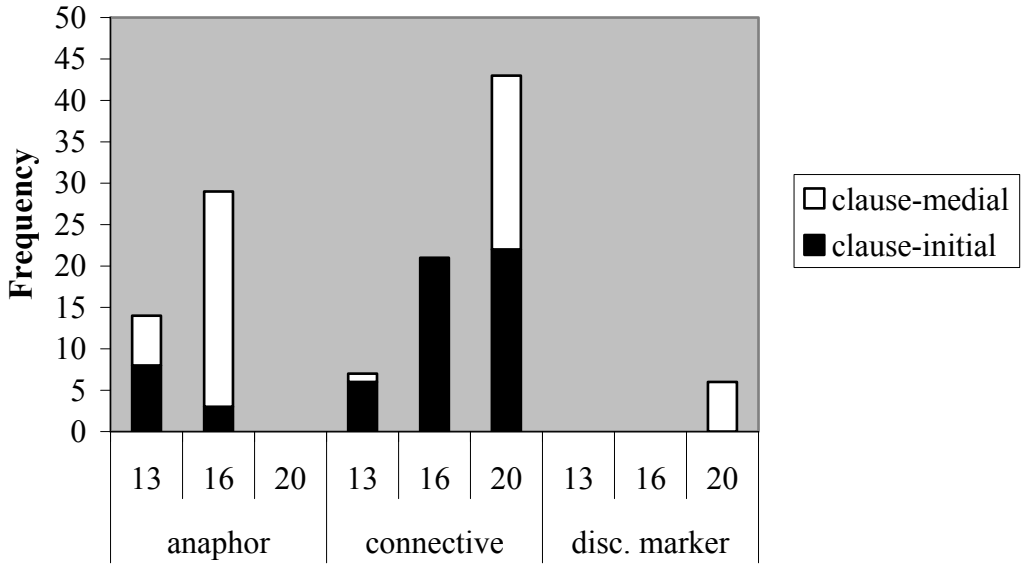


Figure 7.4. *Dus*: Interaction between functions and position

Positioning, then, gives a partial distinction between the three conceptual functions: in the 13th and 16th centuries the anaphoric use has a preference for the clause-medial position, whereas the connective use does not. In the 20th century the discourse marker use is restricted to the clause-medial position, whereas the connective use is not.

7.4.4 Domains of use in relation to the syntactic properties of *dus*

My second interaction analysis focused on the connective function of *dus*. I investigated whether use in different domains could be related to differences in categorical status and/or positioning. Table 7.4 introduces a more detailed picture of the connective use of *dus* in relation to categorical status. Statistical analysis of these data (see Appendix F-3) indicates that these two factors interact (χ^2 (2) = 25.6; p < .001). As a causal connective marking speech-act relations, *dus* has a preference for non-adverbial use (z = 2.83; p = .004). That the categorical status of *dus* does not relate to different domain interpretations in yet another way is in line with the observation in Pander Maat & Sanders (1995: 354) that there are no systematic effects in this area.

Table 7.4. *Dus*: Distribution of adverbial and non-adverbial use over the three domains

		13th century	16th century	20th century	Total
Content	adverbial	2	9	6	17
	non-adverbial	0	0	2	2
Epistemic	adverbial	5	3	23	31
	non-adverbial	0	0	11	11
Speech act	adverbial	0	0	0	0
	non-adverbial	0	9	1	10
Total		7	21	43	71

As both Figure 7.4 (in section 7.4.3) and Table 7.5 below show, the connective *dus* occurs in clause-medial position about half of the time in the 20th century, whereas it does not show up

in that position at all during the 16th century. This change is significant (χ^2 (2) = 20.0; p < .001).

Table 7.5. *Dus*: Distribution of the three domains over the positions in three periods

	13th century		**16th century**		**20th century**		**Total**
	Initial	**Medial**	**Initial**	**Medial**	**Initial**	**Medial**	
Content	1	1	9	0	7	1	19
Epistemic	5	0	3	0	14	20	42
Speech act	0	0	9	0	1	0	10
Total	6	1	21	0	22	21	71

The distribution in Table 7.5 does not result in significant interactions between domains of use and positioning. However, it can be observed that in its speech-act use the connective *dus* is restricted to the clause-initial position. This is in line with the observation that speech-act *dus* has a significant preference for non-adverbial use.

7.4.5 Accessibility in relation to the syntactic properties of *dus*

The third and final interaction analysis concerns an investigation of the connective fragments based on accessibility. This analysis is restricted to the connective fragments from the 20th century, because the connective use from the 13th and 16th centuries hardly showed any variation in terms of positioning (see Table 7.5 in section 7.4.4).

The rise of the discourse marker use already indicates that *dus* can signal the information status of the proposition in which it occurs (at least in the 20th century). In that case, *dus* only signals information status. In other fragments, *dus* only serves the connective function, marking a conclusion that is completely new (or not accessible) to the reader or listener. Analysis of the 20th-century connective fragments shows that there are also fragments in which *dus* fulfills a double role: both as a marker of a causal coherence relation and as a marker signaling information status of the clause. In such fragments *dus* marks a consequence that the writer (or speaker) assumes to be familiar to the reader (or listener). For example, in (27) *dus* signals that the conclusion is already available to the reader, which is obvious, because the information in the *dus*-clause has been presented earlier in the text.

(27) [Context: A mother leaves her house and is about to look for her two children, thinking they ran off to another part of the island. At the moment she gets into her car, she sees them playing at the beach near their house.]
Het is zó normaal, zó vanzelfsprekend, die twee spelende kinderen in de baai, zó precies zoals het moet zijn, dat haar gevoel van verrassing en opluchting meteen wegebt. Ze stapt uit de auto, en loopt met de handen in haar zakken naderbij. Die belhamels waren de hele ochtend ***dus*** *gewoon waar ze hoorden: hier bij haar.* (Verborgen gebreken, 1996)
'It is *so* normal, *so* natural, those two children playing in the bay, *so* exactly the way it should be, that her feeling of surprise and relief drains away immediately. She gets out of the car, and walks closer with her hands in her coat pockets. Those rascals had been where they belonged all morning: here with her.'

The conclusion that the children had not ran off but were where they belonged, can already be inferred from the information that they were just playing at the beach near the house. This conclusion can thus be regarded as an accessible conclusion. In other fragments, the *dus*-

clause presents the only 'logical' consequence of information presented earlier. In such cases (see (28) for an example), this conclusion can also be regarded as accessible to the reader, since the reader could have inferred the conclusion using world knowledge that is commonly available to all language users.

(28) [Topic is the number of Sundays in a year; in 1995 there were 53 instead of 52 Sundays.]
Als ik het goed heb uitgerekend, herhaalt zich deze situatie in het veelbesproken jaar 2000, want dan valt nieuwjaarsdag op een zaterdag en de eerste zondag ***dus*** *op 2 januari.* (MC, 1995)
'If I calculated it correctly, this situation repeats itself in the much-discussed year 2000, because then New Year's Day will fall on a Saturday and thus the first Sunday on the second of January.'

Table 7.6 presents data on the interaction between accessibility and categorical status. Statistical analysis of these data shows that the factor 'accessibility' does not interact with categorical status (χ^2 (1) = 0.1; p = .25).

Table 7.6. *Dus*: Accessibility status of the connective fragments in relation to categorical status (only 20th century)

	Not accessible	**Accessible**	**Total**
Adverbial use	13	16	29
Complementizer use	7	7	14
Total	20	23	43

Table 7.7 depicts the distribution of accessible and non-accessible *dus*-fragments over the two positions.

Table 7.7. *Dus*: Accessibility status of the connective fragments in relation to positioning (only 20th century)

	Not accessible	**Accessible**	**Total**
Clause-initial	14	8	22
Clause-medial	6	15	21
Total	20	23	43

Statistical analysis of these data reveals that accessibility and position interact (χ^2 (1) = 5.4; p < .025): accessible conclusions are accompanied more often by *dus* in clause-medial position.

7.4.6 Conclusions about *dus*

From the diachronic analysis of *dus* it can be concluded that the conceptual and syntactic properties of *dus* interact. First of all, the "S1-use" of *dus* in the 13th was preserved for the cataphoric function of *(al)dus*. With the disappearance of this conceptual function, *dus* has lost its ability to appear in S1 as well.

Secondly, it can be concluded that categorical status is useful to discriminate between the three conceptual functions: the anaphoric use of *dus* is always adverbial, just like the discourse marker use. The connective function of *dus*, on the other hand, can either be adverbial or (from the 16th century on) non-adverbial. To be more precise about this interaction between connective use and categorical status, it can be concluded that especially

the speech-act use of *dus* has a preference for the non-adverbial use. In the other two domains *dus* does not have a preference in terms of categorical status.

A third conclusion is that the position of *dus* does not interact with different domains within the connective function, but that it does interact with the three functions (anaphor, connective, discourse marker). The discourse marker use is restricted to the clause-medial position; the anaphoric use has a preference for this position, whereas the connective use has a preference for the clause-initial position. Especially in the 16th century, the different preferences of anaphoric and connective *dus* result in a nice division of labor. It seems as if the 16th-century language users stressed the intended function by selecting a specific position for *dus*: clause-initial in the case of connective use, and preferably clause-medial in the case of anaphoric use.

Is it remarkable that the 20th-century connective use shows an increase in clause-medial positions? On the one hand it is not: in the 20th century, positioning is no longer needed to discriminate between connective use and anaphoric use, because the anaphoric use of *dus* has disappeared. On the other hand it is: with the rise of the discourse marker use, the language user could have "chosen" to preserve the clause-initial position for connective *dus*, and the clause-medial position for discourse marker *dus*. A clear distinction in positioning possibilities would have supported the conceptual distinction between the two. However, the increase in clause-medial connective *dus* can be explained on the basis of an accessibility perspective. The connective clauses with *dus* in a clause-medial position often present conclusions that are accessible to the hearer or reader. These *dus*-fragments with an accessible conclusion show a positioning behavior that is similar to that of discourse markers, which have a preference for the clause-medial position as well. From this perspective, then, it is less remarkable that in the 20th century the connective use of *dus* is not restricted to the clause-initial position anymore.

7.5 Results of the diachronics of *daarom*

This section provides the results of the diachronic analysis of *daarom*. Section 7.5.1 presents the conceptual developments, section 7.5.2 treats the syntactic developments, and section 7.5.3 discusses several interactions between the conceptual and syntactic properties of *daarom*. The main findings concerning *daarom* are summarized in section 7.5.4.

7.5.1 Conceptual analysis of *daarom*

During each of the three periods the word *daarom* mainly shows up as a causal connective (see Table 7.8).

Table 7.8. *Daarom*: conceptual functions in three periods

Period	Connective use	Non-connective use	Total
13th century	44	6	50
16th century	45	5	50
20th century	50	0	50
Total	139	11	150

An example in which *daarom* introduces S2 as having a cause-effect relationship to S1 can be found in (29). *Daarom* in S2 introduces an event (the fact that the thief calls his girlfriend for help) as being causally linked to the occurrence of the event previously described in S1 (the fact that the loot did not fit into his car).

(29) [$_{S1}$ *De buit (…) bleek zo groot te zijn dat hij die niet allemaal in zijn auto kwijt kon.*]
[$_{S2}$ ***Daarom*** *belde hij zijn vriendin (…) om hulp.*] (MC, 1995)
'The loot (…) appeared to be so big that he couldn't dispose of all of it in his car.
That's why he called his girlfriend for help.'

In the 13th and 16th centuries, the word *daarom* could also be used in another way (see the column *non-connective use* in Table 7.8). An illustration of this *non-connective use* is given in (30). In this example *daer om* refers back to an object (a golden cup) mentioned earlier in the previous clause and not to the preceding proposition as a whole.

(30) *Die portwerder hadde groot ghepeins*
[$_{S1}$ *ende was in anxte in sire herten binnen, / hoe hi den cop moghe ghewinnen. (...)*]
[$_{S2}$ *Ende seide, dat hi gherne soude / dusent maerc* ***daer om*** *gheven van goude*]
(Floris ende Blanchefloer, 1260)
'The gatekeeper was thinking hard
and worried his mind / how he could get hold of the cup (...)
And said that he would gladly / give a thousand marcs of gold **for that**'

This specific anaphoric use is not possible anymore in the 20th century: in a modern Dutch variant of (30) the word *daarvoor* 'for that' would be used instead of *daarom*. This results in a significant change (χ^2 (2) = 7.8; $p < .025$, see Appendix G-1). However, there are present-day utterances in which it is still possible to refer to a constituent with *daarom*, despite the fact that these are not attested in my 20th-century sample. For example, in (31) *dar om* refers to *dat grael* 'the grail' in the preceding clause. The modern Dutch equivalent of this sentence would also contain *daarom*.

(31) *Oec sach ic weder ende vort / dat grael vor mi draghen*
mine horde nieman ***dar om*** *vraghen* (Roman van Perchevael, 1275-1300)
'I also saw that the grail was carried back and forth in front of me
I heard nobody ask for it'

Apart from the clearly anaphoric fragments (in which *daarom* cannot be interpreted as a causal connective at all), the category *non-connective use* also includes another conceptual function. The two occurrences of this function are given in (32) and (33). In both examples *daarom* functions as a relative adverb introducing a relative clause that restricts the interpretation of a noun phrase. In (32) the relative clause introduced by *daer omme* modifies the interpretation of the noun phrase *sijn begheeren* 'his longing'. In (33) the *daarom*-clause restricts the interpretation of the noun phrase *die drie saken* 'the three cases'.

(32) *Die heren die dancten hem seer ende heyndric boven al. ende hi vraechde hem wie hi was ende wat sijn begheeren was* ***daer omme*** *dat hi daer quam.*
(Historie van Margarieta van Lymborch, 1516)
'The lords thanked him greatly and Hendric above all, and he asked him who he was and what request it was he came for.'

(33) *Die drie saken* ***daer om*** *dat hi / soud hebben gheweest verdumt (…)*
(Sinte Lutgard, 1275-1300)
'The three cases for which he should have been cursed (…)'

Modern Dutch *daarom* cannot express this conceptual function anymore; instead, the word *waarom* 'why/for which' would be used.

The most prominent use of *daarom* is its use as a causal connective. As the following examples illustrate, *daarom* can be used in all three domains. In the content example in (34), *daer om* introduces the consequence of the fact that the horseman didn't have much money. The epistemic example in (35) presents a conclusion about friendship: people who leave their friends when these become poor cannot be regarded as real friends. The speech act 'prepare yourself' in (36) is supported by the remark in the preceding clause.

(34) *Die dit liedeken dichte / Dat was een ruyter fijn*
Sinen buydel was seer lichte / ***daer om*** *drinct hi selden wijn*
(Antwerps liedboek, 1544)
'The person who wrote this song / He was an excellent knight
His purse was very light/ that's why he seldomly drinks wine'

(35) *Die arm is die is al ene. dan besuokt die man sinen vrint. Mar man vinder dan luttel.*
Want si vlin den armen. Jnde ***dar vombe*** *en ist gene regte minne.* (NM, 1275-1300)
'Whoever is poor is on his own. Then the man looks for his friend, but one finds few, because they flee from the poor. And that's why it is not real friendship.'

(36) *En denct niet dat ic vervaert ben voer die Sarasinen,*
daer om *lieve swager maect u bereet* (Droefliken strijt van Roncevale, 1510-1530)
'And do not think that I am afraid of the Arabs, so – dear brother in law – prepare yourself'

As Table 7.9 shows, *daarom* has a constant preference for use in the content domain. Statistical analysis of these data reveals that the distribution over domains is not constant (χ^2 (4) = 18.6; $p < .001$): the number of speech acts decreases in the 20th century ($z = 1.83$; $p = .03$, one-sided).

Table 7.9. *Daarom:* Domains of use in three periods

Period	Content	Epistemic	Speech act	Total
13th century	27	10	7	44
16th century	28	9	8	45
20th century	42	8	0	50
Total	97	27	15	139

Within the connective function, *daarom* can be used both anaphorically and cataphorically. The examples in (34)-(36) already illustrated the anaphoric use of *daarom*: in these fragments *daarom* occurs in S2, referring back to an antecedent mentioned in the previous text. Examples of the cataphoric use of *daarom* can be found in (37) to (39). In these fragments, the clause containing the antecedent follows the *daarom*-clause.

(37) [S1 ***daer omme*** *hebbe wi desen wissel ghedaen doer sbeters wille*], [S2 *dat wi niet en wovden, dat die heren van scalvnen enghen kijf moichten hebben ieghens onsen lieden.*]
(CG1-975, 1290)
'That's why we have made this exchange for the good that we did not want the lord of Schalfnen to be in any conflict with us.'

(38) [S1 ***Dar vombe*** *wil ic setten en deel guoder worde die ic gehort hebe in maniegen steden in gescrigte. Inde die giene ak die ic seluer gevinden kan*] [S2 *vomb dat. dat gedinchenisse is ene lidende sake inde die hastelik vergait.*]

(Nederrijns Moraalboek, 1270-1290)

'That's why I want to put in writing a part of the good words that I heard in many places, and also those that I can remember myself, because remembrance is a fleeting thing that disappears quickly.'

(39) [S1 *dat deed hi* ***daer omme***], [S2 *op dat hi daer mede haer herte berrurde*]

(Leven van Sente Kerstine, 1275-1300)

'He did that for that (reason), so that he would touch her heart with that'

In sum, there are 14 cataphoric fragments: twelve in the 13th century and two in the 16th century. Some of these fragments show a double marking of the causal relation. For example, in (38) the causal relation is also signaled by *vomb dat* (*omdat* 'because') in S2. In (39) the causal relation is marked with *opdat* 'in order to'.

This section on the conceptual developments of *daarom* shows that from the 13th century on, *daarom* is mainly used as a causal connective. As a causal connective it has a preference for use in the content domain. Over the years, *daarom* has lost its ability to be used as a relative adverb. In addition, it has partially lost its ability to refer to a constituent within a previous clause.

7.5.2 Syntactic analysis of *daarom*

The diachronic analyses reveal that *daarom* has changed at the syntactic level. The conceptual difference between the anaphoric and the cataphoric use of *daarom* already indicated a first change at the syntactic level. In the 13th and 16th centuries *daarom* could either occur within S2 (see the examples in (34) to (36)) or in S1 (see (37) to (39)); in the 20th century the occurrence of *daarom* is restricted to S2.

A second syntactic change has been mentioned earlier in section 5.5.1 (on the diachronic analysis of *omdat*). This change involves the disappearance of cataphoric instances in which *daarom* is directly followed by *dat*, which resulted in the combination *daar-om-dat*. An example of this combined variant is given in (40). In modern Dutch this relation would be marked with *omdat*.

(40) *Die rede dar wi sunde vomb duon. dat is* ***dar omme dat*** *wi lange wanen leuen.*

(Nederrijns Moraalboek, 1270-1290)

'The reason why we commit sins, that is because we suppose to live long.'

Table 7.10 shows the results of the syntactic analysis of *daarom* based on positioning.

Table 7.10. *Daarom*: Positioning in three periods

	S1	**S2**		**Total**
		Initial	**Medial**	
13th century	12	31	7	50
16th century	2	35	13	50
20th century	0	34	16	50
Total	14	99	37	150

It appears that *daarom* has a preference for use in the clause-initial position. Within this position, *daarom* is mainly used as an adverb, triggering inversion of the subject and the finite verb. However, there are seven fragments in which its categorical status is ambiguous between adverb and complementizer (one in the 13th century and six in the 16th century). All of these ambiguous cases involve imperatives without a subject. A typical example is given in (41).

(41) *ic sal v corteliken al in mijn rijcke ontfaen*
daer om *strijt blijdelike* (Suverlijc boecxken, 1508)
'I will receive you in my kingdom shortly, therefore: fight cheerfully'

Unlike *dus*, *daarom* does not develop the syntactic ability to occur as a complementizer. Over the years, then, *daarom* remains an adverb with a preference for the clause-initial position within the second of the combined clauses.

7.5.3 Interaction between conceptual and syntactic properties of *daarom*

In what way do the conceptual and syntactic properties of *daarom* interact? A first interaction has already been mentioned: only the cataphoric fragments, in which *daarom* refers to an antecedent in the following clause, *daarom* can occur in S1. In all other cases *daarom* is positioned within the second of the combined clauses.

The distribution of *daarom* over the clause-initial and the clause-medial position within S2 is presented in Table 7.11.

Table 7.11. *Daarom*: Positioning within S2 in three periods[14]

	Connective		**Anaphor**		**Total**
	Initial	**Medial**	**Initial**	**Medial**	
13th century	29	3	1	4	37
16th century	34	9	0	4	47
20th century	34	16	0	0	50
Total	97	28	3	8	134

Statistical analysis of this distribution reveals that the distribution over the two positions is not constant (χ^2 (2) = 6.6; $p < .05$): in the 20th century there is an increase in use in the clause-medial position ($z = 2.32$; $p = .02$, see Appendix G-1).

In addition, the logit analysis shows that the two conceptual functions have different positioning preferences (χ^2 (1) = 16.2; $p < .001$). Unlike connective *daarom*, anaphoric *daarom* has a strong preference for the clause-medial position ($z = 3.54$, $p < .001$). In the 13th and 16th century, this mapping between positioning and function is not one-to-one, since there are twelve connective fragments in clause-medial position. A more detailed investigation of these twelve fragments reveals that nine of them involve cases in which *daarom* is ambiguous between the connective use and the anaphoric use. An example of such an ambiguous fragment is presented in (42). If *daer om* in (42) is interpreted anaphorically, S2 should be translated as 'he (Reinout) was mad about that', where 'that' refers to the whole proposition in S1. In the causal connective interpretation, *daer om* expresses a cause-effect relationship

[14] In this analysis, the two fragments in which *daarom* functions as a relative adverb have been disregarded.

between the entire proposition in S1 and the following S2. In that case *daer om* takes S1 as the anaphoric ground for "his anger".

(42) [$_{S1}$ *ic heb tegen Reynout een camp an ghenomen om dat hi mi huden die verradenis dede. ende nam mi Ywijn;*] [$_{S2}$ ***daer om*** *was hi toernich*]
(Historie van den vier heemskinderen, 1508)
'I have opened a combat with Reinout because he betrayed me today and he took Ywijn away from me; that's why he was angry / he was angry about that'

Disregarding the ambiguous fragments would leave only three connectives in clause-medial position. This results in a clearer contrast between the connective function and the anaphoric function in relation to their positioning preferences: in the 13th and 16th centuries, the connective function has a preference for the clause-initial position, whereas the non-causal anaphoric function has a preference for the clause-medial position.

A final analysis involves the interaction between the positioning of connective *daarom* and its conceptual characterization based on domains. Table 7.12 shows the distribution of the domains over the two positions.[15] Statistical analysis reveals that the distribution over the two positions changes over time (χ^2 (2) = 15.1; p < .001): in the 20th century there is an increase in use in the clause-medial position (z = 2.45; p = .01, see Appendix G-2). The positioning of *daarom* cannot be related to use in a certain domain.

Table 7.12. *Daarom*: Positioning within S2 in three periods (only connective function)

	Content		**Epistemic**		**Speech act**		**Total**
	Initial	**Medial**	**Initial**	**Medial**	**Initial**	**Medial**	
13th century	15	1	8	0	5	0	29
16th century	17	2	8	0	8	0	35
20th century	28	14	6	2	0	0	50
Total	60	17	22	2	13	0	114

7.5.4 Conclusions about *daarom*

The diachronic analysis of *daarom* shows that from the 13th century on *daarom* is used mainly as a causal connective, with a preference for use in the content domain. Over the years, *daarom* has lost its ability to be used as a relative adverb. In addition, it has partially lost its ability to refer to a constituent within a previous clause. Over the years, *daarom* remains an adverb with a preference for the clause-initial position within the second clause.

It can be concluded that the conceptual functions and syntactic properties of *daarom* interact. The "S1-use" of *daarom* in the 13th and 16th century is preserved for the cataphoric function of *daarom*. With the disappearance of this conceptual function, *daarom* has lost its ability to appear in S1 as well.

The two positionings cannot be related to different domains within the connective function. However, a contrast is found between the positioning preferences of the connective function and the anaphoric function: in the 13th and 16th centuries, the connective function has a preference for the clause-initial position, whereas the non-causal anaphoric function has a preference for the clause-medial position. In the 20th century, the clause-medial connective use shows a significant increase. This need not be remarkable: with the loss of (or at least the

[15] In this table, the fragments that are ambiguous between a connective and an anaphoric interpretation have been disregarded.

sharp decline in) the anaphoric use, the need to use positioning to discriminate between the two conceptual functions has disappeared.

7.6 Conclusion and discussion

In summary of the previous sections, I will show the general developments of *dus* and daarom (see Table 7.13 on *dus* and Table 7.14 on *daarom*).

Table 7.13. *Dus*: General profile in three periods

	13th century	**16th century**	**20th century**
Function	1. anaphor 2. causal connective	1. anaphor 2. intensifier 3. causal connective	1. causal connective 2. discourse marker
Preferred domain	epistemic domain	(no clear preference)	epistemic domain
Position	initial and medial	initial and medial	initial and medial
Categorical status	adverb	1. adverb 2. ambiguous adverb/complementizer	1. adverb 2. complementizer

Table 7.14. *Daarom*: General profile in three periods

	13th century	**16th century**	**20th century**
Function	1. anaphor 2. causal connective 3. relative adverb	1. anaphor 2. causal connective 3. relative adverb	1. causal connective (2. anaphor)[16]
Preferred domain	content domain	content domain	content domain
Position	initial and medial	initial and medial	initial and medial
Categorical status	adverb	adverb	adverb

How do the conceptual and syntactic changes of *dus* and *daarom* relate to the hypotheses in (43) (repeated from section 7.2)?

(43) <u>Hypotheses for the diachronic study:</u>

a. If *dus* and *daarom* express more than one function during a certain time span, these functions have different positioning preferences.
b. Once *dus* and *daarom* come to serve as connectives, different positionings are related to usage in different domains.
c. Once *dus* comes to serve as discourse marker signaling accessibility, its use as an accessibility marker differs in positioning preferences from its use as connective.

The hypothesis in (43)a is supported for *dus* in several ways. In the 13th century, anaphoric *dus* had a preference for the clause-medial position, whereas connective *dus* had a preference for the clause-initial position. This also holds for the 16th century, where the connective use is even restricted to the clause-initial position. In the 20th century, the discourse marker use is restricted to the clause-medial position, whereas the connective use can occur in both positions. In the 16th and the 20th centuries, the connective use deviates from the other

[16] The anaphoric use of *daarom* in the 20th century is placed between brackets because it was not attested in my corpus and because it is far more restricted than in the earlier centuries.

conceptual functions in that connective *dus* is not restricted to adverbial use, whereas both the anaphoric use and the discourse marker use are restricted in this respect.

The hypothesis in (43)a is also borne out for *daarom*. Positioning is used to discriminate between the connective function and the anaphoric function: in the 13th and 16th centuries, the connective function has a preference for the clause-initial position, whereas the non-causal anaphoric function has a preference for the clause-medial position. In the 20th century, only the connective function remains, which can then be expressed either in the clause-initial or in the clause-medial function.

The hypothesis in (43)b is not supported for *dus* or for *daarom*. The three domains do not have specific position preferences. However, for *dus* it can be stated that the speech-act use shows a preference in terms of categorical status: speech-act *dus* has a preference for the non-adverbial use.

The hypothesis in (43)c is only partially supported: the discourse marker use only differs in positioning preferences from connective fragments that do not contain accessible conclusions. In those 20th-century connective fragments in which *dus* marks the information as accessible, *dus* shows a behavior similar to *dus* in discourse marker fragments: it has a preference for the clause-medial position.

It can be concluded, then, that syntactic characteristics based on positioning only support one difference at the text-linguistic level at the time. In the case of *daarom* and *dus*, positioning is used to support the distinction between connective and non-connective use, and not to support the distinction based on domains of use. In the case of *dus*, positioning can also be related to the accessibility of the information in the *dus*-clause.

In the remainder of this section, two points will be taken up for discussion. The first is the rise of the complementizer use of *dus* (7.6.1); the second is the interaction between positioning and accessibility of the information in the *dus*-clause (7.6.2).

7.6.1 On the rise of the complementizer use of *dus*

The syntactic analysis of *dus* revealed that *dus* has developed its use as a complementizer only from the 16th century on. This makes it interesting to formulate some claims about the rise of this complementizer use.

From the data in my diachronic sample it can be concluded that the development of the complementizer use started around the 16th century. The sample from that century contained several imperatives in which *dus* could not unambiguously be analyzed as an adverb. In these cases the absence or presence of inversion could not be established because the subject was absent. For example, in (44), the finite verb *toont* 'show' can be said to occupy the first position within the imperative clause (so-called V1), which implies that *dus* should be interpreted as a complementizer. Alternatively, the finite verb can be said to be in second position (V2), with *dus* as a topicalized adverb in front of it.[17] Because the subject is absent, no choice between these syntactic analyses can be made.

(44) *En nu die sommige out en cout / sy hebben niet waer mee*
dat sij haer hongerige buyck / sullen versaeden
dus *<u>toont</u> aan haer u lieft* (Spel van sinnen, 1597)
'And now some old and cold ones, they don't have anything with which they can satisfy their hungry stomachs, so show your love to them.'

[17] Note that this second option is only available in Middle Dutch (cf. Van Gestel 1992: 46-47). In Modern Dutch imperatives the finite verb obligatorily occupies the first position.

The occurrence of these ambiguous cases can be traced back to the more flexible word order possibilities of the Middle Dutch language: "(…) as Gerritsen's (1982) study shows, Middle Dutch imperatives showed considerable fluctuation between verb-second and verb-first orders" (Burridge 1993: 231). Van Gestel et al. (1992: 46) also give some examples of imperatives in which one constituent occurs before the verb (see (45) and (46)).

(45) ***Nu*** *sit weder op u ors* (Van Gestel et al. 1992: 46)
'Now get on your horse again'
(46) ***Oec*** *weet: enheefstu die blader niet, so nem dat saet wel ghesoden in water*
(Van Gestel et al. 1992: 46)
'Also know: if you don't have the leaves, so take the seed drenched with water'

The word order within imperative clauses became less flexible: nowadays, imperatives always exhibit V1. Van Gestel et al. (1992: 47) speculate on the cause of the loss of the V2 word order in modern Dutch imperatives by suggesting that the loss of flection led to the fact that the imperative had to look for an alternative, syntactic way to distinguish itself from other verb forms.

My hypothesis, then, is that imperatives played in important role in the categorical reanalysis of *dus*. Because the possibility to front constituents in imperatives disappeared, a different syntactic categorization was needed for *dus* in such clause-initial positions. The positioning of *dus* could only be considered grammatical if the word was reinterpreted as a complementizer. Further research is needed to test this hypothesis about the role of imperatives; this research should focus on diachronic data from the centuries in between the 16th and 20th centuries.

7.6.2 On the rhetorical use of *dus*

In section 7.4.5 I introduced the significant interaction between accessibility and positioning. In this section I discuss two reasons why the mapping between accessibility and the clause-medial position is not one-to-one.

Firstly, there are exceptional fragments in which *dus* marks an accessible conclusion, but still appears in the clause-initial position. This is the case in fragments like (47), in which the conclusion is questioned. By placing *dus* in clause-initial position, speaker B does not present her conclusion as accessible, and hence leaves room for speaker A to disagree with her.

(47) A: *Ik ben er gisteren pas gekomen en ze vertrokken meteen naar hun buitenhuis.*
'I just arrived there yesterday and they went off to their country house immediately.'
B: *Dus je bivakkeert daar helemaal alleen?* (Au pair, 1992)
'So you're lodging there completely on your own?'

Note that this explanation for the "deviant" clause-initial position of *dus* is not the only candidate. An alternative explanation is that the positioning of *dus* here is due to the fact that the author describes an interaction between two persons, and that *dus* might be an attractive candidate to begin the next speaker turn with (cf. Schiffrin 1987 on turn takers). Such a hypothesis should be tested on the basis of spoken corpora.

Secondly, there are exceptional fragments in which *dus* marks a new conclusion, but in which it still occurs in the clause-medial position. It seems as if this clause-medial use of *dus* serves a rhetorical function (cf. Van den Hoven 1997 on *zodat* 'so that'). The speaker presents his position "to the reader [or hearer, JEV] as if it is a matter of common knowledge, which

no one in his right mind would question" (cf. Leech 1974: 61 on manipulative language in advertisements and political statements). For example, in the context of (48), the conclusion that "other municipalities who want to join in should also pay" is a new one, a conclusion of Tiemersma himself. By placing the connective *dus* in the clause-medial position, the speaker presents his conclusion as obvious, ruling out discussion beforehand. Something similar happens in (49). The clause-medial position of *dus* triggers the interpretation that ending up in a political debate is an inevitable situation. This interpretation is possible, but not necessary with *dus* in clause-initial position.

(48) *In de hal van de oliemaatschappij BP werken al vanaf september rond de 40 deelnemers aan het Jeugd Werk Garantieplan (JWG). Nu nog komen ze allen uit Hoogeveen, maar de roep om deelname uit de gemeenten Ruinen, Zuidwolde, Beilen, Westerbork en een deel van Oosterhesselen is te horen. 'Dan moeten die gemeenten* ***dus*** *wel méébetalen', weet ir. J. Tiemersma, directeur van de sociale werkplaats in Hoogeveen Howerco.*
(MC, 1995)
'From September on about 40 participants have been working in the hall of the oil company BP at the Plan of Youth Work Guarantee (YWG). Till now they all came from Hoogeveen, but the call for participation from the municipalities Ruinen, Zuidwolde, Beilen, Westerbork and a part of Oosterhesselen can be heard. "So then these municipalities will have to share in the costs," knows engineer J. Tiemersma, manager of the sheltered workshops in Hogeveen Howerco.'

(49) [About a proposal to give priority to employees with psychological problems]
Pijpstra toont zich geen voorstander van dit voorstel. 'Sociaal-psychologische hulp wordt betaald uit de ABWZ, waarvoor iedereen in Nederland premie betaald. Het zou niet zo eerlijk zijn mensen voor te trekken. Je raakt daarover ***dus*** *verzeild in een politieke discussie. Er bestaat immers kans op tweedeling van de maatschappij.'* (MC, 1995)
'Pijpstra is not an advocate of this proposal. "Socio-psychological help is paid out of the ABWZ, for which everyone in the Netherlands pays a contribution. It would not be so fair to give these people preferential treatment. Therefore you end up in a political debate. After all, there is a possibility of a dichotomy within the society."'

In terms of Ariel (1988: 590), then, *dus* "exploits its ability to function as an accessibility marker". As a result, the mapping between accessibility and positioning is not one-to-one.

From this chapter it can be concluded that the positioning of *dus* and *daarom* supports their intended function at the text-linguistic level. To be more specific, positioning is mainly used to support the distinction between connective and non-connective functions and not for conceptual distinctions within the connective function. The following chapter will focus on such conceptual distinctions within the connective function: it investigates the interaction between the syntactic properties of causal connective fragments and their conceptual characterization based on subjectivity.

APPENDICES TO CHAPTER 7

Appendix F – Logit analyses of the diachronic development of *dus*

Overview of the content of Appendix F:

General remark

Different numbers of words were needed to select 50 *dus*-fragments per period: 146606 words for the 13th century, 51516 words for the 16th century, and 95769 words for the 20th century.

F – 1. Logit analysis of the three conceptual functions of *dus* in relation to categorical status

I – Remarks

- In the logit analysis the 13th-century *aldus*-fragments were disregarded.
- In the 20th century one discourse marker fragment without a verb was disregarded.

II – Data: Distribution of *dus* in three periods (all conceptual functions)

		Connective	Anaphor	Discourse marker	Total
13th century	adverbial	7	14	0	21
	non-adverbial	0	0	0	0
16th century	adverbial	12	29	0	41
	non-adverbial	9	0	0	9
20th century	adverbial	29	0	6	35
	non-adverbial	14	0	0	14
Total		71	43	6	120

III – Results logit analysis *dus*

Logit (fr)	χ^2 model	df	p model	χ^2 factor	df	p factor
1. constant	270.27	17	< .001	-	-	-
+ 2. period	205.75	15	< .001	64.52	2	< .001
+ 3. conceptual function	142.50	13	< .001	63.25	2	< .001
+ 4. categorical status	94.45	12	< .001	48.05	1	< .001
+ 5. period x function	31.33	8	< .001	63.12	4	< .001
+ 6. function x cat. status	7.296	6	< .5	24.03	2	< .001

IV – Parameter estimates *dus* for model 6

Parameter	Estimate	s.e.	z-score	p
constant	-8.10	0.17	-46.75	< .001
period: 16th century	-0.10	0.27	-0.36	0.72
period: 13th century	-2.23	0.41	-5.50	< .001
function: anaphor	-4.99	2.24	-2.23	0.03
function: discourse marker	-1.61	0.45	-3.58	< .001
categorical status: non-adverbial	-0.73	0.25	-2.89	0.004
period x function: 16th anaphor	5.70	2.25	2.53	0.01
period x function: 16th discourse marker	-2.70	2.29	-1.18	0.24
period x function: 13th anaphor	6.06	2.28	2.65	0.01
period x function: 13th discourse marker	-1.62	2.31	-0.70	0.48
function x cat. status: non-adverbial anaphor	-4.24	1.85	-2.29	0.02
function x cat. status: non-adverbial disc. marker	-2.30	1.89	-1.22	0.22

F – 2. Logit analysis of the three conceptual functions of *dus* in relation to positioning

I – Remarks

- In the logit analysis the 13th-century *aldus*-fragments were disregarded.

II – Data: Distribution of *dus* in three periods (all conceptual functions)

	Connective		Anaphor		Discourse marker		Total
	Initial	Medial	Initial	Medial	Initial	Medial	
13th century	6	1	8	6	0	0	21
16th century	21	0	3	26	0	0	50
20th century	22	21	0	0	1	6	50
total	49	22	11	32	1	6	121

III – Results logit analysis *dus*

Logit (fr)	χ^2 model	df	p model	χ^2 factor	df	p factor
1. constant	246.99	17	< .001	-	-	-
+ 2. period	181.63	15	< .001	65.35	2	< .001
+ 3. conceptual function	121.58	13	< .001	60.05	2	< .001
+ 4. position	121.58	12	< .001	0.01	1	< .95
+ 5. period x function	57.13	8	< .001	64.45	4	< .001
+ 6. function x position	32.35	6	< .001	24.79	2	< .001
+ 7. period x position	7.01	4	< .25	25.34	2	< .001

IV – Parameter estimates *dus* for model 7

Parameter	Estimate	s.e.	z-score	p
constant	-9.96	0.38	-26.34	< .001
period: 16th century	2.10	0.44	4.79	< .001
period: 20th century	1.59	0.43	3.67	< .001
function: anaphor	0.01	0.53	0.02	0.98
function: discourse marker	-3.62	2.27	-1.59	0.12
position: clause-medial	-4.57	1.12	-4.09	< .001
period x function: 16th anaphor	-1.58	0.75	-2.12	0.03
period x function: 16th discourse marker	-1.36	3.21	-0.42	0.67
period x function: 20th anaphor	-9.23	2.50	-3.69	< .001
period x function: 20th discourse marker	0.39	2.47	0.16	0.87
period x position: 16th clause-medial	1.80	0.74	2.44	0.02
period x position: 20th clause-medial	4.50	1.15	3.90	< .001
function x position: clause-medial anaphor	4.57	1.01	4.54	< .001
function x position: clause-medial disc. marker	2.01	1.14	1.76	0.08

F – 3. Logit analysis of the three domains of use of *dus* in relation to categorical status

I – Remarks

None.

II – Data: Distribution of *dus* in three periods (only connective function)

		Content	Epistemic	Speech act	Total
13th century	adverbial	2	5	0	7
	non-adverbial	0	0	0	0
16th century	adverbial	9	3	0	12
	non-adverbial	0	0	9	9
20th century	adverbial	6	23	0	29
	non-adverbial	2	11	1	14
Total		19	42	10	71

III – Results logit analysis *dus*

Logit (fr)	χ^2 model	df	p model	χ^2 factor	df	p factor
1. constant	146.84	17	< .001	-	-	-
+ 2. period	95.10	15	< .001	51.74	2	< .001
+ 3. domain	73.03	13	< .001	22.07	2	< .001
+ 4. categorical status	64.28	12	< .001	8.75	1	< .005
+ 5. domain x cat. status	38.72	10	< .001	25.57	2	< .001
+ 6. period x domain	7.60	6	< .5	31.12	4	< .001

IV – Parameter estimates *dus* for model 6

Parameter	Estimate	s.e.	z-score	p
constant	-9.51	0.36	-26.20	< .001
period: 16th century	0.75	0.49	1.55	0.12
period: 13th century	-1.76	0.78	-2.27	0.02
domain: epistemic	1.26	0.41	3.06	0.002
domain: speech act	-5.41	2.05	-2.64	0.01
categorical status: non-adverbial	-2.05	0.72	-2.85	0.004
period x domain: 16th epistemic	-2.52	0.77	-3.29	0.001
period x domain: 16th speech act	1.98	1.12	1.77	0.08
period x domain: 13th epistemic	-0.56	0.91	-0.62	0.54
period x domain: 13th speech act	-0.37	2.54	-0.14	0.89
domain x cat. status: epistemic non-adverbial	1.03	0.80	1.29	0.20
domain x cat. status: speech act non-adverbial	5.56	1.97	2.83	0.01

F – 4. Logit analysis of the three domains of use of *dus* in relation to positioning

I – Remarks

None.

II – Data: Distribution of *dus* in three periods (only connective function)

	Content		Epistemic		Speech act		Total
	Initial	Medial	Initial	Medial	Initial	Medial	
13th century	1	1	5	0	0	0	7
16th century	9	0	3	0	9	0	21
20th century	7	1	14	20	1	0	43
Total	17	2	22	20	10	0	71

III – Results logit analysis *dus*

Logit (fr)	χ^2 model	df	p model	χ^2 factor	df	p factor
1. constant	144.79	17	< .001	-	-	-
+ 2. period	92.73	15	< .001	52.05	2	< .001
+ 3. domain	70.62	13	< .001	22.12	2	< .001
+ 4. position	60.59	12	< .001	10.03	1	< .005
+ 5. period x domain	29.48	8	< .001	31.11	4	< .001
+ 6. period x position	9.52	6	< .25	19.96	2	< .001

IV – Parameter estimates *dus* for model 6

Parameter	Estimate	s.e.	z-score	p
constant	-10.06	0.38	-26.22	< .001
period: 16th century	1.41	0.51	2.77	0.01
period: 13th century	-1.32	0.82	-1.61	0.17
domain: epistemic	1.45	0.39	3.68	< .001
domain: speech act	-1.98	1.02	-1.95	0.05
position: clause-medial	-0.04	0.31	-0.14	0.89
period x domain: 16th epistemic	-2.52	0.77	-3.29	0.001
period x domain: 16th speech act	1.98	1.12	1.77	0.08
period x domain: 13th epistemic	-0.51	0.92	-0.55	0.58
period x domain: 13th speech act	-0.32	2.55	-0.12	0.90
period x position: 16th clause-medial	-4.21	1.86	-2.26	0.02
period x position: 13th clause-medial	-1.58	1.04	-1.52	0.13

Appendix G – Logit analyses of the diachronic development of *daarom*

Overview of the content of Appendix G:

General remark

Different numbers of words were needed to select 50 *daarom*-fragments per period: 118943 words for the 13th century, 125975 words for the 16th century, and 232378 words for the 20th century.

G – 1. Logit analysis of the two conceptual functions of *daarom* in relation to positioning

I – Remarks

- In the logit analysis the cataphoric fragments have been disregarded (12 in the 13th century and 2 in the 16th century)
- In the logit analysis the anaphoric fragments in which *daarom* functions as a relative adverb (1 in the 13th century and 1 in the 16th century) have been disregarded.

II – Data: Distribution of *daarom* in three periods

	Connective		Anaphor		Total
	Initial	Medial	Initial	Medial	
13th century	29	3	2	4	38
16th century	34	9	1	4	48
20th century	34	16	0	0	50
Total	97	28	3	8	150

III – Results logit analysis *daarom*

Logit (fr)	χ^2 model	df	p model	χ^2 factor	df	p factor
1. constant	185.99	11	< .001	-	-	-
+ 2. period	179.66	9	< .001	6.33	2	< .05
+ 3. function	61.08	8	< .001	118.60	1	< .001
+ 4. position	31.24	7	< .001	29.84	1	< .001
+ 5. period x function	23.40	5	< .001	7.84	2	< .025
+ 6. function x position	7.19	4	< .25	16.21	1	< .001
+ 7. period x position	0.57	2	< .975	6.62	2	< .05

IV – Parameter estimates *daarom* for model 7

Parameter	Estimate	s.e.	z-score	p
constant	-8.32	0.19	-44.99	< .001
period: 16th century	0.03	0.25	0.12	0.90
period: 20th century	-0.51	0.25	-2.03	0.04
function: anaphor	-3.49	0.95	-3.66	< .001
position: clause-medial	-2.31	0.59	-3.89	< .001
period x function: 16th anaphor	-1.24	0.89	-1.39	0.17
period x function: 20th anaphor	-4.77	2.36	-2.03	0.04
period x position: 16th medial	1.01	0.69	1.47	0.14
period x position: 20th medial	1.54	0.67	2.32	0.02
function x position: anaphor medial	3.84	1.09	3.54	< .001

G – 2. Logit analysis of the three domains of *daarom* in relation to positioning

I – Remarks

- In the logit analysis the cataphoric connective fragments as well as the ambiguous fragments were disregarded.

II – Data: Distribution of *daarom* in three periods (only connective function)

	Content		Epistemic		Speech act		Total
	Initial	Medial	Initial	Medial	Initial	Medial	
13th century	15	1	8	0	5	0	29
16th century	17	2	8	0	8	0	35
20th century	28	14	6	2	0	0	50
Total	60	17	22	2	13	0	114

III – Results logit analysis *daarom*

Logit (fr)	χ^2 model	df	p model	χ^2 factor	df	p factor
1. constant	148.86	17	< .001	-	-	-
+ 2. period	148.02	15	< .001	0.84	2	< .975
+ 3. domain	90.24	13	< .001	57.78	2	< .001
+ 4. position	35.91	12	< .001	54.33	1	< .001
+ 5. period x domain	17.32	8	< .05	18.59	4	< .001
+ 6. period x position	2.23	6	< .9	15.08	2	< .001

IV – Parameter estimates *daarom* for model 6

Parameter	Estimate	s.e.	z-score	p
constant	-8.96	0.25	-35.41	< .001
period: 16th century	0.03	0.34	0.09	0.93
period: 20th century	-0.05	0.31	-0.16	0.87
domain: epistemic	-0.68	0.43	-1.58	0.11
domain: speech act	-1.14	0.51	-2.25	0.02
position: clause-medial	-3.15	0.93	-3.38	< .001
period x domain: 16th epistemic	-0.17	0.60	-0.29	0.77
period x domain: 16th speech act	0.29	0.66	0.44	0.66
period x domain: 20th epistemic	-0.98	0.58	-1.69	0.09
period x domain: 20th speech act	-4.20	2.30	-1.83	0.07
period x position: 16th medial	0.44	1.16	0.38	0.70
period x position: 20th medial	2.40	0.98	2.45	0.01

Chapter 8

Subjectification in the diachronics of the causal connectives

With the diachronic data mentioned in the previous chapters, it is relatively easy to gain insight into a related discussion, the occurrence of subjectification in connective development and its relation to syntactic changes. In this chapter I discuss the diachronic development of the four causal connectives *daarom*, *dus*, *omdat*, and *want* from a subjectification perspective. In doing so, I have three aims: a) to provide objectifiable and quantifiable operationalizations of the notion of 'subjectivity' in relation to causal connectives; b) to find out whether subjectification is involved in the diachronic conceptual developments; c) to investigate whether there is a relation with syntactic changes.

> *"It [= subjectification] is a recurrent and highly important type of semantic extension and is often a central factor in the evolution from 'lexical' to 'grammatical' elements."* (Langacker 1990: 5)

8.1 Introduction[1]

The diachronic developments of the connectives discussed in the three previous chapters illustrate three types of conceptual change connectives might go through; a) the actual rise of connectives, b) changes within the connective function, and c) shifts away from the connective function. An example of the first change, the actual rise of the connective function, can be found in the change from anaphoric *dus* 'thus/this way' to connective *dus* 'so'. The second change, within the connective function, is exemplified by *want*, which has lost its ability to express purely temporal coherence relations. Finally, the shift away from the connective function is characteristic of the change from causal connective *dus* 'so' into a discourse marker use in which *dus* functions as a marker of information status.

These three types of development can be placed in a grammaticalization framework, a theory that starts from the idea that lexical elements can develop into functional elements and that existing functional elements can develop new grammatical functions (see, among others, Heine, Claudi & Hünnemeyer 1991: 4; Hopper & Traugott 1993: 94; Bybee, Perkins & Pagliuca 1994: 13). Several authors have claimed that grammaticalization often goes hand in hand with a specific change in meaning: so-called *subjectification*, an increase in subjectivity (cf. Langacker 1990: 5; Traugott 1995: 31).[2] This subjectification hypothesis has also been

[1] I would like to thank Ninke Stukker for making her diachronic corpus of *dus* and *daarom* fragments available to me. Part of this study is the result of joined work with her, published earlier in Evers-Vermeul & Stukker (2003). The two operationalizations of the notion 'subjectification' as mentioned in section 8.2 are the outcome of our collaboration. Furthermore, Stukker performed the SOC-analysis of *dus* and both subjectivity analyses of *daarom*.

[2] Verhagen (2000) refines Traugott's claim by subdividing *subjectivity* into two types: *speaker/hearer subjectivity* and *character subjectivity*. According to him, there is an increase in speaker/hearer subjectivity in the grammaticalization process, but there is also a decrease in character subjectivity. In the current chapter, Traugott's term *subjectivity* is restricted to Verhagen's *speaker/hearer subjectivity*.

applied to the diachronic development of connectives (see section 3.6 in Chapter 3). For example, Traugott (1995) has shown that a process of subjectification took place in the diachronic change of English *while*, and she (1995: 39) even claims: "historically almost all grammatical markers of clause combining have developed out of a more 'objective' function".

Looking at the connective data investigated so far, the subjectification hypothesis indeed seems to hold for the three types of conceptual change connectives might go through. However, the exact range of the subjectification hypothesis in relation to the diachronic development of connectives is not clear yet. Does it apply to all types of connectives? And does it hold for different uses within a specific connective function? For example, Sweetser (1990) has claimed that linguistic expressions that show a multiple-domain usage get their original meaning in the content-domain and that only later on is this meaning extended to the other domains. In other words: "there is a tendency to borrow concepts and vocabulary from the more accessible physical and social world to refer to the less accessible worlds of reasoning, emotion and conversational structure" (Sweetser 1990: 31). Does the subjectification hypothesis even hold for subtle changes within the use as a causal connective? And finally, how does subjectification relate to syntactic changes (e.g. changes in word order or categorical status)? For example, Keller (1995) claims that the change of German *weil*, which involves both a conceptual and a syntactic component, can be characterized as a case of subjectification. These questions were used as a starting point for the investigation in this chapter.

The current chapter presents an analysis of the historical development of the Dutch causals *want*, *omdat*, *dus*, and *daarom* based on subjectification. For each of the four connectives, two questions will be answered:

(1) <u>Research questions of this chapter:</u>
   a. Do the diachronic conceptual changes involve so-called *subjectification*?
   b. Are syntactic changes a necessary prerequisite for the occurrence of subjectification?

Although the main focus here will be on changes within the use as causal connective, my analyses also pay attention to the other two types of conceptual change: the rise of the connective usage, and shifts away from this clause combining function.

In order to operationalize the term *subjectivity* in an objective and quantifiable way, I will use Sweetser's domains of use as an analytical instrument (see also the conceptual analyses in Chapters 6 and 7). However, some researchers have recently rejected Sweetser's model as a means of accounting for the distribution and interpretation of specific causal markers (cf. Pander Maat & Sanders 2000; Pander Maat & Degand 2001; Pit 2003) and have argued instead in favor of subjectivity. For example, Pander Maat & Degand (2001) put forward a subjectivity scale, one of *speaker involvement*. Given this theoretical development, the current chapter will use this recent "speaker involvement" approach as a second analytical instrument. This chapter may shed some light on the competition between the two approaches, since it allows for a comparison between them.

The organization of this chapter is as follows. First, I discuss the two analytical instruments used for establishing the relative degree of subjectivity of connective fragments. The first operationalization of the notion 'subjectivity' can be related to the "speaker involvement" approach (see section 8.2). It concentrates on the linguistic realization of specific subjective elements within the causally related propositions. The second method (see section 8.3) is in line with the domains approach; it measures subjectivity at the conceptual level, focusing on the causal relation as a whole. Following these operationalizations, the

sample and methodology in this study are accounted for (8.4), after which the results from the corpus study are presented (8.5) and discussed (8.6).

8.2 Measuring subjectivity part 1: SOC types

Subjectivity "implies some degree of integration of the perceiver in the description of an object or a process" (Cuenca 1997: 5). This integration can appear indirectly from the presence of all kinds of linguistic elements that signal subjectivity (J. Sanders & Spooren 1997: 91; Pit 2003). Some typical examples are evaluative expressions (*unfortunately*) and modal elements (adverbs such as *probably* and *perhaps* or verbs (*must*) that express likelihood judgments). What all these factors have in common is that they all express to which degree an utterance is on the speaker's account.

Integration of the speaker into a connective expression can also appear directly from the fact that he is linguistically realized in one of the causally related propositions. For example, the speaker is visibly integrated if he mentions himself as the *I*-subject who is responsible for the construction of the causal relation. Such a fragment is very subjective. My first operationalization focuses on this direct form of integration: I analyze whether the speaker is also the one who is responsible for the construction of the causal relation. Section 8.2.1 explains the relation between the speaker and the 'constructor of the causal relation'; section 8.2.2 treats the relative degree of subjectivity of the different linguistic realizations of this constructor.

8.2.1 Responsibility and subjectivity

In the concept of subjectivity, animacy is crucial, since only causal relations that originate in some mind can contain subjective elements. Pander Maat & Sanders (2000, 2001) call the animate being that constructs the causal relation the 'Subject of Consciousness' or SOC (see (2)). They use this concept to distinguish between objective and subjective causal relations.

(2) <u>Subject of Consciousness or SOC:</u>
person who is responsible for constructing the causal relation

This SOC can be the speaker (like the *I*-figure in (3)), but it might as well be another person who is responsible (like the *he*-subject in (4)).

(3) *Als u niet antwoordt arresteer* ***ik*** *u omdat u weigert informatie te geven.*
(Verborgen gebreken, 1996)
'If you don't answer I will arrest you because you refuse to give information.'

(4) *(…)* ***hi*** *sal hem doen hangen of quader doot doen sterven om dat hise verraden heeft.*
(Historie vanden vier heemskinderen, 1508)
'(…) he will hang him or give him some other evil death because he has betrayed them.'

Note that in both connective fragments it is the speaker who is responsible for reporting the causal relation. My search for differences in subjectivity, then, does not focus on reporting the causal relation in a text, but on constructing the reported causal relation in reality. To illustrate this point, (5) gives an example in which the speaker is both responsible for uttering the causal relation and for the construction of the causality in reality. In (5) the speaker (included in the *we*-subject) gives a motivation for the fact that she uses the word *moeder* 'mother' to talk about her mother. She justifies her own act in reality and thus constructs the causal relation herself. The speaker here is also the SOC. In (6), on the other hand, the speaker only

reports about an event in which he is not at all involved and in which he consequently cannot be held responsible for the real-world causality. In other words: the speaker is not the SOC.

(5) *Moeder als naam voor je moeder klonk mij vreemd in de oren. Wij gebruikten het wel om over onze moeder te praten, want een moeder was wat ze was, het was een soort beroep.* (De vriendschap, 1995)
'Mother as a name for your mother sounded strange to me. We did use it when talking about our mother, because she was a mother, it was a kind of profession.'

(6) *Eind vorig jaar is een zending met lantaarns, laarzen, breiwol en weefgetouwen niet in Pakrac aangekomen, omdat het vervoer niet werd toegelaten.* (MC, 1995)
'At the end of last year a supply of lanterns, boots, knitting wool and looms did not arrive in Pakrac, because the transport was refused entry.'

The identification of an SOC allows me to tell whether or not a causal relation is subjective. But in order to answer my research question, I need to establish the relative degree of subjectivity. Remember that subjectivity has been defined as the degree of speaker integration in the causal relation. Following Pander Maat & Sanders (2000: 68), the relative degree of subjectivity is defined here as the relative distance between the speaker expressing the causal relation and the SOC presented as being responsible for that causal relation.

(7) <u>Degree of subjectivity based on SOC type:</u>
distance between SOC and the present speaker

The smaller the distance between SOC and speaker, the more subjective the relation is. Conversely, with an increasing distance between speaker and SOC, the degree of integration of the speaker into the causal relation decreases. For example, in fragments (3) and (5) the SOC coincides with the speaker, resulting in relatively subjective connective fragments. However, in (4) the speaker and the *he*-subject (the SOC in this fragment) are not the same person. Hence, the distance between speaker and SOC is larger than in (3) and (5). The speaker reports a causal event from a certain distance. Hence, the degree of subjectivity is smaller. In the next section, these differences in degree of subjectivity will be discussed in more detail.

8.2.2 SOC types and their degree of subjectivity

The distance between speaker and SOC is dependent on two factors. The first is the presence or absence of an SOC and accordingly the possibility for the speaker to identify with this SOC. The second factor is the way in which a present SOC is linguistically realized. Table 8.1 gives an overview of the categories I distinguish in my SOC analysis. The SOC types in this table are ordered by decreasing subjectivity. Four of them are adopted from the classification in Pander Maat & Sanders (2000: 68-69); Ninke Stukker and I added category 5 and the distinction between 3 and 4 in order to establish a gradual scale of subjectivity (cf. also Pander Maat & Degand's (2001) scale of speaker involvement). This gradualness can be important to trace subtle changes in subjectivity. In the remainder of this section I will clarify each SOC type and give an account for its relative degree of subjectivity.

Table 8.1. Overview of SOC types, ordered by decreasing subjectivity

SOC type	**Examples of linguistic realization**
1. implicit speaker	1. Ø
2. explicit speaker (including cited speakers)	2. *I*, *we*, *you and I*, generic *you*
3. 2nd or 3rd person pronominal	3. *he*, *she*, *they*, *you*
4. 3rd person nominal	4. *the man*, *Clara*
5. unspecified	5. *one*, passive voice
6. no SOC	6. does not apply

A causal relation is maximally objective when there is no SOC (SOC-type 6). Both in the constructed example in (8)[3] and in the corpus fragment in (9) there is no conscious person who is responsible for the construction of the causal relation. In both fragments, the speaker is responsible for uttering the text, but not for the causal process described in it. In fragment (9) neither the speaker nor the subject in the consequent (the sculptures of the idols) can be labeled as the one responsible for the construction of the causal relation. The sculptures only "experience" the causal effect, and do not function as an agent. This kind of causal relations can be labeled totally objective.[4]

(8) The temperature rose quickly because the sun was shining.
(9) *Ioseph ende Maria van daer traken / In Egipten / daer die afgoden al braken.*
Om dat si Iesum moesten laten/ Behouden die ste hy is van
(Devoot ende profitelyck boecxken, 1539)
'Joseph and Mary traveled to Egypt, where the idols broke completely, because they had to leave Jesus alone together with the place he was in'

The following two utterances give examples of the other – maximally subjective – end of the scale (SOC-type 1). Here the speaker coincides with the SOC. The ∅-symbol marks that the speaker has remained implicit.

(10) The temperature is probably going to rise, ∅ because the sun is shining.
(11) *Een duidelijke politieke lijn ontbrak en mede daardoor hebben de onderhandelingen tot een fatale oorlog geleid. ∅ Het is dus een onmogelijke opgave om duidelijk vast te stellen wat de regering en het Nederlandse volk nu eigenlijk wilden.* (MC, 1995)
'A clear political line was missing and also because of that the negotiations led to a fatal war. ∅ It is therefore an impossible task to clearly establish what the government and the Dutch people actually wanted.'

The speaker is the one responsible for the reasoning process, constructing the causal relation by connecting a conclusion with an argument (or the other way around as in (11)). In these kinds of relations the speaker is maximally integrated in the causal process.

[3] The examples in (8), (10) and (12) are constructed by Ninke Stukker.
[4] It is possible that a fragment does not have an SOC, although it contains an *I*-subject. This is the case in (i) (thanks to Wilbert Spooren for bringing up this example). Here the *I*-figure himself does not cause his own injuries. These might, for example, be due to the fact that his car has slipped because of ice and then crashed against a tree).
(i) *Ik raakte gewond doordat ik mijn gordel niet droeg.*
'I got injured because I didn't wear my safety belt.'

SOC-type 2 is a second SOC type with the speaker as the person who is responsible for the causal relation. The crucial difference between cases like (10) and (11) on the one hand and (12) and (13) on the other, is that the speaker remains implicit in the former cases, whereas he is linguistically present in the latter cases.

(12) **I** think the temperature is going to rise, because the sun is shining.
(13) *Zonder onderling contact en lering droogt de bron uit, en daarom weet **ik** dat nogal wat preken opgestuurd worden vanuit Nederland.* (Het verbroken zegel, 1991)
'Without mutual contact and teaching a well runs dry, and that's why **I** know that quite some sermons were sent from the Netherlands.'

Just as in SOC-type 1, SOC and speaker are one and the same person here, but the distance between SOC and speaker has increased somewhat. By mentioning himself explicitly as an actor within the causal relation, the speaker creates a certain distance between his role of communicator and his role of actor within the communicated causal relation. This makes the degree of subjectivity in the explicit speaker SOC smaller than in the implicit variant. Langacker (1990: 316) clarifies the explicit-implicit distinction by comparing it to looking *at* glasses versus looking *through* glasses. Glasses that are taken off only function as the object of perception, and not as part of the perceptual apparatus itself. By contrast, when glasses are worn, they fade from the observer's awareness despite their role in determining the nature of the perceptual experience. In that case they are part of the subject of perception – they are a component of the perceiving apparatus, but they are not themselves perceived. Similarly, when SOC's are not explicitly mentioned, we conceive of the causal relationship through their eyes; this results in a maximally subjective relation.

SOC-type 3 (2nd or 3rd person pronominal) and SOC-type 4 (3rd person nominal) both have a lower degree of subjectivity than SOC-type 2. In the case of 2nd or 3rd person SOC's, there is a larger distance between speaker and SOC, because the SOC is not the speaker, but someone else. In examples (14) the SOC *hi* 'he' is about to perform an act, motivated by the situation in the second segment of the relation. In (15) the SOC *patrouilles* 'patrols' is the actor in the clause expressing the consequent. This SOC performs an act in order to reach what is mentioned in the first clause. In both examples, the 3rd person actor is responsible for the causal relation in reality.

(14) *(…) **hi** sal hem doen hangen of quader doot doen sterven om dat hise verraden heeft.*
(Historie vanden vier heemskinderen, 1508)
'(…) **he** will hang him or give him some other evil death because he has betrayed them.'
(15) *Het project 'Watertoerisme 1995' moet de veiligheid en de leefbaarheid op en aan het water bevorderen en moet de overlast door het watertoerisme zoveel mogelijk tegengaan. Daarom scheren er regelmatig **patrouilles** kris-kras door het merenlandschap.*
(MC, 1995)
'The project 'Water tourism 1995' has to promote the security and livability on and near the water and prevent inconvenience due to water tourism as much as possible. That's why **patrols** regularly skim criss-cross over the lake landscape.'

It can be argued that the SOC's *he* and *patrols* differ in their degree of subjectivity, although both are 3rd person SOC's. This difference can be established by looking at the degree to which the speaker identifies himself with the actor in the reported situation: the speaker can adopt the perspective of the 3rd person actor (cf. Pander Maat & Sanders 2001: 256). This type

of identification is possible in both examples, albeit to different degrees. In example (14) the speaker can identify with the pronominal *hi* 'he'. Such pronominal expressions are used in situations in which the antecedent has already been introduced and thus can be regarded as 'known'. These pronominals lend themselves more easily for identification than relative 'unknown' nominal expressions such as *patrouilles* 'patrols' in (15) (cf. Kuno 1987: 204-206 and Ariel 1990).[5] The distance between speaker and pronominal SOC's (SOC-type 3) is smaller than the distance between speaker and nominal SOC's (SOC-type 4), resulting in a gradual decrease in subjectivity between examples (14) and (15).

SOC-type 5 shows the largest distance between speaker and SOC. This SOC type can be found in fragments in which the SOC has been pushed to the background due to the formulation, as in (16). In (16), the α-symbol marks that the SOC, the persons who take action to get hold of some money, has not been realized linguistically. Still, this SOC is conceptually present in the passive construction *werden ondernomen* 'were undertaken'. Passives, but also cases in which the SOC does not refer to a specific person (e.g. *men* 'one') are grouped together within this 'unspecified' SOC type.

(16) *(...) veel van de kinderen die lid waren van de GENK kwamen uit gezinnen waar men het moest doen met een minimumloon. Daarom werden **α** er geregeld acties ondernomen om aan geld te komen.* (De kunstrijder, 1989)
'Many children who were members of GENK came from families that had to live on minimum wages. That's why actions to collect money ***α*** were undertaken regularly.'

The six SOC types differ in their accompanying distance between speaker and SOC. These SOC types can be used to establish the degree of subjectivity of connective fragments from different periods, providing an objectifiable way to test the subjectification theory put forward at the beginning of this chapter. In line with this subjectification theory, which predicts an increase in subjectivity, the following subjectification hypothesis can be formulated.

(17) <u>Subjectification hypothesis based on SOC type:</u>
If a causal connective shows a change at the conceptual level, I expect:
a. an increase in subjectivity;
b. this increase to show from the fact that in later centuries this connective is more often combined with higher ranked SOC types (i.e. with SOC types with a relatively smaller distance between speaker and SOC) than in earlier centuries.

8.3 Measuring subjectivity part 2: domains of use

The operationalization of the term *subjectivity* in section 8.2 is related to the "speaker involvement" approach. The current section retakes the domains approach as a second method for establishing the degree of subjectivity. Analyzing connective fragments with both methods may shed some light on the competition between the two approaches.

---

[5] Kuno (1987: 204-206) refers to this form of identification with the notion 'empathy': "Empathy is the speaker's identification, which may vary in degree, with a person/thing that participates in the event or state that he describes in a sentence" (p. 206). Kuno uses an empathy hierarchy to describe the distances between speaker and different persons with whom the speaker might identify. According to the hierarchy of personal pronouns a pronominal can count on a greater empathy (that is a higher degree of identification) than a nominal expression. This view is in line with my claim that (14) is more subjective than (15).

For my second operationalization of the notion 'subjectivity' I use the extended version of Sweetser's (1990) theory on domains of use. In this connective analysis I do not focus on the linguistic realization of specific subjective elements within (one of) the conjuncts, but on the conceptual relation between the conjuncts as a whole. I distinguish between *non-volitional*, *volitional*, *epistemic*, and *speech act* relations, as illustrated by the constructed a-examples and the corpus-based b-sentences in (18)-(21) below (see also Chapter 2 for a more elaborate discussion of these domains). The subdivision within the content-domain allows me to measure subjectivity in a more subtle way.

(18) Content non-volitional
a. The temperature rose quickly because the sun was shining.
b. *Nv benic worden cranc ende out*
Dar omme *si mi nu begeuen / Die sinne die mi daden leuen*
(Sente Lutgart, 1263-1280)
'Now I have become ill and old
That's why the senses that made me live have left me'

(19) Content volitional
a. We went out in the garden because the sun was shining.
b. ***Omdat*** *de pony de fietser bleef nalopen*
bracht deze hem maar onder bij de dieren in het park. (MC, 1995)
'Because the pony kept on following the cyclist
he brought him to the animals in the park.'

(20) Epistemic
a. The temperature is probably going to rise, because the sun is shining.
b. *De sloffen maakten een wrang geluid, maar ze waren pure noodzaak,* ***want*** *alleen dankzij die sloffen had ik het gevoel enige grip te hebben op de vloer.* (De vriendschap, 1995)
'The slippers made a dry sound, but they were pure necessity, because only thanks to those slippers did I have the feeling of having some grip on the floor.'

(21) Speech act
a. Let's have dinner in the garden, because the sun is shining.
b. *Kun je knollen rapen,* ***want*** *ander werk is er niet.* (MC, 1995)
'Can you gather turnips, because there is no other work.'

If subjectivity "implies some degree of integration of the perceiver in the description of an object or a process" (Cuenca 1997: 5), then it can be argued that the different domains can be used as a way to measure the degree of subjectivity of causal (and other) connective fragments. A speaker constructing a content relation gives a description of facts that can be established objectively in reality. A non-volitional content relation like (18) arises without influence of animate beings and thus can be labeled the least subjective. The causality in a volitional content relation (cf. (19)) is due to an actor who takes the situation in the cause segment as a reason to perform the act mentioned in the consequent. The speaker implicitly becomes involved in a volitional relation: in order to report the cause as a valid reason to perform the act, he (temporarily) has to adopt the viewpoint of the actor. A volitional relation can therefore be regarded as more subjective than a non-volitional one, although both relations can be observed objectively in the real world. Even more subjective are epistemic relations (see (20)), in which the consequent is not a state of affairs in reality, but a mental state of the protagonist. The causal relation as a whole is not objectively observable and the speaker who presents the relation has to adopt the perspective of the protagonist. Maximally

subjective are speech-act relations like (21), since they do not concern a reality outside the speech event, but the structure of the current discourse (cf. Pander Maat & Degand 2001: 216-228). This structure is the full responsibility of the speaker. To conclude, speech-act relations show the highest ranking in terms of subjectivity, whereas non-volitional content relations are ranked at the lowest end.

The extended version of Sweetser's domains theory allows me to formulate a second subjectification hypothesis for the diachronic study.

(22) Subjectification hypothesis based on domains:
If a causal connective shows a change at the conceptual level, I expect:
a. an increase in subjectivity;
b. this increase to show from the fact that in later centuries this connective is more often combined with higher ranked (i.e. more subjective) domains than in earlier centuries.

8.4 Methodology

The two complementary methods used to establish the relative degree of subjectivity of connective fragments enable the investigation of the subjectification hypothesis: does this hypothesis hold for diachronic changes within the connective function? To perform the subjectivity analysis, I make use of the same connective fragments as mentioned in the three previous chapters. In this section, I discuss the procedure followed during the subjectivity analyses and the way changes are tested statistically.

Per connective a total number of 150 fragments has been selected (50 per period). Chapter 5 and 7 showed that in the early centuries a combined variant of the connectives *omdat* and *daarom* existed: *daar-om-dat*.

(23) *Die rede dar wi sunde vomb duon. dat is*
dar vombe dat *wi lange wanen leuen.* (NM, 1270-1290)
'The reason why we commit sins, that is because we suppose to live long.'

On the basis of spelling alone it cannot be decided whether this is a variant of *daarom* or of *omdat*. However, the interpretation of this utterance does allow me to make a choice. The words *daar* 'there' and *om* 'for/to' refer forward to the cause mentioned in the clause "that we suppose to live long", which means that *daar* has a cataphoric function (see also section 7.5.1 and 7.5.2 in the previous chapter). In contrast, the separate form *daarom* has an anaphoric function, referring backward. It can be concluded that the combination *daar-om dat* at the conceptual level functions like *omdat*. Therefore, I decided to only incorporate this combination into the *omdat*-samples (5 cases in the 13th century). For the *daarom*-corpus, new fragments were selected to maintain the number of 50 fragments per period.

I performed two conceptual analyses for each fragment, one in terms of SOC type and one in terms of domains of use. Using the first operationalization (see section 8.2), the degree of subjectivity was measured by looking at the distance between the 'Subject of Consciousness' (SOC) and the speaker in the causal relation. More specifically, I have looked at the lexical realization of the SOC, making a distinction between six categories (ordered here by decreasing subjectivity): *implicit speaker*, *explicit speaker*, *2nd or 3rd person pronominal*, *3rd person nominal*, *unspecified SOC*, and *no SOC*. For each fragment, I determined whether there was an SOC, and if so, how it was realized linguistically in the consequent conjunct of the causal fragment. The SOC type could only be established in those fragments in which *dus*, *daarom*, *want*, and *omdat* functioned as a causal connective (since

this operationalization deals with the person who is responsible for the construction of the causal relation). Fragments in which these words occurred in a different function (like the temporal use of *want*) were disregarded in the SOC analysis.

My analysis based on domains of use distinguishes between four categories of causal relations (ordered by decreasing subjectivity): *speech act*, *epistemic*, *content volitional*, and *content non-volitional*. To establish the domain type objectively, I have used an adapted version of the 'Basic operation paraphrase test' (Sanders 1997), as has already been discussed in Chapter 4. An overview of the paraphrases is given in Table 8.2.

Table 8.2. Overview of the paraphrases used in the domains analysis

Domain	**Paraphrase**
1. Speech act	1a. Situation P causes SOC's speech act Q
	b. SOC's speech act Q is caused by situation P
2. Epistemic	2a. Situation P causes SOC's conclusion Q
	b. SOC's conclusion Q is caused by situation P
3. Content volitional	3a. Situation P causes SOC's act Q
	b. SOC's act Q is caused by situation P
4. Content non-volitional	4a. Situation P causes situation Q
	b. Situation Q is caused by situation P
5. Other use	None of the paraphrases applies

The two tools to establish the relative degree of subjectivity were tested in a pilot study with four researchers, which led to highly similar judgments. In cases of divergent judgments in the current study, additional criteria were invoked or a discussion between the two analysts followed until consensus was reached (see section 4.3.3 in Chapter 4).

On the basis of the protocol mentioned above, I determined the distribution of the connective fragments over the different SOC and domain types. This distribution was then tested statistically with a logit analysis. In general, the significance or insignifance of changes can be taken as direct evidence in favor of, or against a hypothesis. However, in the current study both the presence and the absence of significant changes have been treated with more caution. Three additional factors were taken into account in the interpretation of the changes. First of all, a significant diachronic increase in subjective contexts was only considered a case of subjectification if it was accompanied by a decrease in the more objective contexts.[6] Introducing this accompanying decrease as an extra criterion was necessary because it is possible to find a significant increase in a relatively subjective category (like the epistemic domain) caused by a decrease in an even more subjective category (like the speech-act domain).

Secondly, diachronic increases or decreases in subjectivity were only taken as decisive evidence if they were lasting and not temporary. For example, if a connective showed a significant decrease in very subjective use only in the 16th century, but not in the 20th century, this was not taken as counterevidence to the subjectification hypothesis.

Finally, subjectification is not an obligatory phenomenon in diachronic development. As Traugott (2001: 7) formulates it herself: "change does not have to occur". This implies that it is not possible to draw conclusions on the validity of the subjectification hypothesis from the

[6] It was not necessary to find a significant decrease in one of the objective categories, since the increase in more subjective use can also result from insignificant decreases in several of the more objective categories.

absence of subjectification trends. Therefore, I will only reject the subjectification hypothesis if I find tendencies in the reverse direction: 'objectification'. Objectification is defined here as a diachronic pattern in which a connective shows a significant *increase* in objective contexts in combination with a *decrease* in the more subjective contexts.

8.5 Results

What does the corpus study reveal with respect to the questions asked at the beginning of this chapter? Before I turn to the subjectivity results of changes within the connective function, I present a more general subjectivity analysis of the diachronic developments of *dus* and *daarom*, the two words that also serve non-connective functions (see Chapter 7). In section 8.5.1 I discuss the different degrees of subjectivity of the anaphoric and the connective function of *daarom* and *dus*, as well as the discourse marker function of *dus*. Section 8.5.2 (on SOC types) and 8.5.3 (on domains) focus on subjectification within the connective function of *want*, *omdat*, *dus* and *daarom*. Section 8.5.4 investigates the relation between subjectification and changes in categorical status. Details on the statistical analyses mentioned in these sections can be found in Appendix H (on SOC types) and I (on domains).

8.5.1 Subjectification across functions?

As the previous chapter already indicated, *dus* and *daarom* appear in conceptual functions other than that of causal connective (see sections 7.4.1 and 7.5.1). Both words occur as anaphors; in addition, *dus* can be used as a discourse marker signaling information status. In this section I present an analysis of their cross-functional changes in terms of subjectivity.

Assuming that the connective function developed out of the anaphoric function, it can be stated that both *dus* and *daarom* show subjectification in their diachronic development. In the 13th and 16th centuries, *dus* and *daarom* could occur as anaphoric expressions. An example of the anaphoric use of *dus* is given in (24).

(24) *Doe seide thobias de goede. / Suchtende ende weenende sere (…)*
Dat mijn gheest ontfanghen si. / Met goeden vrede biddic di.
Mi es beter die dod dan dat lijf.
***Dus** bat die oude omdat sijn wijf. / Sulke ouertale vp hem sprac.* (Rijmbijbel, 1275-1300)
'Then Tobias the good said, / sighing and crying much (…),
"That my spirit will be received with good peace, I pray to you.
It is better for me to die than to live."
Thus prayed the old [man] because his wife spoke such coarse language over him.'

The anaphoric function is more objective than the connective function. In the anaphoric use, both *dus* and *daarom* have a deictic function: they refer indirectly to objectively observable phenomena in reality. In this use, the attitudes or beliefs of the speaker (or writer) do not play any role. Therefore, the anaphoric use of *dus* and *daarom* can be considered as more objective than the clause combining use. In the 20th century, the anaphoric uses of *dus* and *daarom* have disappeared.[7] This results in an increase in the more subjective connective use of these words. Hence, it can be concluded that *dus* and *daarom* show subjectification tendencies across different conceptual functions.

[7] The anaphoric use of *daarom* is still grammatical, but it occurs less frequently. The anaphoric use of *dus* has been transferred to the more specialized word *aldus*.

For *dus*, another case of subjectification across functions can be found. As I discussed in the previous chapter (see section 7.4.1), 20th-century *dus* shows up in the function of discourse marker. The text in (25), taken from a newspaper interview with the winner of a cooking contest, gives a nice illustration of this discourse marker use. In this fragment *dus* neither expresses a causal relation, nor is it used as an anaphoric expression. Its function is best described as signaling informational status: the information presented in the *dus*-clause is marked as already familiar to the listener.

(25) *Het ging erom wie het mooiste, het beste produkt maakte. Dat was ik* ***dus****.* (MC, 1995)
'The point was who could make the most beautiful, the best product. And that was me.'

In its use as a discourse marker, *dus* signals the knowledge state of the speaker (or writer), who considers the information presented as accessible to the hearer (or reader). Therefore, I consider this function as more subjective than the connective use of *dus* (cf. Traugott & Dasher 2002 for a similar line of reasoning on the Japanese equivalent of *dus*: *sate*). The significant increase of this use in the 20th century can thus be considered as a case of subjectification across functional categories.

It can be concluded that both *dus* and *daarom* show subjectification in their diachronic cross-functional development. In the following two sections, I will focus on subjectification within the connective use of *want*, *omdat*, *daarom* and *dus*. Section 8.5.2 deals with subjectification based on SOC type: section 8.5.3 discusses connective subjectification based on domains of use.

8.5.2 Subjectification based on SOC type?

Figure 8.1 shows the distribution of *want* and *omdat* over the different SOC types.

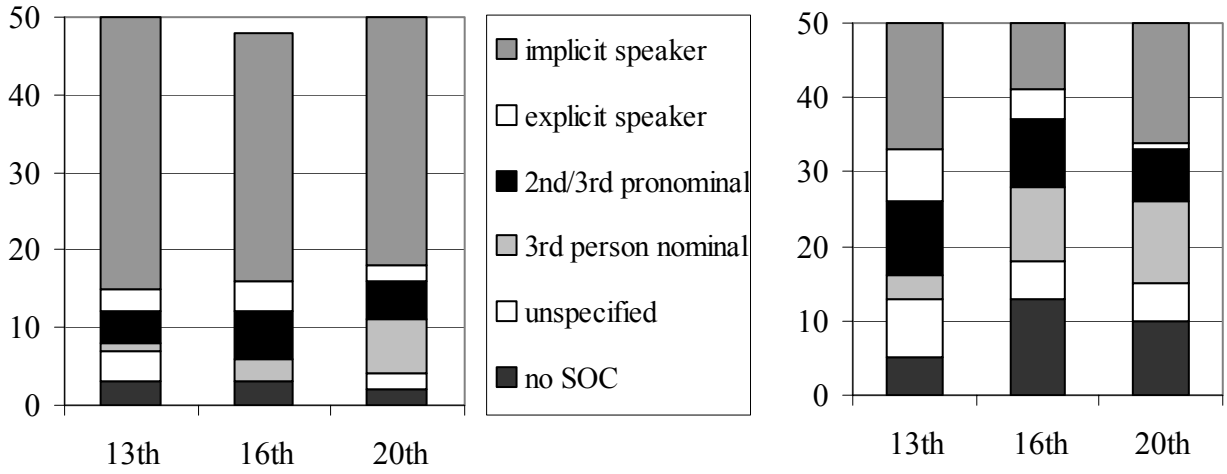


Figure 8.1. Distribution of *want* (left) and *omdat* over the SOC types in three periods

The connective *want* is accompanied more often by an implicit speaker SOC than by any of the other SOC types (χ^2 (5) = 189.1; $p < .001$, see Appendix H-1). As can be inferred from the high number of implicit speaker SOC's in Figure 8.1, *want* is already rather subjective in the 13th century, and this degree of subjectivity remains rather constant. Only in the 20th century can a significant change be observed, namely, an increase in 3rd person nominal SOC's (χ^2 (1) = 4.4; $p < .05$). This is one of the less subjective uses, so the significant increase could point

to an objectification trend. However, to establish objectification, it is not only necessary to observe a significant increase of a more objective category, but also a decrease in one or more of the subjective categories. For *want*, none of the other SOC types shows a substantial decrease in frequency, so it is impossible to decide whether the increase in 3rd person nominal SOC's comes from the more subjective categories at the top of the diagram, or from the less subjective ones at the bottom. To conclude, then, *want* does not show subjectivity changes in one direction or the other.

The connective *omdat* shows a large spread over the different SOC types. Although this distribution does not remain constant during the three periods under investigation (χ^2 (10) = 18.8; p = .04, see Appendix H-2), *omdat* does not show any subjectification trends. On the contrary: several changes in the opposite direction, so-called objectification trends can be found. In the 16th century there is a significant increase in the more objective SOC types *3rd person nominal* (z = 2.37) and *no SOC* (z = 2.38). Both increases are accompanied by a decrease in the use in more subjective SOC categories. However, given their temporary nature, these increases are not regarded as real counterevidence to the subjectification hypothesis.

The SOC results of *daarom* and *dus* are summarized in Figure 8.2.

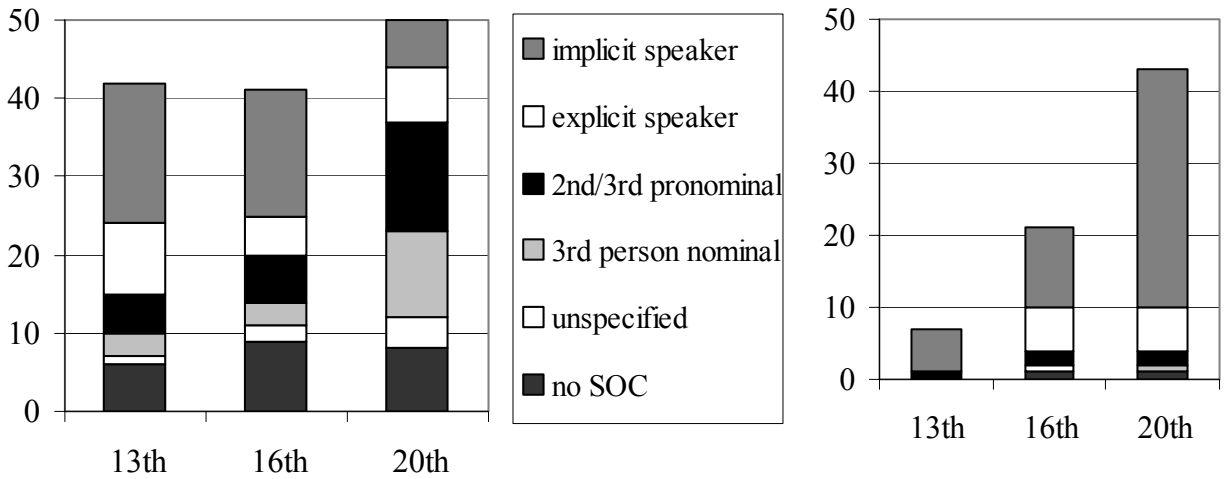


Figure 8.2. Distribution of *daarom* (left) and *dus* over the SOC types in three periods

The third connective, *daarom*, does not show any subjectification trends based on SOC type, but again a change in the direction of objectivity can be found (χ^2 (10) = 22.2; p < .001, see Appendix H-3). In the 20th century three non-speaker SOC types show a significant increase: *3rd person pronominal* (z = 3.03; p = .002), *3rd person nominal* (z = 2.98; p = .002), and *unspecified SOC* (z = 2.05; p = .04). These SOC types are at the more objective side of the subjectivity continuum. The increases in the more objective categories are accompanied by a decrease in the most subjective category. So, *daarom* shows objectification in the 20th century.

As Figure 8.2 furthermore shows, the connective use of *dus* has a continuous preference for implicit speaker SOC's (χ^2 (5) = 116.6; p < .001, see Appendix H-4). There are neither subjectification nor objectification trends; the SOC profile of *dus* remains relatively constant.

All in all, based on SOC type no subjectification trends can be found within the clause combining function of the four connectives. On the contrary, for one connective, *daarom*, a so-called objectification trend was found (see Table 8.3).

Table 8.3. Overview of the results on SOC type

Connective	Subjectification?	Objectification?
want	no	no
omdat	no	no
daarom	no	yes: increase *3rd person pronominal*, *3rd person nominal*, and *unspecified SOC* in the 20th century
dus	no	no

8.5.3 Subjectification based on domains?

The analysis of the four connectives based on domains confirms the picture of the SOC analyses. Figure 8.3 shows the distribution of *want* and *omdat* over the different domains of use.

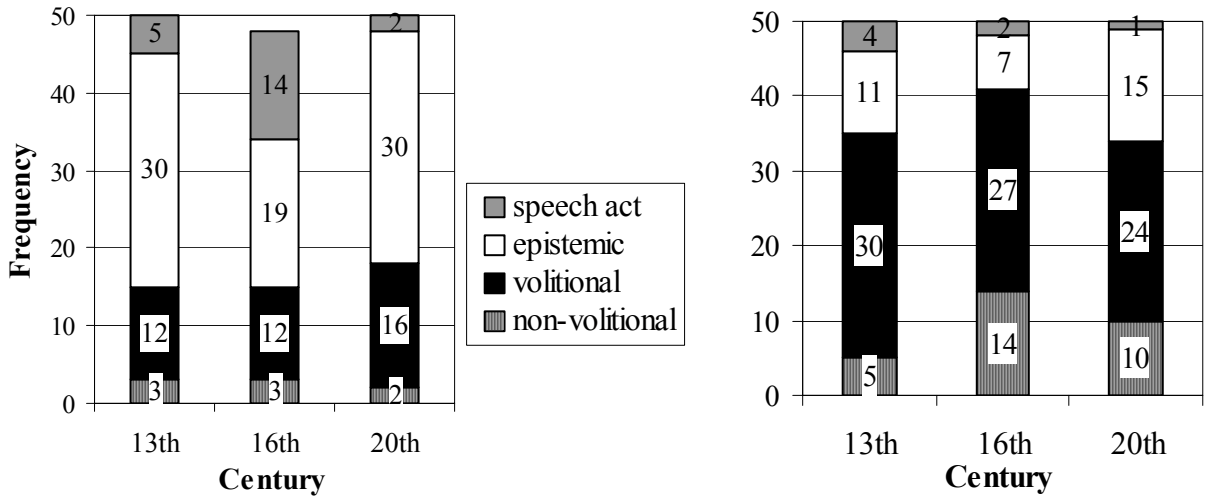


Figure 8.3. Distribution of *want* (left) and *omdat* over the domains of use in three periods

The connective *want* has hardly changed during the selected time span of 800 years; only in the 16th century was a significant increase in speech-act use found (χ^2 (1) = 12.2; p < .001, see Appendix I-1). This increase seems to point to subjectification. However, this subjectification was not a lasting phenomenon, since the number of speech-act fragments decreased again in the 20th century. In order to find an alternative explanation for the temporary increase in speech acts during the 16th century, I have looked at the origin of these speech-act fragments in more detail. This analysis revealed that seven of the fourteen speech-act fragments came from the same page in a moralistic source – *Devoot ende profitelyck boecxken* 'Devout and profitable book' – in the sample of rhyme texts. On this page, advice and orders like (26) are presented for a 'good life-style'. This advice and these orders are frequently supported with arguments, which results in the high number of speech-act relations.

(26) *God kent sijn schapen ende sij hooren na sijn stem*
Nyemant en machse trecken wt sijnder hant / Sijn woert aenhoert
***want** God ghetuycht van hem* (Devoot ende profitelyck boecxken, 1539)
'God knows his sheep and they listen to his voice
No one can draw them from his hands / Listen to His (Jesus') word
because God testifies to him'

I consider this high number of speech acts on one page as an author-specific effect. All other pages showed a lower number of *want*-fragments (varying from three to zero) as well as more variation in their domains interpretations. With this alternative explanation, the temporary change in the use of *want* cannot be regarded as evidence either against or in favor of the unidirectionality of the subjectification hypothesis.

Figure 8.3 indicates that the domains profile of *omdat* is relatively stable across ages (χ^2 (6) = 10.1; p < .25, see Appendix I-2).[8] This diachronic picture remains the same if the finalistic *omdat*-fragments are disregarded in the statistical analysis (see Appendix I-3). Chapter 6 (see section 6.4.1) already revealed that the domains profile of *omdat* in its finalistic use differs from the profile in its 'normal' use. Compared to the 'normal' use of *omdat*, the finalistic use shows a relatively high number of speech act relations. In addition, a more detailed analysis of the content use reveals that finalistic *omdat* is restricted to volitional relations, whereas 'normal' *omdat* can mark both volitional and non-volitional relations. However, the disappearance of finalistic *omdat* does not result in a significant increase in non-volitional use. It can be concluded then, that *omdat* neither shows subjectification nor objectification.

Figure 8.4 shows the distribution of *daarom* and *dus* over the different domains of use.

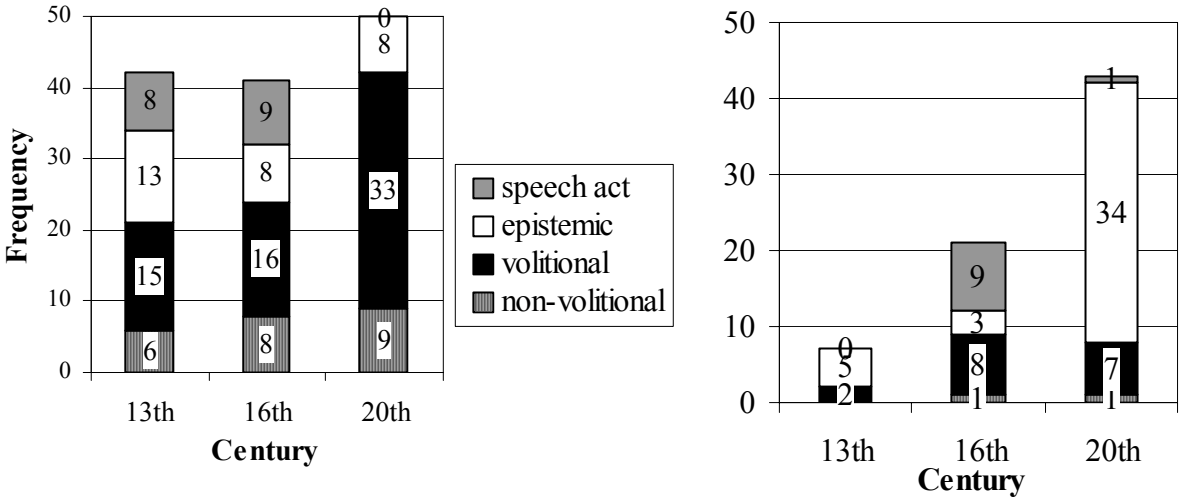


Figure 8.4. Distribution of *daarom* (left) and *dus* over the domains of use in three periods

As Figure 8.4 depicts, the connective use of *daarom* is not stable across ages. The first significant change is the increase in the volitional content use of *daarom* in the 20th century (χ^2 (1) = 16.6; p < .001, see Appendix I-4). Note that the significant decrease in the non-connective use (χ^2 (1) = 10.4; p < .001) in this period cannot account for this increase alone. The increase in volitional content use is also accompanied by a significant decrease in the speech-act domain (χ^2 (1) = 7.6; p < .01), the most subjective end of the subjectivity continuum. From the combination of these observations it can be concluded that the causal connective *daarom* shows objectification.

[8] Note that *omdat* during all periods occurs both in the non-volitional and in the volitional content domain. This is in line with observations by De Rooij (1982), but it is in contrast to the modern Dutch grammarian's rule that non-volitional (cause-result) relations should be marked with *doordat* 'because (of the fact that)', which is restricted to this domain.

For *dus*, the distribution over the domains is not constant (χ^2 (8) = 89.9; p < .001, see Appendix I-5); in the 20th century the epistemic use of *dus* increases significantly (z = 5.24; p < .001). However, this period also shows an increase in the number of fragments in which *dus* is used as a causal connective. This makes it impossible to decide whether the increase in the relatively subjective epistemic domain should be regarded as a case of subjectification or whether it results from the overall increase in connective fragments. A similar argumentation holds for a difference between the 13th and the 16th century: the number of volitional content fragments increases significantly (z = 2.15; p = .03). Since this increase is accompanied by an overall increase in the connective use of *dus*, there is no decisive evidence in favor of the subjectification hypothesis.

The results of the subjectivity analyses based on domains can be summarized as follows.

Table 8.4. Overview of the results on domains

Connective	Subjectification?	Objectification?
want	no	no
omdat	no	no
daarom	no	yes: increase volitional content use in the 20th century
dus	no	no

8.5.4 Subjectification in relation to syntactic changes

Section 8.5.2 and 8.5.3 revealed that none of the selected lexical items showed subjectification within their use as causal connectives. However, in section 8.5.1 I argued that changes towards and away from the connective use of *daarom* and *dus* could be seen as instances of subjectification. In the current section I will try to answer the second research question of this chapter: Are syntactic changes a necessary prerequisite for subjectification to occur?

A first possibility involves the relation between subjectification and changes in categorical status. From the results of *daarom* it can be concluded that subjectification is not exclusively tied to grammatical changes. In addition to its anaphoric meaning, this word gains a connective meaning, which results in an increase in subjectivity. However, the categorical status of *daarom* does not change: *daarom* remains an adverb, both in its function as an anaphoric expression and as a connective (see section 7.5.3).

The diachronic changes in the use of *dus* also suggest that subjectification should not be restricted to changes in categorical status. During the rise of the connective use out of the anaphoric use (in or before the 13th century), *dus* remains an adverb. *Dus* gains the syntactic possibility to function as a complementizer only later on – starting with the ambiguous fragments in the 16th century. With its increasing preference for the connective function, *dus* also shows an increase in its use as a complementizer. The change in categorical status, then, cannot be regarded as a necessary prerequisite for subjectification to occur.

A second possibility involves the relation between subjectification and changes in positioning. In the case of *daarom*, the only significant syntactic change is an increase in the use in clause-medial position in the 20th century. This syntactic change appears to follow up on a change at the text-linguistic level: with the loss of the anaphoric use, the need to use positioning to discriminate between the two conceptual functions has disappeared.

A similar story holds for the significant changes in the positioning of *dus*. In the 13th century, both anaphoric *dus* and connective *dus* could occur in clause-initial as well as in clause-medial positions (see section 7.4.3). With the increase in connective use, the anaphoric

use developed a preference for the clause-medial position, just as the connective function developed a preference for the clause-initial position. This syntactic change does not precede the change at the text-linguistic level, and hence cannot be regarded as a condition for the occurrence of subjectification.

It can be concluded that subjectification is not exclusively tied to changes in positioning or categorical status. This shows that such syntactic changes are not a necessary prerequisite for subjectification.

8.6 Conclusion and discussion

Table 8.5 shows a summary of the main results.

Table 8.5. Overview of the subjectification and objectification tendencies

Connective	Evidence in line with subjectification hypothesis?	Evidence against subjectification hypothesis?
want	no	no
omdat	no	no
daarom	yes (change to connective function)	yes (change within connective function)
dus	yes (change to and from connective function)	no

The first research question of this chapter was: Do the diachronic conceptual changes involve so-called *subjectification*? A first answer to this question is that I did not find subjectification within the connective function of the four words under investigation. There were no lasting increases in their use in the more subjective domains. On the contrary, I found evidence against the subjectification hypothesis: *daarom* showed so-called objectification trends. However, subjectification does occur if functions other than the connective function are taken into account: subjectification is found at changes to or from the use as causal connective. Two examples are the decrease of the anaphoric use of *dus* and *daarom* and the increase in the new discourse marker use of *dus*; both developments lead to an overall subjectification in the use of these words.

The second research question was: Are syntactic changes a necessary prerequisite for subjectification to occur? The diachronic data on *dus* and *daarom* revealed that subjectification should not be tied to syntactic changes. It appears, then, that changes in the positioning or categorical status are not a necessary condition for subjectification.

In the remainder of this section, three points will be taken up for discussion: the controversy between SOC and domains of use (8.6.1), the relative stability of the connectives (8.6.2), and the range of subjectification (section 8.6.3).

8.6.1 On the SOC-domains controversy

In this chapter I presented the domains theory (Sweetser 1990) and the subjectivity scale based on the *subject of consciousness* (Pander Maat & Sanders 2000) as parallel operationalizations of the degree of subjectivity in causal relations. This is not an obvious thing to do, since in the past years these theories have been presented as competitors in answering the question which theory can describe and explain the use of connectives best. In my opinion, this polarization is not fruitful. In section 8.2 I argued that these theories should be considered indicators of subjectivity at different levels: the domains theory is concerned

with the conceptual relation as a whole, whereas the SOC approach focuses on the linguistic realization of a specific subjective element (the SOC) within this relation.

The results in this chapter provide further evidence that the two approaches can be used side by side: both operationalizations give a similar indication of the degree of subjectivity of the selected connectives (see 8.5.2 and 8.5.3). For example, non-volitional content relations have no SOC per definition, speech-act relations and epistemic relations are almost always combined with implicit speaker SOC's and the explicit SOC's are mainly preserved for volitional content relations. The parallel between the results based on SOC types and domains of use does not seem perfect: the objectification of *want* after the 16th century – caused by the author-specific preference for speech-act use in the 16th century – only shows up in the domains analysis. However, in this case the decrease in speech-act relations is accompanied by an increase in the number of epistemic relations. This observation implies that the constancy in the distribution of SOC types is not remarkable after all: both epistemic and speech-act relations are combined with implicit speaker SOC's.

From the parallel between the findings based on domains and based on SOC's, it can be concluded that there is no need to force a choice between the two approaches. The two can be said to be interwoven.

8.6.2 On the stability of connectives

It is remarkable that the four connectives showed relatively constant profiles of use over a period of 800 years, both in terms of SOC type and in terms of domains. Perhaps this should not surprise us too much, since the specific profile of a certain connective guarantees its right to exist. Drastic change might result in a complete overlap with the profile of another causal connective, which could subsequently lead to the disappearance of one of the two connectives. The four connectives in this study indeed show a certain degree of specialization (cf. Knott & Sanders 1998); they have distinct conceptual profiles: *dus* and *daarom* function as forward causals, whereas *want* and *omdat* are markers of backward causality. Within these directions one relatively subjective and one relatively objective connective can be found. These pairs can be complemented with two even more objective connectives, *daardoor* and *doordat*, thus covering the complete continuum from objective to subjective.[9]

Still, certain conceptual shifts were found. From these changes it can be concluded that specialization of connectives indeed plays a role. This can be shown from the divergence of *aldus* and *dus*. In Middle Dutch these words could be exchanged freely; in modern Dutch *aldus* can only fulfill the anaphoric function, whereas *dus* can only be used as a causal connective or as a discourse marker. A second type of specialization has occurred with *opdat* and *omdat*. In Middle Dutch both connectives could explicate causal relations with a desirable, not yet realized consequent; in modern Dutch *omdat* has lost this possibility. As a result, *opdat* now has a specific profile that is clearly distinct from that of *omdat*. Even a third change, the objectification of *daarom* may be the result of specialization: it is likely that *daarom* has become more objective because the connective *dus* came to be used in the relatively subjective areas more often (cf. Stukker 2005).

From the relative stability of the conceptual profiles it can be concluded that subjectification should not be seen as a kind of tendency that occurs automatically. Rather, it

[9] It is true that connectives can be used in contexts other than their prototypical context of use, even in a context that is considered the prototypical context of another connective. In such contexts, however, their frequency is always relatively low (cf. Degand & Sanders 1999). The use in non-prototypical contexts can often be explained on the basis of rhetorical motives (see Van den Hoven 1997).

appears that connectives change when they come into competition with another connective expressing a similar relation. The change resulting from this "specialization process" can either be a case of subjectification, or of objectification.

8.6.3 On the range of subjectification

On the basis of my results some remarks can be made on the range of the subjectification theory as proposed by Traugott (1995). It can be stated that subjectification only takes place at changes across functions, e.g. at the transition of lexical to functional (from deictic or anaphoric use to connective use), or at the transition from one text-linguistic function to another (from connective to discourse marker use). Contrary to Traugott's (1995) findings on *while*, subjectification does not occur between different uses within one connective function (like the domains within the use as causal connective). On the contrary, within the connective function, changes can even go in the opposite direction.

There are two possible explanations for the absence of subjectification tendencies at changes within the connective function. A first explanation has already been presented in section 8.6.2: connective changes based on domains may result from a "specialization process", which can either be a case of subjectification or of objectification. The second explanation is that none of the connectives under discussion showed real changes in their domains of use in the sense that they came to be used in a domain in which they could not occur earlier. The connectives *want*, *omdat*, and *daarom* could be used in all three domains from the 13th century on. It is likely that the same holds for *dus* as well, although this connective was not attested in the content domain during the 13th century. It may be the case that subjectification only occurs when lexical items gain new meanings, and not when lexical items show a shift in the distribution over the meanings they can already express.

The results in this chapter can be used to refine the subjectification hypothesis in a second way. This modification concerns the moment in the grammaticalization process at which subjectification occurs. Traugott (1995: 47) suggests that subjectification mainly occurs at the beginning of the grammaticalization process, during the shift from lexical element to function word. From this perspective, the diachronic development of *dus* is especially interesting. The rise of the discourse marker use, which is a case of subjectification, can be seen as a second stage in the grammaticalization process (a linguistic element that is grammaticalized to a certain extent, gains another grammatical function). This implies that subjectification can also occur at later stages in the grammaticalization process (which Traugott also shows herself in Traugott & Dasher 2002).

In this chapter I have tested the subjectification hypothesis by analyzing four Dutch causal connectives. However, more quantitative research – both on Dutch as well as on other languages – is needed to investigate the range of subjectification in a more detailed way. For example, no claims can be made on changes from the causal function to other types of coherence relations, since the current study is restricted to an analysis of causal connectives. Traugott's (1995) work gives an indication that subjectification does occur at such transitions: the shift from temporal *while* into contrastive *while* goes hand in hand with an increase in subjectivity. It seems a promising undertaking, then, to perform such analyses.

A final point regarding the methodology of this study requires mention here. An important goal of this research was to develop more objective and quantifiable methods to test the subjectification hypothesis. With the parallel operationalizations, I have tried to make the concepts of 'subjectivity' and 'subjectification' more concrete. My empirical research based on these operationalizations shows how a quantitative approach can contribute to the development of a theory on the range of subjectification.

APPENDICES TO CHAPTER 8

Appendix H – Logit analyses of the diachronic SOC developments

Overview of the content of Appendix H:
H – 1 Logit analysis of the diachronic SOC development of *want*
H – 2 Logit analysis of the diachronic SOC development of *omdat*
H – 3 Logit analysis of the diachronic SOC development of *daarom*
H – 4 Logit analysis of the diachronic SOC development of *dus*

H – 1. Logit analysis of the diachronic SOC development of *want*

I – Remarks
Two fragments from the 16th century have been left out of the statistic analyses. These involve *want*-fragments with a temporal meaning.

II – Data: Distribution of *want* over the SOC types in three periods

	Period			Total
	13th century	16th century	20th century	
1. implicit speaker	35	32	32	99
2. explicit speaker	3	4	2	9
3. explicit 2nd or 3rd person pronominal	4	6	5	15
4. explicit 3rd person nominal	1	3	7	11
5. unspecified	4	0	2	6
6. no SOC	3	3	2	8
Total	50	48	50	148
number of words needed for 50 fragments	34253	14790	75799	

III – Results logit analysis SOC types *want*

Logit (fr)	χ^2 model	df	p model	χ^2 factor	df	p factor
1. constant	259.69	17	< .001	-	-	-
+ 2. period	200.61	15	< .001	59.08	2	< .001
+ 3. SOC type	11.47	10	< .5	189.10	5	< .001
+ 4. period x SOC: 3rd person nominal 20th century	7.03	9	< .25	4.44	1	< .05

IV – Parameter estimate SOC types *want* for model 4

Parameter	Estimate	s.e.	z-score	p
constant	-6.90	0.15	-45.22	< .001
period: 16th century	0.80	0.20	3.97	< .001
period: 20th century	-0.90	0.21	-4.32	< .001
SOC type: explicit speaker	-2.40	0.35	-6.89	< .001
SOC type: 3rd person pronominal	-1.89	0.28	-6.81	< .001
SOC type: 3rd person nominal	-2.83	0.51	-5.52	< .001
SOC type: unspecified	-2.79	0.42	-6.68	< .001
SOC type: no SOC	-2.52	0.37	-6.85	< .001
period x SOC: 3rd person nominal 20th century	1.34	0.65	2.06	0.04

H – 2. Logit analysis of the diachronic SOC development of *omdat*

I – Remarks

No remarks.

II – Data: Distribution of *omdat* over the SOC types in three periods

	Period			Total
	13th century	16th century	20th century	
1. implicit speaker	17	9	16	42
2. explicit speaker	7	4	1	12
3. explicit 2nd or 3rd person pronominal	10	9	7	26
4. explicit 3rd person nominal	3	10	11	24
5. unspecified	8	5	5	18
6. no SOC	5	13	10	28
Total	50	50	50	150
number of words needed for 50 fragments	83400	95505	61574	

III – Results logit analysis SOC types *omdat*

Logit (fr)	χ^2 model	df	p model	χ^2 factor	df	p factor
1. constant	44.30	17	< .001	-	-	-
+ 2. period	39.40	15	< .001	4.90	2	< .1
+ 3. SOC type	18.83	10	< .05	20.57	5	< .001
+ 4. period x SOC type	0	0	1	18.83	10	< .05

IV – Parameter estimate SOC types *omdat* for model 4

Parameter	Estimate	s.e.	z-score	p
constant	-8.50	0.24	-35.03	< .001
period: 16th century	-0.77	0.41	-1.87	0.06
period: 20th century	0.24	0.35	0.70	0.48
SOC type: explicit speaker	-0.89	0.45	-1.98	0.05
SOC type: 3rd person pronominal	-0.53	0.40	-1.33	0.18
SOC type: 3rd person nominal	-1.74	0.63	-2.77	0.01
SOC type: unspecified	-0.75	0.43	-1.76	0.08
SOC type: no SOC	-1.22	0.51	-2.41	0.02
16th century x explicit speaker	0.08	0.75	0.10	0.92
16th century x 3rd pronominal	0.53	0.62	0.86	0.39
16th century x 3rd nominal	1.84	0.78	2.37	0.02
16th century x unspecified	0.17	0.70	0.24	0.81
16th century x no SOC	1.59	0.67	2.38	0.02
20th century x explicit speaker	-1.89	1.12	-1.68	0.09
20th century x 3rd pronominal	-0.30	0.60	-0.49	0.62
20th century x 3rd nominal	1.36	0.74	1.84	0.07
20th century x unspecified	-0.41	0.67	-0.61	0.54
20th century x no SOC	0.75	0.65	1.16	0.25

H – 3. Logit analysis of the diachronic SOC development of *daarom*

I – Remarks

The *daarom*-fragments in which *daarom* did not function as a connective have been left out of the statistical analysis.

II – Data: Distribution of *daarom* over the SOC types in three periods

	Period			Total
	13th century	16th century	20th century	
1. implicit speaker	18	16	6	40
2. explicit speaker	9	5	7	21
3. explicit 2nd or 3rd person pronominal	5	6	14	25
4. explicit 3rd person nominal	3	3	11	17
5. unspecified	1	2	4	7
6. no SOC	6	9	8	23
Total	42	41	50	133
number of words needed for 50 fragments	148679	134016	232378	

III – Results logit analysis SOC types *daarom*

Logit (fr)	χ^2 model	df	p model	χ^2 factor	df	p factor
1. constant	52.92	17	< .001	-	-	-
+ 2. period	49.71	15	< .001	3.21	2	< .25
+ 3. SOC type	22.20	10	< .025	27.51	5	< .001
+ 4. period x SOC type	0	0	1	22.20	10	< .025

IV – Parameter estimate SOC types *daarom* for model 4

Parameter	Estimate	s.e.	z-score	p
constant	-9.02	0.24	-38.26	< .001
period: 16th century	-0.01	0.34	-0.04	0.97
period: 20th century	-1.55	0.47	-3.28	< .001
SOC type: explicit speaker	-0.69	0.41	-1.70	0.09
SOC type: 3rd person pronominal	-1.28	0.51	-2.53	0.01
SOC type: 3rd person nominal	-1.79	0.62	-2.87	0.004
SOC type: unspecified	-2.89	1.03	-2.82	0.01
SOC type: no SOC	-1.10	0.47	-2.33	0.02
16th century x explicit speaker	-0.47	0.66	-0.72	0.47
16th century x 3rd pronominal	0.30	0.70	0.43	0.67
16th century x 3rd nominal	0.12	0.89	0.13	0.90
16th century x unspecified	0.81	1.27	0.64	0.52
16th century x no SOC	0.52	0.63	0.83	0.41
20th century x explicit speaker	0.85	0.69	1.23	0.22
20th century x 3rd pronominal	2.13	0.70	3.03	0.002
20th century x 3rd nominal	2.40	0.80	2.98	0.003
20th century x unspecified	2.49	1.21	2.05	0.04
20th century x no SOC	1.39	0.72	1.93	0.05

H – 4. Logit analysis of the diachronic SOC development of *dus*

I – Remarks

The fragments in which *dus* did not function as a connective (i.e. the anaphoric and discourse marker use of *dus*) have been left out of the statistical analysis.

II – Data: Distribution of *dus* over the SOC types in three periods

	Period			Total
	13th century	16th century	20th century	
1. implicit speaker	6	11	33	50
2. explicit speaker	0	6	6	12
3. explicit 2nd or 3rd person pronominal	1	2	2	5
4. explicit 3rd person nominal	0	0	1	1
5. unspecified	0	1	0	1
6. no SOC	0	1	1	2
Total	7	21	43	71
number of words needed for 50 fragments	146606	51516	95769	

III – Results logit analysis SOC types *dus*

Logit (fr)	χ^2 model	df	p model	χ^2 factor	df	p factor
1. constant	176.21	17	< .001	-	-	-
+ 2. period	124.68	15	< .001	51.53	2	< .001
+ 3. SOC type	8.09	10	< .75	116.60	5	< .001

IV – Parameter estimate SOC types *dus* for model 3

Parameter	Estimate	s.e.	z-score	p
constant	-8.16	0.23	-35.29	< .001
period: 16th century	-2.09	0.43	-4.90	< .001
period: 20th century	0.09	0.27	0.35	0.72
SOC type: explicit speaker	-1.42	0.32	-4.43	< .001
SOC type: 3rd person pronominal	-2.30	0.47	-4.91	< .001
SOC type: 3rd person nominal	-3.73	0.92	-4.04	< .001
SOC type: unspecified	-3.73	0.92	-4.05	< .001
SOC type: no SOC	-3.17	0.70	-4.50	< .001

Appendix I – Logit analyses of the diachronic domain developments

Overview of the content of Appendix I:

I – 1. Logit analysis of the diachronic domain development of *want*

I – Remarks

The two *want*-fragments labeled 'other use' have been disregarded in the statistic analysis.

II – Data: Distribution of *want* over the domains of use in three periods

	Period			Total
	13th century	16th century	20th century	
1. speech act	5	14	2	21
2. epistemic	30	19	30	79
3. content volitional	12	12	16	40
4. content non-volitional	3	3	2	8
5. other use	0	2	0	2
Total	50	50	50	150
number of words needed for 50 fragments	34253	14790	75799	

III – Results logit analysis domains of use *want*

Logit (fr)	χ^2 model	df	p model	χ^2 factor	df	p factor
1. constant	151.97	11	< .001	-	-	-
+ 2. period	93.08	9	< .001	58.89	2	< .001
+ 3. domain	15.25	6	< .025	77.83	3	< .001
+ 4. period x domain: speech act 16th century	3.05	5	< .25	12.20	1	< .001

IV – Parameter estimate domains of use *want* for model 4

Parameter	Estimate	s.e.	z-score	p
constant	-9.37	0.37	-25.23	< .001
period: 16th century	0.53	0.22	2.35	0.02
period: 20th century	-0.80	0.20	-3.97	< .001
domain: content volitional	1.61	0.39	4.16	< .001
domain: epistemic	2.29	0.37	6.18	< .001
domain: speech act	0.18	0.52	0.34	0.73
period x domain: speech act 16th century	1.70	0.50	3.37	< .001

I – 2. Logit analysis of the diachronic domain development of *omdat*

I – Remarks

- The category 'other use' has been disregarded in the statistical analysis.
- The finalistic *omdat*-fragments are included in this statistical analysis.

II – Data: Distribution of *omdat* over the domains of use in three periods

	Period			Total
	13th century	16th century	20th century	
1. speech act	4	2	1	7
2. epistemic	11	7	15	33
3. content volitional	30	27	24	81
4. content non-volitional	5	14	10	29
5. other use	0	0	0	0
Total	50	50	50	150
number of words needed for 50 fragments	83400	95505	61574	

III – Results logit analysis domains of use *omdat*

Logit (fr)	χ^2 model	df	p model	χ^2 factor	df	p factor
1. constant	92.93	11	< .001	-	-	-
+ 2. period	88.03	9	< .001	4.90	2	< .1
+ 3. domain	10.10	6	< .1	77.93	3	< .001
+ 4. period x domain	0	0	1	10.10	6	< .25

IV – Parameter estimate domains of use *omdat* for model 4

Parameter	Estimate	s.e.	z-score	p
constant	-7.93	0.18	-43.43	< .001
period: 16th century	-0.24	0.27	-0.91	0.36
period: 20th century	0.08	0.27	0.29	0.77
domain: content non-volitional	-1.79	0.48	-3.71	< .001
domain: epistemic	-1.00	0.35	-2.85	0.004
domain: speech act	-2.02	0.53	-3.79	< .001
period x domain: 16th cont. non-volitional	1.14	0.58	1.94	0.05
period x domain: 16th epistemic	-0.35	0.55	-0.63	0.53
period x domain: 16th speech act	-0.59	0.91	-0.65	0.52
period x domain: 20th cont. non-volitional	0.92	0.61	1.50	0.13
period x domain: 20th epistemic	0.53	0.48	1.11	0.27
period x domain: 20th speech act	-1.16	1.15	-1.01	0.31

I – 3. Logit analysis of the diachronic domain development of non-finalistic *omdat*

I – Remarks

The category 'other use' as well as the finalistic fragments have been disregarded in the statistical analyses.

II – Data: Distribution of *omdat* over the domains of use in three periods

	Period			Total
	13th century	16th century	20th century	
1. speech act	2	0	1	3
2. epistemic	9	7	15	31
3. content volitional	23	20	24	67
4. content non-volitional	5	14	10	29
5. other use	0	0	0	0
Total	39	41	50	130
number of words needed for 50 fragments	83400	95505	61574	

III – Results logit analysis domains of use *omdat*

Logit (fr)	χ^2 model	df	p model	χ^2 factor	df	p factor
1. constant	91.74	11	< .001	-	-	-
+ 2. period	81.30	9	< .001	10.43	2	< .01
+ 3. domain	8.66	6	< .1	72.64	3	< .001
+ 4. period x domain	0	0	1	8.66	6	< .25

IV – Parameter estimate domains of use *omdat* for model 4

Parameter	Estimate	s.e.	z-score	p
constant	-8.20	0.21	-39.31	< .001
period: 16th century	-0.28	0.31	-0.90	0.37
period: 20th century	0.35	0.29	1.19	0.23
domain: content non-volitional	-1.53	0.49	-3.09	0.002
domain: epistemic	-0.94	0.39	-2.39	0.02
domain: speech act	-2.44	0.74	-3.31	< .001
period x domain: 16th content non-volitional	1.17	0.60	1.94	0.05
period x domain: 16th epistemic	-0.11	0.59	-0.19	0.85
period x domain: 16th speech act	-2.86	3.26	-0.88	0.37
period x domain: 20th content non-volitional	0.65	0.62	1.05	0.29
period x domain: 20th epistemic	0.47	0.51	0.91	0.36
period x domain: 20th speech act	-0.74	1.26	-0.58	0.56

I – 4. Logit analysis of the diachronic domain development of *daarom*

I – Remarks
No remarks.

II – Data: Distribution of *daarom* over the domains of use in three periods

	Period			Total
	13th century	16th century	20th century	
1. speech act	8	9	0	17
2. epistemic	13	8	8	29
3. content volitional	15	16	33	64
4. content non-volitional	6	8	9	23
5. other use	8	9	0	17
Total	50	50	50	150
number of words needed for 50 fragments	148679	134016	232378	

III – Results logit analysis domains of use *daarom*

Logit (fr)	χ^2 model	df	p model	χ^2 factor	df	p factor
1. constant	89.69	14	< .001	-	-	-
+ 2. period	80.91	12	< .001	8.78	2	< .025
+ 3. domain	36.97	8	< .001	43.95	4	< .001
+ 4. per. x domain: volitional 20th century	20.39	7	< .005	16.57	1	< .001
+ 5. per. x dom: speech act 20th century	12.79	6	< .05	7.60	1	< .01
+ 6. per. x domain: other use 20th century	2.41	5	< .9	10.37	1	< .001

IV – Parameter estimate domains of use *daarom* for model 6

Parameter	Estimate	s.e.	z-score	p
constant	-9.86	0.25	-39.36	< 0.001
period: 16th century	0.10	0.20	0.52	0.60
period: 20th century	-0.48	0.31	-1.52	0.13
domain: content volitional	0.69	0.29	2.38	0.02
domain: epistemic	0.23	0.28	0.83	0.41
domain: speech act	0.09	0.33	0.28	0.77
domain: other use	0.09	0.33	0.28	0.77
period x domain: volitional 20th century	0.79	0.39	2.03	0.04
period x domain: speech act 20th century	-4.41	3.18	-1.39	0.17
period x domain: other use 20th century	-4.41	3.18	-1.39	0.17

I – 5. Logit analysis of the diachronic domain development of *dus*

I – Remarks
No remarks.

II – Data: Distribution of *dus* over the domains of use in three periods

	Period			Total
	13th century	16th century	20th century	
1. speech act	0	9	1	10
2. epistemic	5	3	34	42
3. content volitional	2	8	7	17
4. content non-volitional	0	1	1	2
5. other use	43	29	7	79
Total	50	50	50	150
number of words needed for 50 fragments	146606	51516	95769	

III – Results logit analysis domains of use *dus*

Logit (fr)	χ^2 model	df	p model	χ^2 factor	df	p factor
1. constant	244.40	14	< .001	-	-	-
+ 2. period	218.33	12	< .001	26.06	2	< .001
+ 3. domain	89.94	8	< .001	12.84	4	< .025
+ 4. period x domain	0	0	1	89.94	8	< .001

IV – Parameter estimate domains of use *dus* for model 4

Parameter	Estimate	s.e.	z-score	p
constant	-7.48	0.19	-40.29	< .001
period: 13th century	-0.65	0.24	-2.71	0.01
period: 20th century	-2.04	0.42	-4.85	< .001
domain: speech act	-1.17	0.38	-3.07	0.002
domain: epistemic	-2.27	0.61	-3.74	< .001
domain: content volitional	-1.29	0.40	-3.22	0.001
domain: content non-volitional	-3.37	1.02	-3.31	< .001
period x domain: speech act 13th century	-4.89	3.19	-1.54	0.12
period x domain: epistemic 13th century	0.12	0.77	0.15	0.88
period x domain: volitional 13th century	-1.78	0.83	-2.15	0.03
period x domain: non-volitional 13th century	-2.70	3.32	-0.81	0.41
period x domain: speech act 20th century	-0.78	1.14	-0.68	0.50
period x domain: epistemic 20th century	3.85	0.74	5.24	< .001
period x domain: volitional 20th century	1.29	0.67	1.93	0.05
period x domain: non-volitional 20th century	1.42	1.48	0.96	0.34

Part III – Developments in the acquisition of connectives

Chapter 9

On the methodology of the acquisition studies

In my search for interactions between conceptual and syntactic properties of connectives, I have studied two types of developmental data. In Chapter 5 to 8 I discussed the diachronic development of several Dutch connectives. In Chapter 9 to 12 I will focus on connective development in the area of first language acquisition. The current chapter provides some theoretical and methodological background for the chapters that follow.

"For what does it mean to acquire the meaning of the word?"
(Nelson 1991: 281)

9.1 Introduction

Data from first language acquisition form the second window on possible interactions between conceptual and syntactic properties of connectives and the clauses they host. Byrnes & Gelman (1991: 4) already stressed the suitability of this kind of developmental data: "A developmental approach is inherently concerned with process, and so is especially suitable for detecting the relations among the ingredients of a complex system (…). Studying development is thus a powerful methodological tool and potentially more revealing about the organization of a system than any attempt to infer it from the adult end product."[1]

In Chapters 10 to 13, I will discuss the acquisition of several Dutch connectives from different perspectives. For example, I will be looking at the orders of acquisition of particular connectives, and at the frequencies with which children use these connectives. In the current chapter, two preliminary questions about connective acquisition are answered.

(1) <u>Main questions of this chapter:</u>
   a. What does it mean to 'acquire' a connective?
   b. What is an appropriate way to establish 'acquisition'?

Section 9.2 will show that acquisition involves both the emergence and the further development of connectives. Section 9.3 argues that different methods can be used to investigate different parts of the acquisition process. These answers are needed to provide some methodological background for the following chapters, which each focus on a different aspect of connective acquisition in Dutch. Section 9.4, which introduces previous findings on the acquisition of Dutch connectives, will show the need for additional research into the acquisition of connectives. Finally, section 9.5 presents an outline of Chapter 10 to 12, also accounting for the choice of their topics.

[1] De Haan (1987) is somewhat more cautious in using child data as independent (i.e. the only conclusive) evidence: "Given the state of our knowledge of child grammars, it is rational to take evidence from child language only seriously, for the time being, as <u>confirming</u> evidence" (p. 37).

9.2 Connective 'acquisition'

Connectives are common elements in every day language: all adults frequently use connectives when they produce texts. It is clear, however, that this proficiency is not there from the start: connectives are not part of children's first utterances nor of their first conjoined clauses (cf. Diessel 2004: 158). So, connectives are among the many linguistic elements that have to be acquired. But what exactly does it mean to 'acquire' a certain connective? An answer to this question is needed to determine which methods are appropriate for establishing connective acquisition. In the first subsection, 9.2.1, I work out three components involved in the process of mastering connectives. In section 9.2.2 I go on to discuss what full mastery of a connective implies.

9.2.1 Child development and linguistic mapping

The acquisition of connectives is not an isolated process; it is tightly related to the overall cognitive and linguistic development of the child. In the process of connective acquisition (and, in fact, of each linguistic element) three components are involved: the child's conceptual or cognitive development, its syntactic development and the linguistic mapping between concepts and words (see Wagner 1998: 104-106 for a similar proposal concerning the acquisition of linguistic elements expressing time). Each of these components may be used to account for certain characteristics of the process of connective acquisition.

To begin with, the connective development of children is in part determined by the general cognitive development they go through (see, among many others, Piaget 1969). Ingram defines this cognitive development as "the infant's growing knowledge of the world around him" (1989: 115). In order to be able to linguistically mark coherence relations between clauses, children have to be familiar with the concepts underlying these relations. For example, children can only express a variety of temporal relations if they have some sense of notions like *temporal perspective* (present, past, or future), *sequence* (temporal order) and *length* (duration of or interval between events) (cf. Nelson 1991: 287-291).[2] The link between cognitive development and the rise of connectives in child speech has also been noted in Braunwald's (1985) report on the connective acquisition of her daughter Laura: "Connectives first appear in Laura's speech at the same time as a general developmental advance in her ability to relate her concrete objective experiences to internal psychological processes. The content of her language indicates that Laura is discovering intentionality in the concrete sense of an emerging awareness that people are separate individuals with wills, feelings, and minds of their own" (p. 520-521).

Secondly, the connective development of children is also related to their progress in syntactic abilities. For example, there is no sense in acquiring subordinating connectives if the child does not produce complex clauses at all. In several publications, beginning with Slobin (1973), Slobin has demonstrated that even when a meaning is potentially accessible to children (given their cognitive development), their expression of it may be delayed, at least in a conventional way, because of complexity in the formal mechanism used to encode it.[3]

---

[2] Nelson (1991) claims that the relation between language and thought in development is bidirectional: cognitive development influences linguistic development, but linguistic development also influences cognitive development.

[3] Note that the syntactic development of children cannot be regarded completely independently of their cognitive development. For instance, the syntactic ability to produce preposed adverbial clauses is restricted by the child's memory span.

The third component involved in the acquisition of connectives is the linguistic mapping between concepts and words. Children have to learn the language-specific codings of the various coherence relations. For instance, in order for children to learn what concepts go with which pieces of morphology, they need linguistic evidence (Wagner 1989: 106). This is a kind of lexical learning; children have to find out word for word which meanings can be expressed and what the syntactic properties of the words are. This acquisition process goes on until they reach adult-like proficiency.

Each of the three factors discussed above may be used to account for certain characteristics of the process of connective acquisition. For example, in Chapter 10 the order of emergence of Dutch connectives will be related to the cognitive development children undergo.

9.2.2 From emergence to full mastery

The concept of 'acquisition' has been defined in different ways in research concerning connective acquisition, and these differences have also led to different results. For example, Bloom et al. (1980) observed that *but* is already 'acquired' by children aged three. In contrast, Piaget (1969) claimed that English children have only 'acquired' the adversative *but* when they are seven or eight years old. This large difference in age can be accounted for by looking at the different components of 'acquisition': the emergence of connectives, the process of further development, and finally, full mastery of connectives. Bloom et al. (1980) equate 'acquisition' to the first emergence of connectives, which explains why they come up with the age of three. Piaget (1969), however, focuses on full mastery of the connective, which implies more than mere emergence.

To clarify the discussion on the notion of 'acquisition', Nelson's (1991) subdivision seems useful. She claims: "four components are involved in the acquisition of a new word: (a) acquisition of the phonological form of the word; (b) acquisition of its extension to referents or its use in discourse; (c) acquisition (by inference or through explication) of its semantics – the meaning it conveys; (d) understanding of its relations to other forms in the lexicon (including possible paraphrases)" (p. 283). Hence, between emergence and full mastery, a process of further 'acquisition' takes place. Although the phonological form of a certain connective may appear early in a grammatical child utterance, this does not necessarily imply that the child instantaneously understands all the meanings it may convey. And – as is already known from the literature – the early use of many connectives is indeed rather restricted. For example, Wing & Scholnick (1981) notice: "Although many subordinating conjunctions appear in speech before the age of six, studies of comprehension show that children rarely appreciate the full logical implications of these conjunctions until much later" (p. 347). Furthermore, Braunwald (1997) states that although preschool aged children use causal connectives correctly in their language production, they do so "within a restricted range of linguistic and social contexts" (see Braunwald 1997: 120, as well as the many references there). Connectives can therefore be said to adhere to a general acquisition pattern mentioned in Berman (1996): "Across time, use of any given form is extended and hence reconstrued in a variety of interrelated ways" (p. 345).

To define the end point of acquisition four criteria can be used: the three criteria listed in Berman (1996) and one criterion mentioned in Nelson (1991). According to Berman, full mastery or complete acquisition of a linguistic form implies that the child can use it (1) with different semantic and/or pragmatic denotations;[4] (2) in extended syntactic contexts; and (3)

[4] For ease of exposition, I have grouped together Berman's "different semantic denotations" and "different discourse roles in connecting parts of a text".

with different stylistic levels or usage registers (cf. Berman 1996: 345). In the case of connective acquisition, this developmental process from emergence to full mastery takes time, sometimes many years (e.g. in the case of *but* and *although*, see Piaget 1969). A fourth criterion, found in Nelson (1991), can be added to Berman's criteria for full mastery. In Nelson's view, full mastery also implies that children understand its relations to other forms in the lexicon. For connectives this means that children are able to make a deliberate choice between (a) leaving the coherence relation implicit; (b) using an underspecified connective (i.e., a connective which is less specific than the intended relation, cf. Spooren 1997); or (c) selecting a specified connective from a set of more or less equally suitable connectives, taking into account the specific linguistic context (e.g. narrative or argumentative) as well as the social context (e.g. formal or informal).

9.3 Methods to establish connective 'acquisition'

Having defined the notion of connective acquisition, the following question can now be answered: What is an appropriate way to establish 'acquisition'? Given the three definitions of 'acquisition' (emergence, further development, and full mastery), it is likely that there is not simply one correct way to establish connective acquisition. And indeed, several methods can be found from the literature. Below, I will discuss proposals concerning acquisition based on quantitative emergence criteria (section 9.3.1), qualitative emergence criteria (9.3.2), percentages of a parental target (9.3.3), and developmental curves (9.3.4).

9.3.1 Quantitative emergence criteria

Several researchers looking at longitudinal connective data have tried to establish an order of emergence.[5] They often determine this order on the basis of a quantitative emergence criterion based on productivity (see Table 9.1).

Table 9.1. Emergence criteria in longitudinal studies providing English connective data

Researcher(s)	Emergence criterion
Bloom et al. (1980: 239-240)	five different occurrences, irrespective of phrasal or clausal use
Braunwald (1985: 514, 517)	five occurrences, irrespective of meaning
Diessel (2004: 172)	one occurrence

Diessel (2004) indeed uses the very first occurrence of the connective expressing a clausal coherence relation as an indication of emergence. This is in line with the method in Clancy, Jacobsen & Silva (1976: 72), who used the first utterance expressing a specific conjunction type to determine developmental sequences of coherence relations in different languages. The other researchers mentioned in Table 9.1 are more conservative (see also Bloom, Lightbown & Hood 1975): they set a criterion of productivity to establish acquisition, namely "the

[5] This *order of emergence* concerns the development of new forms for a particular function and should be distinguished from the *order of development*, the development of more functions for the same form (see Berman 1996: 346). As Brown & Hanlon (1970) clarify, we must "distinguish between two sorts of order of emergence in child speech. The first sort is an order among constructions which are, all of them, mature adult forms. (…) The second sort of order concerns constructions which are equivalent semantically but which exist in one or more immature or childish forms as well as, eventually, the adult form" (p. 192-193). This discussion is thus restricted to the order of emergence of the first type.

occurrence of five or more different utterances in at least two successive observations" (Bloom et al. 1980: 239).[6]

Do these methods result in different acquisition orders? In order to make a direct comparison between them I have looked at the connective data from the child Peter, who has been studied by both Bloom et al. (1980: 242) and Diessel (2004: 198). Table 9.2 shows the relevant acquisition orders.

Table 9.2. Peter's orders of emergence based on fifth or first use (with ages in years;months)[7]

Order of emergence	Criterion of five occurrences	Criterion of one occurrence
1	and (2;3)	and (1;11)
2	because (2;9)	because, so (2;5)
3	when (2;11)	when (2;7)
4	so, but (3;2)	but (2;8)

As can be inferred from this table, applying the productivity criterion not only results in a different acquisition order (cf. the different rankings of the connective *so*), but also in an emergence delay of at least four months (which might also be partially due to the selection of recordings in the study by Bloom and her colleagues).

Applying productivity criteria to establish acquisition has received some critique. Bowerman (1975) mentions that setting an arbitrary frequency criterion is not a particularly good measure, since it will be highly influenced by sample size. Furthermore, as Ingram (1981) points out, it does not take into consideration crucial information about the distributional properties of words. Brown, Cazden & Bellugi-Klima (1969: 144) already discuss the danger of an arbitrary frequency threshold. They observe that the frequencies they studied in mother-to-child English were astonishingly stable across the three mothers in their study. Therefore, they claim "there seems to be something like a standard frequency profile for mother-to-child English". Because they also found that frequencies in child speech tend to match adult frequencies (within the limits of the child's competence), this implies that an alternative explanation for a certain 'acquisition order' is possible. "Highly stable orders of construction emergence, in terms of an arbitrary frequency criterion, are not inconsistent with the possibility that the children in question know how to form all the constructions from the start but produce them with unequal frequency" (Brown et al. 1969: 145).

As Brown et al. (1969: 145) suggest, there are "various ways out of the trap, all involving the utilization of data that are better indices of knowledge or competence than is an arbitrary frequency of production." One of these is to "consider child frequencies against a background of known stable adult frequencies and so set frequency criteria that are not entirely arbitrary".[8] For example, Brown & Hanlon (1970: 184) counted different sentence types of question-tags in 1400-utterance samples drawn from the mothers and found that the lowest output rate was 6 in 2100 utterances. As a threshold of emergence they then adopted the value "6"; the recording in which the child produced a sentence type six times or more was considered to be

[6] Because Braunwald (1985) applies this method to diary notes, she slightly modifies this criterion into "five or more examples from different speech events over a period of time" (p. 514).

[7] The connectives *and then* and *then* have been disregarded in this table, since these are only studied by Bloom and her colleagues.

[8] Van Kampen's (1997, 2001) method of "balanced conversation" is very similar to this approach. Here, the child must reach 90% of the conversation partner (in the given file). This may lead to an acquisition curve, if the connective is sufficiently present in the data (see also Evers & Van Kampen 2001).

the recording of 'emergence'. This approach is an improvement when compared to the arbitrary frequency criteria mentioned above, but it can still be questioned. Despite the label 'emergence', it seems as if Brown & Hanlon establish full mastery with their frequency criterion. This idea is confirmed by the observation that – despite the fact that they call their criterion for acquisition "not unreasonable" – Brown & Hanlon set additional qualitative criteria to exclude non-productive occurrences (see p. 185-186). In the end, quality instead of quantity is the major determinant of acquisition.

An arbitrary productivity criterion seems less appropriate to establish an order of connective acquisition based on full mastery. However, it may still be useful if the type of acquisition is taken into account, that type being the very beginning of connective usage. In line with the words of Bloom et al. (1980: 259), the acquisition reported here "can be described only in terms of emergence, (…) rather than in terms of achievement." No matter which frequency criterion is set, the children at this very early stage of the acquisition process will not be fully productive in their use of connectives. Over time, they will extend their connectives to other linguistic contexts or situations and will go on fine-tuning their usage in other respects (cf. Braunwald 1997 on the development of *because* and *so*), until they reach adult-like proficiency.

Since both Diessel and Bloom focus on the very beginning of the development, there must be another reason why Bloom and her colleagues use a frequency criterion. It seems to me that – like Brown & Hanlon (1970) – they try to guarantee the quality of this early connective usage by applying a quantitative criterion. With five occurrences they avoid the risk of basing their order of emergence on utterances that do not necessarily presuppose knowledge of the meaning represented by the connective. In my view, however, it is not necessary to use a quantitative approach to rule out such so-called non-productive utterances.[9] By setting qualitative criteria in advance, it is equally possible to guarantee the productivity of the first usage.

9.3.2 Qualitative emergence criteria

In order to establish an order of emergence, it seems reasonable to stay as close as possible to the earliest emergence by using first occurrence complemented with certain qualitative criteria. In Chapter 10 the criteria in (2) will be used to establish the order of emergence of Dutch connectives.

(2) <u>Method used to establish the emergence of a connective:</u>
First occurrence in which the connective is being used
a. in a correct way
b. as a word combining two clauses
c. in a creative way.

The first criterion, that the connective must be used in a *correct* way, is needed to avoid that the acquisition of a connective be established on the basis of a fragment that is not grammatical from an adult perspective. Firstly, this implies that only connective fragments in which the related utterances are intelligible are considered "correct". Secondly, it implies that it is possible to interpret the connective in its particular context as expressing an appropriate

[9] Ingram (1989: 76-77) labels these non-productive utterances "unanalyzed wholes" (that is, utterances produced without any knowledge of the internal structure) or "analyzed utterances without productivity" (such as idiomatic expressions).

coherence relation that we know from adult language users. For example, an utterance like (3) would be excluded because the meaning of the sentence is not clear, whereas (4) would be ruled out because the connective marks the wrong relation (the child marks the relation with a contrastive connective instead of a causal one).[10]

(3) Mother: *En 't is na 't avondeten.*
'And it's after dinner.'
Thomas: *Nee.* ***Want*** *vanavond <xxx>.* (Thomas, 2;10.2)
'No. Because tonight xxx (inarticulate speech).'

(4) Mother: *En dan zing je heel erg hard. Nou daar schrikken toch de konijntjes toch van, hè?*
'And then you sing very loud. Well, that makes the rabbits scared, doesn't it?
Hein: ***Maar*** *die gaat weglopen.* (Hein, 2;6.8)
'But that one runs away.'

This approach can be considered an improvement compared to the one in Braunwald (1985: 517), who examines Laura's initial use of a form irrespective of a specific meaning. As a result, her five occurrences of the word *so* also include idiomatic use (*think so*) and intensifier use (*so nice*) (see her examples on p. 518).

The second criterion, that a word should be used to combine clauses, excludes two types of utterances. The first involves the use of a form to link a child utterance to a nonlinguistic event or situation. This *contextual use* (exemplified in (5)) is also excluded by Bloom et al. (1980: 240). The other type is one in which the connective connects two constituents instead of two clauses. This so-called *phrasal use* of words like *and* and *but* will be excluded in Chapter 10, in contrast with Bloom et al. (1980: 240), who do consider examples like (6).

(5) (Kathryn has opened a box of figures and taken them out; picking up a box of furniture)
and let's see dis (= this) (Kathryn, 2;2) (Bloom et al. 1980: 240)

(6) and Mommy's gonna get me chair **and** table (Gia, 2;6) (Bloom et al. 1980: 240)

To exclude phrasal uses as in (7), the sentences introduced by *en* 'and' and *maar* 'but' (the only connectives which could be used this way) should contain both a subject and a verb.

(7) Father: *Is dit een brandweerauto.*
'Is that a fire engine?'
Daan: *Nee, niet brandweerauto* ***maar*** *vrachtwagen.* (Daan, 2;7.24)
'No, not fire engine, but truck.'

Note that the second criterion does not imply that children must create a coherence relation between two clauses within the same utterance. Children can also establish a coherence relation between two clauses in separate utterances (see (8)).[11] Nor does the second criterion even imply that children must create a coherence relation between two clauses they uttered themselves. Children could also establish a coherence relation between their own connective utterance and a clause previously produced by a parent (see (9), an adapted version of (8)).

[10] In this qualitative approach, examples like (3) and (4) are disregarded. However, such examples may prove to be revealing with respect to specific developments during the acquisition of connectives.
[11] In the CHILDES format a period indicates the end of an utterance (MacWhinney 1995: 60).

(8) Niek: *Ik ben al in het ziekenhuis (ge)legen.*
'I have already been in hospital.'
Father: *Ja.*
'Yes.'
Niek: ***Toen** ik nog baby was.* (Niek, 3;10.3)
'When I was still a baby.'
(9) Father: *Jij hebt al in het ziekenhuis gelegen.*
'You have already been in hospital.'
Niek: ***Toen** ik nog baby was.*
'When I was still a baby.'

The crucial factor here is that the child must have uttered the connective clause. This approach is similar to the one in Bloom et al. (1980: 237), but it deviates from the one proposed by Kyratzis et al. (1990: 210) who disregard connective responses to why-questions.

According to the third criterion, the connective must be used in a creative way. *Creative* means that the child creates or formulates the instances of the coherence relation himself.[12] This implies that the connective utterance should not be a fixed expression (i.e. a line from a song like (10) or a frequent utterance from an adult conversation partner) and that it should not be a direct imitation of a sentence previously uttered by an adult, as in (11).

(10) Lines from a Dutch children's song:
Poes Minet ging naar bed, heeft zijn slaapmuts opgezet.
Ging (=gleed) ***toen*** *uit op zijn snuit. Oh, wat een domme domme guit!* (Daan, 3;1.14)
'Puss Minet went to bed, put his nightcap on his head.
Then slipped and fell on his snout. Oh, what a silly silly rascal!'
(11) Direct imitation:
Father: *En toen?*
'And then?'
Niek: ***En toen** xxx niet leuk.* (Niek, 3;8.30)
'And then xxx [= inarticulate speech, JEV] not funny.'

This criterion of creativity should lead to the exclusion of utterances that might be regarded as memorized wholes.

9.3.3 Proportion of the parental target

A qualitatively oriented criterion is useful if the focus is on the emergence of connectives. A qualitative approach also seems appropriate for tracking further connective developments. For example, Brown et al. (1969: 145) propose to consider the pattern of omissions in the total distributional range of a form. However, Brown et al. also have some suggestions for using certain quantitative criteria (rather than using an arbitrary frequency of production). One of the options they mention is to "consider frequencies of forms in contexts that make them obligatory" in the parental target (p. 145). Each of these obligatory contexts can be considered a learning trial for the child, and the degree of acquisition can be measured by computing the proportion of instances in which the child performs appropriately (as this proportion changes

[12] In the acquisition literature (e.g. Ingram 1989: 77) this criterion of *creative* use is often labeled *productive* use. In order to avoid confusion I have chosen to use the term *creative* for the qualitative criterion I propose here; the term *productive* often refers to quantitative emergence criteria (see section 9.3.1).

over time). This method has been applied frequently, both in first and in second language acquisition research.

One of the areas in which proportions of the parental target have been computed repeatedly is the acquisition of inflectional morphology. For example, Brown et al. (1969: 148) only speak of emergence if the child supplies the inflection in at least ninety per cent of the contexts in which they are clearly required.[13] The ordering of morphemes in the L1 acquisition of English in De Villiers & de Villiers (1973) is also based on the lowest sample at which each morpheme first occurs in 90% of obligatory contexts (see Zobl & Liceras 1994: 168). Finally, Andersen (1978), looking at morpheme orders in the L2 acquisition of English by Spanish speakers, gives the implicational order at 80% criterion for correct use.

This method of computing proportions of forms in obligatory contexts seems very appropriate for more or less obligatory parts of the target language (such as verb second and inflectional morphology). For instance, the morphological rules for the insertion of a plural form are very straightforward (at least for English), and the cases in which these rules should be obliged can be easily determined. However, is this method equally appropriate for determining whether certain connectives have been acquired or not? The answer seems to be "no", for two reasons. First of all, connectives (like discourse anaphora) are more or less optional elements. Because coherence relations may remain implicit, there are many cases in which the use of a connective is not obligatory.[14] Furthermore, a coherence relation may be marked with a less specific connective. This results in underspecification, a phenomenon that also occurs frequently in the adult language (cf. Spooren 1997). Another complicating factor is that the types of connectives in the adult target fluctuate per context, i.e. their frequencies may depend on the text type, just as relations do (Sanders 1997). To conclude then, it seems impossible to establish obligatory contexts for the use of connectives, leave alone proportions of connectives in obligatory contexts.

A second reason to dismiss this quantitative method is that these proportions are used because of the rule-based nature of phenomena like inflectional morphology. Again, a quantitative method is being used to guarantee the quality of the child's utterances. In child language, it is likely that (for example) plural word forms appear relatively early. These early-inflected occurrences do not necessarily imply that the child has acquired the "rules for plural formation"; the child may have stored certain plural forms as unanalyzed wholes in the lexicon. Correct performance can thus be reached in a different way. In other words: one correct representation does not necessarily imply correct knowledge, or correct competence. This is why a percentage of appropriate forms in obligatory contexts is being used to measure rule-based acquisition. Referring to a psycholinguistic experiment by Berko (1958), Zobl & Liceras (1994: 165-166) state that "rule-governed representations of linguistic forms manifest themselves in productive occurrence" and that "a low rate of accuracy may relate to difficulty in achieving a rule-based representation." Correct connective use, on the other hand, does not seem to depend on the application of rules similar to the rules for producing plural forms. The conceptual and syntactic properties of each connective have to be acquired separately. Therefore, in the case of connectives it does not seem necessary to set a certain quantitative criterion based on proportions of correct usage. The only danger of disregarding quantitative

[13] The term "emergence" is confusing here. In the work of Brown et al. (1969) this term does not seem to point to the earliest use of inflection morphology, since the 90% correct use almost implies full mastery.

[14] The most optimistic position would be that – given a certain context and a certain word order in the connective clause – a specific connective is the only one that can be used.

criteria is that certain connectives can be stored in the memory lexicon within larger wholes, such as lines from songs. This is why such "fixed expressions" should be excluded in a qualitative approach, using the criterion of creative use.

9.3.4 Developmental curves

Although research into the emergence of connectives is interesting and useful in itself, it does not give a complete picture of the process of connective acquisition. It ignores the developmental process of *decontextualization* (Braunwald 1997). She observes that new abilities often "emerge in a specific and limited context of use. Development involves the gradual generalization of these context-specific abilities to novel contexts of use" (Braunwald 1997: 124). Therefore, only establishing an order of emergence would result in an incorrect 'yes/no'-idealization of the developmental data. If it is obligatory to characterize a connective as either "acquired" or "not acquired", this would not account for 'more or less' considerations (Van den Bergh, Herrlitz & Klein Gunnewiek 1999: 22).

Van den Bergh et al. (1999: 25) argue: "language development can be described more adequately by a continuous scale on which the mastery is allowed to vary from 'no mastery at all' to 'complete mastery'." This is why they favor an analysis based on developmental curves showing probabilities of occurrence. An additional advantage of using developmental curves in tracking the gradual process from emergence to full mastery is that the curves allow the researcher to study the development from two perspectives. The researcher can study mean patterns of development, i.e. the general acquisition pattern per connective, as well as the way individuals deviate from this mean development (Van den Bergh et al. 1999: 31). This is important, since there may be differences in the development of connectives in the level of mastery and in the rate of change over time (cf. Klein Gunnewiek 1999 for observations along this line concerning the L2 acquisition of word order patterns).

Again, the analysis based on developmental curves is a quantitative approach that has to be supplemented with qualitative data of the connective use during the period under investigation. Only qualitative analyses can show whether children make mistakes in the use of certain connectives, or whether they use these connectives in a restricted range of all the meanings they can express.

At the beginning of this chapter, I formulated the question "What is an appropriate way to establish 'acquisition'?" The discussion in section 9.3 has shown that several methods are appropriate for establishing connective acquisition in all its varieties. A full picture can be gained from longitudinal data on connective acquisition by (a) looking at the first correct and creative use of connectives; (b) analyzing intra- and interindividual patterns of acquisition using developmental curves; and (c) analyzing the connective utterances produced at different ages in a qualitative way.

9.4 Previous research into the acquisition of Dutch connectives

In each of the following chapters, I will first discuss some general observations from the literature on the acquisition of connectives, and then present my findings on different aspects of the early acquisition of Dutch connectives. In the current section, I focus on literature concerning the acquisition of Dutch connectives, in order to show that the shortage of data in this area requires a more detailed investigation of this acquisition process. I introduce previous findings on connective acquisition by Dutch children aged 0 to 4 (see section 9.4.1) as well as previous observations on older children (section 9.4.2).

9.4.1 Early Dutch connective acquisition in the literature (0-4 years)

The early acquisition of Dutch connectives has hardly received any attention in the literature. Several studies describing the linguistic development of Dutch children aged zero to six years only devote between three to seven lines to this topic. The remarks in these studies are often very general, only mentioning five or six connectives explicitly, and they are only supported by exemplary connective examples uttered by young children, without any reference to connective frequencies or a description of the methodology used to reach these conclusions. For example, Schaerlakens & Gillis (1987: 154) claim that complex clauses start to appear during their 'differentiation phase' (which starts at about age 2;6). According to them, this is at first restricted to coordinated clauses, but later on also appears in subordinate constructions with *als… dan* 'if… then' and *omdat* 'because'. In the same vein, Bol & Kuiken (1988: 63) claim that clauses coordinated by Dutch *en* 'and' occur from the age of two and a half on. From the age of three coordinations with *of* 'or', *want* 'because', and *maar* 'but', also appear, as well as subordinations (e.g. marked with *als*, 'if'). And finally, Schlichting (1996) mentions that children only become productive in uttering coordinating clauses (with *en*, *maar* and *want*) and subordinating clauses (with *als* and *omdat*) by the age of three and a half (see Goorhuis-Brouwer 1997: 61 for findings and formulations similar to the ones mentioned above).

Equally little is known about the syntactic aspects of Dutch connective acquisition; the syntactic development of connectives has not been investigated systematically. For Dutch, it has been observed that children master the so-called "V2-position" relatively early. This is the functional position that – according to most analyses of adult Dutch – hosts the connective in subordinate clauses (see, among others, De Haan 1987). One remarkable observation is that when very young children start producing subordinate clauses, they sometimes leave out the complementizer (cf. Krikhaar 1992). Some typical examples are given in (12) and (13), in which the Ø-symbol marks the absence of the complementizers *dat* 'that' (which is obligatory in Dutch) and *als* 'when' respectively (examples taken from Krikhaar 1992: 5).

(12) *Ik denk Ø jij boven was.* (Laura, 3;3)
'I think you were upstairs.'

(13) *Ik mag straks Sarah spelen, Ø zij gegeten heeft.* (Laura, 3;5)
'I can play with Sarah soon, she has eaten.'

Why did the Dutch connective acquisition receive relatively little attention? Several factors might play a role. First of all, connective acquisition becomes a more interesting phenomenon after the age of three and a half. However, language acquisition by children aged four or older has received relatively little attention overall. The small number of observations may also be due to the fact that language acquisition has traditionally been studied at different levels of description: phonology, lexicon, morphology and syntax. The study of connective acquisition would fall into the last category. However, a proper investigation of connectives needs to consider discourse as well, going beyond the syntax of single clauses and thus beyond the primary subject of interest investigated by most syntacticians. Furthermore, connectives are in a way not obligatory; the coherence relations they explicate can also be derived without the connectives being present. In studies treating the acquisition of functional categories, focus has often been on obligatory phenomena like determiners and inflectional morphology. Finally, Dutch acquisition researchers who did study the acquisition of functional categories like complementizers, mostly worked within a generative framework. This restricted their field of attention to the absence or presence of certain functional projections in child language

(cf. Wijnen 1997 for an overview of relevant literature); as a result the variety of connectives that can occur in these projections hardly received attention.

9.4.2 Later Dutch connective acquisition in the literature (4 years up)

The scarce data available on Dutch connective acquisition after the age of four have all been gathered in cross-sectional studies. Four studies will be discussed here: Roelofs (1998), Van Hell, Verhoeven & Wengelin (1999), Spooren, Tates & Sanders (1996), and Spooren (1997).

Roelofs (1998) reports on the pragmatic development of children aged four to eight, based on cross-sectional data gathered within a narrative and a conversational task. Across ages, she found stability in the number of coordinating conjunctions, both in the narrative and in the conversational context. For subordinating conjunctions, she observed an increase over time in narrative contexts. However, this increase should be subscribed to an increase in the use of the grammatical subordinator *dat* 'that'; the use of semantic subordinators remains infrequent in all age groups (Roelofs 1998: 149).

Some of these findings are in line with the results of Van Hell et al. (1999), who collected written narratives in four different age groups: 4th graders from 9 to 10 years old, 6th graders from 11 to 12 years old, high school students who were 15 or 16 years old and a group of adults. They performed both qualitative and quantitative analyses of the narratives. Their results show that the age groups did not differ in their use of coordinating conjunctions; the four groups used coordinating conjunctions equally as often and in all different types of coordination: coordinate *en* 'and', adversative *maar* 'but', causal *want* 'because', and consecutive *dus* 'so'. They furthermore conclude that the use of subordinating conjunctions develops as one gets older. Quantitative analyses showed that 4th graders and 6th graders used fewer subordinating conjunctions than the high school students or the adults did.[15] Qualitative analyses revealed that the adult group used all the different types of subordinating conjunctions, whereas the high school students used 64% of these types and the fourth and sixth graders only 45%.

In an experimental study on the acquisition of connectives and relations Spooren et al. (1996)[16] found that the coherence relations between utterances produced by primary school children were strongly determined by the discourse task in which the children were involved: a picture description task resulted mainly in propositional (content) relations; when asked for an opinion the children mainly produced pragmatic (epistemic and speech act) relations. A second finding in their study is that children aged 6-7 produce significantly fewer pragmatic relations than children aged 11-12. With age, children become more proficient in marking argumentative relations with connectives.

Using the same data as Spooren et al. (1996), Spooren (1997) investigated the degree of underspecification in different age groups. His study shows that, with age, children become more specific with respect to their relation marking, taking into consideration the needs of the conversational partner. At age 6-7, 75% of the relations expressed with a connective were underspecified, at age 11-12, 65% (Spooren 1997: 160). It is remarkable that children of both age groups frequently used general connectives, heavily relying on the hearer to infer the correct relation. A more specific use of the connectives occurs only when the child has mastered coherence relations to such a degree as to be able to fine-tune the language code to the needs of the hearer.

---

[15] However, for the temporal (*voordat* 'before', *terwijl* 'while', *nadat* 'after') and causal conjunctions (*omdat* 'because'), no differences among the four age groups were observed.

[16] The results of this experiment are also reported in Spooren & Sanders (2005).

9.5 Preview

Sanders et al. (1992, 1993) have shown that coherence relations and connectives can be classified with four conceptual primitives (see Chapter 2). Can this conceptual system also be used to explain the acquisition of Dutch connectives? From the previous section it becomes apparent that more research into the acquisition of Dutch connectives is needed in order to answer this question.

The acquisition part of this thesis (Part III) follows a bottom-up approach. Given the scarce availability of detailed analyses of connective acquisition by young Dutch children, I will pay much attention to my findings concerning this acquisition process as well as to an explanation of these acquisition facts in terms of cumulative complexity (see Chapters 10-13). It is not until the final chapter of Part III (Chapter 13) that I will explicitly re-address the main question of this thesis.

Chapter 10 focuses on the very beginning of connective acquisition; it introduces English and Dutch acquisition orders based on first emergence (defined in a qualitative way). This chapter investigates the influence of the inherent cognitive complexity of the connectives as a determinant of acquisition.

Chapter 11 treats the overall developments of the Dutch connectives discussed in Chapter 10, showing quantitative analyses and developmental curves of these connectives. Chapter 11 provides additional support for the cognitive complexity approach in Chapter 10, but it also pays attention to the influence of parental input on the acquisition process.

Chapter 12 deals with the overall developments of Dutch connectives as well, but the focus here is on qualitative changes during the acquisition process. This chapter studies the impact of increasing syntactic complexity on the degree to which children integrate their connective clauses into the matrix clauses.

Chapter 13 discusses the acquisition of Dutch causal connectives. I investigate whether young children already make conceptual distinctions based on domains of use. This chapter stresses the importance of taking into account the different contexts in which children produce their utterances. In Chapter 13 I will also re-address the main question of this thesis, showing what the acquisition data discussed in Chapters 10 to 13 reveal about the interaction hypotheses put forward in Chapter 3. In line with the diachronic part of this thesis, the conclusions on specific form-function relations will be based mainly on the data of the causal connectives.

An overview of the methodologies and connective selections in Chapter 10 to 13 is presented in Table 9.3.

Table 9.3. Overview of the methodologies and connective selections in Chapters 10-13

Chapter	Methodology	Connective selection
10	first emergence	en, toen, maar, want
11	overall development in quantitative terms	en, toen, maar, want
12	overall development in qualitative terms	en, toen, maar, want, omdat
13	first emergence and overall development	want, omdat, dus, daarom

Chapter 10

The emergence of connectives

This chapter deals with the emergence of the connectives in Dutch child language. The main question is: Is there a fixed order in which connectives emerge and if so, how can this order be explained? The most comprehensive study in this field, by Bloom et al. (1980), found that for English there is "variation in form with similarity in content in the development of individual children" (p. 260). Bloom et al. explain the similarity in content on the basis of cumulative semantic or conceptual complexity. However, they do not give an account for the variation among children. The current study takes a multi-dimensional approach to the acquisition process, which makes it possible not only to account for the uniformity, but also for the diversity in the developmental sequences of both English and Dutch children.

> *"It would appear that cognitive complexity sets the pace for acquisition, at least in part."* (Clark & Clark 1977: 338)

10.1 Introduction[1]

Most children build their first texts before the age of three; instead of uttering one clause at the time, they start producing combined clauses. At first the coherence relations between these clauses remain implicit. For instance, the contrastive relation in (1) could have been marked with *maar* 'but', and the causal relation in (2) with *want* 'because'.

(1) *Ik wil niet teken(en). Ik wil verven.* (Josse, 2;8.18)
'I do not want to draw. I want to paint.'

(2) *Even liggen. Beetje moe.* (Matthijs, 2;9.15)
'Lay down for a moment. Bit tired.'

To make these coherence relations explicit, children need to acquire connectives, linguistic units that explicate conceptual relations between combined clauses (see (3) to (5)).

(3) *Jij mag niet eh van drop,* ***want*** *dat is van mij!* (Thomas, 2;10.19)
'You can't have uh licorice, because that's mine!'

(4) *Ik wil geen motor.* ***Maar*** *nou wil ik een politieauto.* (Josse, 2;11.23)
'I don't want a motorbike. But now I want a police car.'

(5) *En toen kwam Bugs Bunny* ***en toen*** *gaan ze de banaan pakken.* (Daan, 3;1.14)
'And then Bugs Bunny came and then they go get the banana.'

The current chapter takes up two questions concerning the acquisition of connectives:

[1] Part of the analyses in this chapter have been performed in collaboration with Johanneke Wilson-Birnie (see Wilson-Birnie 2002).

(6) Research questions of this chapter:
   a. Is there a fixed order in which connectives emerge in child language?
   b. And if so, how can this order be explained?

The most comprehensive study in this field, by Bloom, Lahey, Hood, Lifter, and Fliess (1980, reprinted in a slightly revised version in Bloom 1991), found that English children follow the same route in acquiring coherence relations:

(7) additive < temporal < causal < adversative

For the linguistic markers expressing these coherence relations, (sentential) connectives, their data give a less clear picture. Although *and* always appears first, the four children in their study show a variety of developmental patterns for the connectives *and then*, *because*, *so*, and *but*. In other words, for English there is "consistency among children in acquiring content and variation in acquisition of form" (Bloom 1991: 260). They explain the similarity in the development of coherence relations on the basis of cumulative semantic or conceptual complexity: "adversative sentences were all additive in that two events or states were joined; causal was both additive and temporal; some of the adversative sentences were both additive, temporal and quasi-causal" (Bloom et al. 1980: 258). However, they do not give an account for the variation among children.

The current study complements Bloom et al.'s study in two ways. First of all, it enables crosslinguistic comparison by extending the field of connective acquisition with an analysis of data from twelve Dutch children acquiring their native language. Furthermore, it takes a multi-dimensional conceptual approach to the acquisition process, which makes it possible not only to account for the uniformity, but also for the diversity in the developmental sequences of Dutch and English children.

The organization of this chapter is as follows. In section 10.2 several factors are rejected that might account for the variation among children; it also presents an alternative account based on increasing conceptual complexity including hypotheses based on this notion. In section 10.3 these hypotheses are applied to data already available on the acquisition of English connectives; in section 10.4 the same is done for newly analyzed connective data from twelve very young Dutch children. Section 10.5 concludes this chapter with a summary of the main findings and a discussion.

10.2 Explaining orders of appearance

The 1980-article by Bloom and her colleagues is the first and most comprehensive study in the acquisition field that deals with the emergence of both conceptual relations between clauses and connectives, not only looking at semantic properties, but also analyzing syntactic and pragmatic properties. Another positive aspect is that Bloom et al. give a very detailed account of their thorough methodology. Despite these merits, their approach has some drawbacks as well. One objection relevant to the purposes of my study is that Bloom and her colleagues do not account for the variation in their data on connective acquisition. This gives rise to the question of which factor causes this diversity. In section 10.2.1 I discuss and reject several factors mentioned in the literature; in section 10.2.2 I put forward an alternative, multi-dimensional proposal based on increasing conceptual complexity. This approach can account for both the uniformity and the diversity in the orders of emergence.

10.2.1 Explanations for variation in orders of emergence

Bloom et al. found variation among children in the orders of emergence of English connectives. In the literature concerning variation in language acquisition in general, several factors have been proposed to account for different acquisition tracks. Below, I discuss these factors and argue why they seem less suited to account for the diversity found in the data on connective acquisition.

Ingram (1989: 77-79) treats three possible sources of variation among children: (a) performance variation, (b) linguistic variation, and (c) environmental variation. Performance variation is due to differences in "individual capacities or abilities of the child that lead to preferences for, or better skill at, particular linguistic subsystems" (p. 77). This factor is unlikely to be responsible for the diversity in the development of connectives, since existing child language literature (e.g. Brown 1973) has often found that performance variation leads to variation in age of emergence, but not in order of appearance.

Linguistic variation is the second factor that might cause developmental differences. Linguistic variation is due to the range of possibilities allowed by the grammar of a certain language.[2] This factor can also be ruled out: the only grammatical alternatives in the case of connectives are (a) choosing another connective (or cue phrase) or (b) leaving the coherence relation implicit. It is not likely that children will choose the first option: using other connectives that are suitable to express the intended coherence relation. The connectives that are at the heart of this study are among the very first connectives that children acquire, which implies that linguistic alternatives are simply not available. For example, Evers-Vermeul (2000) has shown that *want* is the first causal connective in Dutch child language and that only later do other causal alternatives like *omdat* and *daarom* appear (see also Chapter 12 on the acquisition of causal connectives). It is clear that this alternative does not lead to a proper understanding of variation among children. The second option based on linguistic variation, leaving the coherence relation implicit, demands a deeper explanation. Why should children not use a certain connective by leaving the coherence relation implicit, while they already use other connectives to explicate relations that are conceptually more complex?

The third factor mentioned by Ingram (1989: 78) is environmental variation or differences in the input language. It is possible that environmental variation is the cause of variation among children, since children need to hear connectives in their parents' language to be able to acquire them. Without input, acquisition cannot take place. The influence of parental input will be taken up for discussion in Chapter 11. The current chapter will start from the idea that environmental variation need not be a major factor in determining the order of acquisition. First of all, the connectives at hand are the linguistic markers of very basic conceptual notions such as causality and polarity, and it remains to be seen whether parents differ significantly in the frequency with which they use these basic connectives.

To conclude, the factor of environmental variation cannot be ruled out completely, but the two other factors, performance variation and linguistic variation can be put aside as possible determinants of the order of emergence of connectives. In the following subsection I present an alternative account, working out the hypothesis that both the diversity and the

---

[2] Ingram (1989: 78-79) himself defines linguistic variation as being due to the range of structural possibilities allowed by Universal Grammar. Because it is not clear to me which parameters should be involved in connective acquisition or how different parameter settings would lead to different acquisition orders for connectives, I interpret linguistic variation in a somewhat broader sense: the variation that is due to different linguistic possibilities within the language.

uniformity in the developmental connective sequences can be explained by reference to the notion of increasing conceptual complexity.

10.2.2 Cumulative complexity as an explanation for variation

A classical, pioneering precedent for the notion of cumulative complexity as a determinant of order of acquisition of a set of forms is Brown (1973). He introduces the notion extensively, in connection with the question why fourteen English grammatical morphemes are acquired in a particular order (see Brown 1973: 254 ff.). He looks at both cumulative semantic complexity and cumulative grammatical complexity, and compares them, as predictors of order of acquisition, to each other and to the relative frequency of the forms in the parental input.

In Chapter 11, I will a first exploration of the role of parental input, and in Chapter 12, I will address the influence of cumulative syntactic complexity on the acquisition process. In the current chapter, I focus on cumulative semantic or conceptual complexity as a determinant of connective acquisition. My approach can be regarded as a multi-dimensional elaboration of the cumulative complexity explanation that Bloom et al. (1980: 258) give for the uniformity in the development of coherence relations. It starts from the idea that there is no inherent need for one fixed order of emergence, but that the interaction between the relevant conceptual primitives themselves leaves room for variation.

To give a more refined account of the relevant semantic or conceptual primitives involved in the process of connective acquisition, I base myself on the cognitive approach of coherence relations as advocated in Sanders, Spooren & Noordman (1992, 1993). They claim that all coherence relations can be described with four binary cognitive primitives: *basic operation*, *polarity*, *source of coherence*, and *order of the segments*. Per primitive one relatively easy and one relatively complex value can be distinguished, which make these primitives useful for an operationalization based on increasing conceptual complexity.

In this section, I use three of these primitives to derive hypotheses about the acquisition order of connectives, because it is likely that a psychologically plausible classification of coherence relations can also be used for the linguistic counterparts of these cognitive entities.[3] Spooren & Sanders (2005) have convincingly shown the relevance of each of these cognitive primitives in accounting for acquisition data.[4] However, their work focused on 6 to 12 year old children; they conclude that an analysis of younger children is needed to test the influence of cognitive complexity on the early acquisition process. In addition, work by Piaget (1969) and others has shown that three year olds have not mastered all types of coherence relations, despite the fact that notions like addition and causality are very basic (compare children's early holophrastic utterances like *ik ook* 'me too', which occur long before children start producing complete utterances). The general idea here, then, will be that conceptual complexity based on the three cognitive primitives is a major determinant in the way young children acquire the linguistic markers of coherence relations.

The relative conceptual complexity can be thought of in terms of processing cost: the production of a relatively complex coherence relation involves a higher 'processing cost' than the production of a relatively easy coherence relation. It is to be expected that in complex

---

[3] Their fourth primitive, *order of the segments*, is disregarded here. This primitive is only relevant for causal connectives and it also interacts with another primitive: *source of coherence* (see Spooren, Sanders & Visser 1994; Spooren, Tates & Sanders 1996).

[4] This 2005-article largely draws from data presented earlier in Spooren et al. (1994) and Spooren et al. (1996).

relations less energy is left for the production of a linguistic element that explicates this relation. Support for this view comes from two sources, both of which show that not only does conceptual complexity play a role during the first acquisition of concepts, but also at a later stage. Firstly, in an experiment eliciting coherence relations and connectives, Spooren, Tates & Sanders (1996) found that the use of markers expressing negative causal relations does not increase with age (both 6-7 year old and 11-12 year old Dutch children continue using negative causals relatively infrequently), even though these children have acquired the notions of causality and polarity separately. This can be explained on the basis of the higher complexity of the combination of these two notions. Furthermore, even adult second language learners – who are very much capable of understanding all the different coherence relations – initially produce underspecified connectives or no connectives at all in their acquisition process, and only later start to produce relation-specific connectives (cf. Spooren 1997). The higher 'processing cost' of more complex relations plays a lasting role. The relative complexity of coherence relations can thus also be used to establish the relative complexity of connectives.

Below, I present each primitive and also discuss how they result in differences in conceptual complexity. The first primitive is the *basic operation*, which makes a distinction between additive and causal relations, such as (8) and (9).

(8) *Kim is een meisje.* ***En*** *jij bent een jongetje he?* (Thomas, 2;8.23)
'Kim is a girl. And you are a boy, aren't you?'

(9) *Ik heb (een) beetje griep.* ***Want*** *ik ben laat gaan slapen.* (Abel, 3;3.8)
'I've got a touch of flu. Because I went to bed late.'

Because a causal relation can be seen as an additive relation with an implication relation added to it, a causal relation is regarded as more complex than an additive relation. This can be nicely represented in terms of features: an additive connective like *en* is underspecified for the feature 'causal' (represented as [α causal]), whereas a causal connective like *want* is positively specified for the same feature (represented as [+ causal]).[5] In line with this difference in conceptual complexity, the following prediction can be made concerning the acquisition of additive versus causal connectives (see (10)). The fact that causal use in prediction (10) is placed below additive use means that the first causal connective should not appear before an additive connective has occurred.

(10) <u>Prediction based on the *basic operation*:</u>

additive [α causal]

↓

causal [+ causal]

The second primitive is *polarity*, which distinguishes between positive ([α negative]) relations as in (11) and negative ([+ negative]) relations such as the one in (12).

[5] This idea of an underspecification in features (represented with an α) is in line with the observation (mentioned in section 2.1 in Chapter 2) that certain 'underspecified' connectives can occur in more specific coherence relations (e.g. additive *and* can be used to express a temporal relation). Hence, choosing an underspecified feature instead of a negative one cannot simply be regarded as notational variation.

(11) Parent: *Nou rijdt 'ie.*
'Now it (lit. 'he') drives.'
Peter: ***En** nou gaat 'ie in het schuur.* (Peter, 2;3.7)
'And now it goes into the barn.'

(12) *'k Wou bij oma een molen maken, **maar** dat kon ik niet.* (Laura, 4;9.10)
'I wanted to make a mill at grandma's, but I couldn't.'

The negative relation can be classified as more complex than the positive. Both clause complexes in (11) and (12) state something; in addition, the negative relation denies a relation between the propositions contained in (12). For connectives, this results in the prediction (see (13)) that the first negative connective should not appear before a positive connective has occurred.

(13) Prediction based on *polarity*:

positive [α negative]
↓
negative [+ negative]

For the two separate primitives, the conceptual complexity account seems to result in relatively strong claims, which predict uniformity in the acquisition processes. The diversity arises when the interaction between these factors is taken into account. The two primitives – *basic operation* and *polarity* – do not operate independently, since each connective can be characterized by both primitives. For example, a positive additive is unspecified for both features, whereas a negative causal is positively specified for both features. My prediction is that negative causals are the most complex, given their double specification, and so should appear last, after negative additives and positive causals, which are both only specified for one feature. This is depicted in the diagram in (14) by placing the negative causal use below both negative additive and positive causal use. Negative additives and positive causals are placed on the same level: they are not ordered in relation to one another, because my complexity theory does not make predictions about the relative complexity of the feature 'negative' versus the feature 'causal'. Of these two, children can either first acquire a negative additive connective or a positive causal connective. However, they should both occur only after a positive additive connective has entered the language of the child, since positive additive is unspecified for both features. The diagram in (14) shows that my complexity theory leaves room for variation in the developmental sequence.

(14) Prediction based on the interaction between *basic operation* and *polarity*:

positive additive
[α negative, α causal]
↙ ↘
negative additive **positive causal**
[+ negative, α causal] [α negative, + causal]
↘ ↙
negative causal
[+ negative, + causal]

A similar story holds for the interaction between *basic operation* and yet another conceptual factor: *temporal order*. Additive relations may or may not show a temporal ordering of the

segments. Similarly, causal relations are not necessarily based on a temporal ordering of the segments. The utterances in (15) and (16) give examples of a temporal causal relation (in this case marked with *and*) and a non-temporal causal relation marked with *cause* (= (13) and (12) in Bloom et al. 1980: 244).

(15) She put a band-aid on her shoe **and** it maked it feel better. (Peter, 3;2)
(16) Get them **cause** I want it. (Eric, 2;5)

From the point of view of cumulative complexity I predict that both temporal additives and non-temporal causals will appear after the first use of a purely (non-temporal) additive connective. In addition, the emergence of temporal additives and non-temporal causals is ordered with respect to the emergence of temporal causals, but not in relation to one another (compare the diagram in (17)). Again, the interaction between the two primitives predicts both uniformity and variety in the developmental sequences of individual children.

(17) <u>Prediction based on the interaction between *basic operation* and *temporal order*:</u>

non-temporal additive
[α temporal, α causal]

↙ ↘

temporal additive
[+ temporal, α causal]

non-temporal causal
[α temporal, + causal]

↘ ↙

temporal causal
[+ temporal, + causal]

The distinction between temporal and non-temporal use of causal connectives can be related to a third primitive in the work of Sanders et al. (1992, 1993): *source of coherence*. This primitive distinguishes between semantic relations (between locutions, i.e. descriptions of real world events or states of affairs) and pragmatic relations (between illocutions, i.e. reasoning or motivations for performing speech acts). Semantic relations (as in (15)) also involve a temporal ordering of the causally related segments, whereas pragmatic relations (as in (16)) do not necessarily involve such a temporal relation. In terms of *source of coherence*, pragmatic causals may also occur directly after the first additive connectives have emerged, because pragmatic use operates independently of the notion of temporality. My hypothesis in (17) differs from the one put forward by Bloom and her colleagues in that the conceptual notions are not ordered along one dimension (temporal < causal), but rather so that each connective is defined on the basis of two separate primitives.[6]

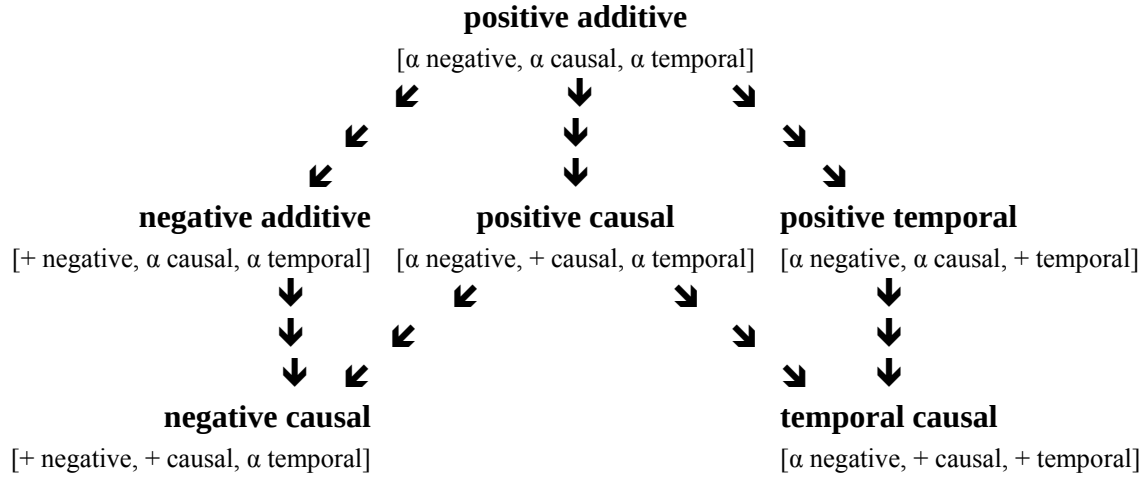
To conclude, then, my hypotheses can be summarized as in (18).[7] Again, placement on the same level implies indeterminacy in order of emergence, whereas placement on a lower level implies that the connective should emerge later than connectives on a higher level.

---

[6] In Chapter 13, this hypothesis about the relative complexity of semantic versus pragmatic relations is worked out in more detail for Dutch causal connectives. That analysis starts from the more refined tripartation based on domains. The discussion here is restricted to the distinction between causal relations that are based on a temporal ordering of the segments versus relations that are non-temporal. This focus on the presence or absence of temporal order in causal relations allows me to formulate precise hypotheses based on cumulative complexity.

[7] In this diagram, the combination of negative and temporal (which would be placed on the lowest row) is disregarded. An example of a connective involving both a negative and a temporal relation is

(18) <u>Hypotheses based on the interaction between *basic operation*, *polarity*, and *temporal order*:</u>



For ease of reference, I have split this diagram into four separate hypotheses (see (19)). The notation "A ≥ B" means that A will not emerge before B has occurred. Sentence symbols separated by commas (e.g. A, B) are to be read as not ordered relatively to one another.

(19) <u>Hypotheses based on increasing cumulative complexity:</u>

a. first causal connective ≥ additive connective
   (= the first causal does not appear before an additive has occurred)
b. first negative connective ≥ positive connective
c. first negative causal ≥ positive causal connective, negative additive connective
d. first temporal (= semantic) causal ≥ (non-causal) temporal ≥ additive connective

The first two hypotheses predict uniformity in the developmental sequences; the other two hypotheses leave room for variation among children.

Hypotheses similar to the strong hypotheses in (10) and (13) have already been put to a test – not for connectives, but for coherence relations – in experiments by Spooren & Sanders (2005) among Dutch children from grade 1 (aged 6-7) and grade 6 (aged 11-12) of an elementary school. Spooren & Sanders (2005: 20) tried to derive an order of acquisition by looking at the frequencies with which both age groups used the more complex relations in a description and a conversation task. They expected that older children would use causal, negative, and negative causal relations (more complex relations than their additive and/or positive counterparts) more frequently than younger children. This turned out not to be the case: older children did not use complex relations like causals more often than younger children. As Spooren & Sanders (2005: 30) indicate themselves, this finding need not be in conflict with the complexity idea, since it is very much possible that even the younger children in their study had already acquired the concept of causality. In line with this possibility Spooren & Sanders suggest: "the utterances of younger children should be studied

---

Dutch *terwijl* 'while'. The temporal relation it expresses is one of simultaneity instead of sequence (which is discussed here). It appears that children acquire this connective relatively late: it is not attested to at all in any of the Dutch corpora.

Another combination that is disregarded in this diagram, is the combination negative temporal causal (with a positive specification of all three features). This would be the most complex combination. However, I do not know of any Dutch connective that forces this specific interpretation.

to find the kind of differences under discussion" (2005: 30). This is exactly what is done in the following two sections. In order to test the hypotheses in (18), which are based on the interaction between the different cognitive primitives, I have performed two analyses. First of all, I have re-examined the English data in Bloom et al. (1980) and two other observational studies on English. Secondly, I have analyzed data on the acquisition of Dutch connectives by very young children (with ages ranging from 1;5 to 5;6).

10.3 The emergence of English connectives

As a first exploration, this section discusses a re-examination of observational data concerning the acquisition of English connectives in terms of my hypotheses. These data are taken from three longitudinal studies that mention a developmental order of early connectives: Bloom et al. (1980), Braunwald (1985) and Diessel (2004).[8] More information on these studies is given in Table 10.1.

Table 10.1. Information on three studies providing the English connective data

Source	Type of data	Relevant connectives
Bloom et al. (1980)	Recordings (± 8-week intervals)	and, (and) then, when, because, so, but
Braunwald (1985)	Diary notes (1-day intervals)	and, when, because, so, but
Diessel (2004)	Recordings (± 2-week intervals)	and, when, because, so, but

As can be seen in Table 10.1, the selection of connectives is not equal across the three studies: whereas Braunwald and Diessel only study the temporal *when*, Bloom et al. also take into account the connective *then* (both in combination with *and*, and as a separate connective).[9] However, as the acquisition orders in Table 10.2 show, this difference need not be problematic: in the data taken from Bloom et al., *and then* and *when* often emerge at the same time, without resulting in different orders of emergence with respect to the other connectives.

Table 10.2. English orders of emergence per child[10, 11]

Child	Researcher	Order of emergence
Eric	Bloom et al.	and < because < so, and then, when < but < then
Gia	Bloom et al.	and < and then, when, then < because, but < so
Kathryn	Bloom et al.	and < and then, so, then < because, but, when
Peter	Bloom et al.	and < because < and then, when < so, but < then
Laura	Braunwald	and < because < when < so < but
Adam	Diessel	and < because, so < but < when
Naomi	Diessel	and < because < but < so < when
Nina	Diessel	and < because < when < but, so
Sarah	Diessel	and < because < so < but < when

[8] All researchers also study some more connectives, e.g. *if*, *for*, *while*, *before*, *after*. Since all of these appear later than the connectives mentioned in this table, these other connectives are disregarded here.
[9] Apart from type of data and connective selection, the three studies differ in a third respect, namely the criterion for determining the date of emergence (see the methodological discussion in Chapter 9).
[10] The orders of emergence of Eric, Gia, Kathryn and Peter are taken from Figure 1 in Bloom et al. (1980: 242). Laura's data come from Braunwald (1985: 513, 518) and the data from the other children are taken from Table 7h *Appearance of the children's conjoined clauses* in Diessel (2004: 198).
[11] Diessel (2004) shows the developmental order of another child: Peter. Because Bloom et al. have already studied this child, these data are disregarded here. See section 9.3.1 in Chapter 9 for a comparison of Diessel's and Bloom's analyses.

The first two, relatively strong predictions (the first causal connective and the first negative connective should not occur before an additive connective has been acquired), are borne out by all child data. Additive *and* is the first connective for all nine children, resulting in a later acquisition of both the earliest causal, *because*, and the earliest negative connective *but*.

My third hypothesis – based on the interaction between polarity and basic operation – is that the first negative causal connective should not appear prior to the occurrence of both a positive causal and a negative additive connective. In all child data, the negative connective *but* does not appear before the first use of the causal *because*, which is in line with the first half of my hypothesis. Note that the data from Gia and Kathryn – where *but* and *because* emerge at the same age – do not provide evidence against my hypothesis: increasing complexity leaves room for simultaneous emergence of forms. The second half of this hypothesis cannot be tested on the basis of the data in the three studies. The reason for this is that none of the studies distinguishes between the negative additive and the negative causal use of *but*. For one child, Laura, the type of first negative use can be derived from the first utterance containing the contrastive *but* as mentioned in Braunwald (1985: 518, Table 2).

(20) (Context: Laura is watching her mother Sue write with a red fountain pen.)
Laura: Do you have a fountain pen, daddy?
Father: Yes.
Laura: **But**, it has a blue lid. Sue's has a red lid. (Laura, 2;6.19)

The negative relation in this fragment can be interpreted as negative additive, providing evidence in line with my hypothesis. This method is not possible with the other two studies, since these do not systematically provide examples of the first use of the connectives. The few *but*-examples they do mention are additive ones, such as (21)-(23).

(21) Adult: It is called the skin of the peanut.
Naomi: **But** this isn't the skin. (Naomi, 2;11) (Diessel 2004: 164)
(22) Adult: David doesn't shave yet.
Adam: Uhuh. **But** I shave. (Adam, 3;8) (Diessel 2004: 164)
(23) Cause I was tired, **but** now I'm not tired. (Kathryn, 2;11) (Bloom et al. 1980: 245)

Although these *but*-examples are suggestive, they do not provide real evidence for my hypothesis, since – as the age of Adam in (22) already indicates – they are not the very first usages of *but* by these children. The fact that none of these studies mentions the emergence of the connective that can only be used as a negative causal, *although*, suggests that my hypothesis should be accepted. Diessel (2004: 151) makes the same observation, noticing that "*although*-clauses did not occur in any of the corpora that have been examined" (cf. Clark 1970, 1973; Clancy, Jacobsen & Silva 1976; Bloom et al. 1980; Eisenberg 1980; Lust & Mervis 1980; McCabe et al. 1983; Braunwald 1985; Peterson & McCabe 1985). Piaget (1969: 37-54) even claims that connectives like *although* are not fully acquired before the age of eleven.

My fourth and last hypothesis concerns the interaction between basic operation and temporal order. It predicts that the first temporal causal should not appear before the first temporal additive (and the first purely additive connective). The fact that two children, Gia and Kathryn, produce *and then* before they start using *because*, is in line with this prediction. Kathryn's data also contain early usage of *so*. The early example in (24), which exhibits

temporal causal use of this connective, emerges at the same age as *and then*. Since *so* does not appear before the first temporal, it does not provide counterevidence to the third hypothesis.

(24) Maybe you can bend him **so** he can sit. (Kathryn, 2;5) (Bloom et al. 1980: 244)

The early usage of *because* by the other seven children might provide counterevidence, depending on the (temporal or non-temporal) nature of the relation these *because*-fragments mark. Unfortunately, the data at hand again do not include examples of the very first usages of *because*.

To conclude, the reanalysis of the English data at hand does not present counterevidence to my hypotheses. This suggests that my increasing conceptual complexity theory, which predicts both uniformity as well as diversity, is on the right track. However, the re-examined English data do not provide information that is detailed enough to draw any firm conclusions about my multi-dimensional approach. Therefore, in order to test my hypotheses about the interaction between the conceptual primitives, I have examined data on the connective acquisition of young Dutch children in the age range of 1;5 to 5;6. This longitudinal analysis nicely complements previous experimental analyses of the connective use by Dutch children aged six and older (see Spooren et al. 1996; Spooren 1997; Roelofs 1998; Van Hell et al. 1999).

10.4 The emergence of Dutch connectives

In this section, I successively discuss the Dutch connectives selected for my study, the corpora that provided the child data and the methodology followed during the analyses. The results are then held up to the complexity hypotheses as laid out in section 10.2.

10.4.1 Connective selection and materials

My research focuses on four of the most frequent Dutch connectives (in adult language, cf. Uit den Boogaart 1975), which represent all the conceptual primitives mentioned above (see Table 10.3). These connectives equal the earliest connectives mentioned in the English studies in section 3, which makes it possible to compare the Dutch results to the English data.

Table 10.3. Dutch connective selection

Connective	English equivalent	Basic operation	Polarity	Temporality
en	and	additive	positive	temporal / non-temporal
maar	but	additive / causal	negative	non-temporal
toen	then/when	additive	positive	temporal
want	because/for	causal	positive	temporal / non-temporal

I have examined transcriptions of spontaneous speech data of twelve monolingual Dutch-speaking children. All these materials are available through the Child Language Data Exchange System or CHILDES (MacWhinney 2000). The transcriptions in these longitudinal corpora are based on audiotape recordings made at home, in an unstructured home setting. The recordings were made with relatively small (often 2-week) intervals. This is a significant improvement compared to the English study by Bloom et al., in which approximately 8-week and 12-week intervals separated each of the observations that provided the data for analysis (Bloom et al. 1980: 259). The relatively high density allows me to give a more precise characterization of the development, probably with fewer connectives emerging at the same

time. The children's age ranges are given in Table 10.4, which also shows the total number of utterances produced by each child.

Table 10.4. Dutch corpus data (with ages in years;months.days)

Child	Age range	Number of utterances	Corpus
Abel	1;10.30 – 3;04.01	11883	Groningen[12]
Daan	1;08.21 – 3;03.30	15229	Groningen
Hein	2;04.11 – 3;01.24	12781	Utrecht[13]
Iris	2;01.01 – 3;06.15	8771	Groningen
Josse	2;00.07 – 3;04.17	12651	Groningen
Laura	1;09.04 – 5;06.12	22323	Van Kampen[14]
Matthijs	1;10.13 – 3;07.02	19864	Groningen
Niek	2;07.00 – 3;10.17	15151	Wijnen[15]
Peter	1;05.09 – 2;08.22	8578	Groningen
Sarah	1;06.16 – 5;02.13	17458	Van Kampen
Thomas	2;03.22 – 2;11.22	12670	Utrecht
Tomas	1;07.05 – 3;01.02	9126	Groningen

10.4.2 Methodology

To determine the Dutch orders of emergence, a method is needed to establish whether a connective has been 'acquired'. As I have argued in Chapter 9 (see section 9.3.1), a quantitative approach (like the productivity criterion set by Bloom and colleagues) is not very suitable to measure the emergence of connectives. At this very early stage of the acquisition process children will not be fully 'productive' in their use of connectives in the sense that they are not yet able to use a connective in a variety of meanings and in a variety of contexts.

In order to establish the Dutch order of emergence, it seems reasonable to stay as close as possible to the earliest emergence by using first occurrence complemented with certain qualitative criteria, as mentioned in (25) (see also section 9.3.2 in Chapter 9).

(25) <u>Method used to establish the emergence of a connective:</u>
First occurrence in which the connective is being used
a. in a correct way
b. as a word combining two clauses
c. in a creative way.

The first criterion implies that uninterpretable or wrong connective use is disregarded. For the connective *maar* 'but', the first criterion implies that utterances such as in (26) and (27), in which this word does not clearly mark a contrastive relation, are disregarded.

(26) Mother: *Er ligt nog (ee)n stukje koekje op jouw beker. Dat is van jou.*
'There is still a cookie on your mug. That's yours.'
Matthijs: *Eh!* ***Maar*** *ik lus(t) ook ee(n) koekje!* (Matthijs, 2;10.8)
'Eh! But I also like a cookie!'

[12] See Bol (1996) and Wijnen & Verrips (1998).
[13] See Wijnen & Elbers (1993).
[14] See Van Kampen (1997).
[15] See Wijnen & Elbers (1993).

(27) Mother: *Is de trompet op de slaapkamer?*
'Is the trumpet in the bedroom?'
Thomas: *Ja,* ***maar*** *Loekie wil zoeken.* (Thomas, 2;7.2)
'Yes, but Loekie has to look.'

The second criterion excludes the contextual use of connectives (in which the child chains the utterance to a nonlinguistic event or situation) as well as the phrasal use of *en* and *maar*. For the latter connectives, only clauses containing a subject and a verb are taken into account. The third criterion excludes fixed expressions (e.g. lines from a song) and direct imitations from the analysis. This criterion of creativity should lead to the exclusion of utterances that might be regarded as memorized wholes.

10.4.3 Results

Table 10.5 shows the Dutch acquisition orders per child based on onetime correct and creative clausal connective use, together with the ages at which these connectives emerge. The ordering of the children is such that children following the same route are grouped together.

Table 10.5. Dutch orders of emergence per child (with ages in years;months.days)

Child	1st connective	2nd connective	3rd connective	4th connective	not acquired
Daan	en (2;4.0)	maar (2;5.11)	toen/want (3;1.14)		
Josse	en (2;8.04)	maar (2;11.9)	toen/want (3;0.20)		
Laura	en (2;2.10)	maar (2;7.19)	toen (3;4.21)	want (3;4.25)	
Niek	en (3;4.9)	maar (3;8.2)	toen (3;8.30)		want (>3;10.17)
Peter	en (2;3.7)	maar/toen (2;4.12)		want (2;8.22)	
Iris	en (3;1.0)	maar (3;1.14)	want (3;2.11)	toen (3;3.23)	
Abel	maar (2;3.23)	en (2;4.9)	want (2;10.0)	toen (2;11.10)	
Matthijs	en (2;4.24)	want (2;11.19)	maar (3;0.9)	toen (3;0.20)	
Tomas	en (2;5.7)	want (2;10.10)	maar (2;10.24)		toen (>3;1.2)
Hein	en (2;4.14)	toen (2;5.19)	maar (2;6.10)	want (2;8.28)	
Thomas	en (2;3.23)	toen (2;7.1)	maar (2;7.20)	want (2;10.19)	
Sarah	en (1;11.15)	toen (2;4.2)	want (2;9.7)	maar (3;0.19)	

How do these results relate to the hypotheses we formulated? My first prediction – based on the basic operation – is that the first causal connective does not appear before an additive connective has occurred. This prediction is borne out: all children start with the additive *en* and only later come up with the causal connective *want* (compare as an illustration Thomas' ages at his first production of *en* (28) and *want* (29)).

(28) Mother: *Ja in mijn bord zit pap.*
'Yes in my plate there's porridge.'
Thomas: ***En*** *Loek eet de appelsap.* (Thomas, 2;3.23)
'And Loek eats the appel juice.'
(29) *Jij mag niet eh van drop* ***want*** *dat is van mij.* (Thomas, 2;10.19)
'You may not eh have licorice because it's mine.'

The second prediction – based on polarity – claims that the first negative connective does not appear before a positive connective has occurred. Most of the data are in line with this

prediction: eleven children start with the positive *en*, and only later on produce the negative connective *maar*. Abel forms a remarkable exception to this acquisition pattern; contrary to my prediction, his first connective is *maar* (for example, compare Abel's ages at the time of his first *maar* in (30) and his first clause combining *en* in (31): 2;3.23 versus 2;4.9). In (30), Abel's *maar*-utterance can be seen as a negation of the adult's claim that "Abel can leave the radio where it is": the fact that Abel wants to build a tunnel in that place, implies that he cannot leave the radio over there. In this fragment, then, *maar* marks a real contrastive relation.

(30) Adult: *Je kan hem* (= een radio) *toch gewoon laten staan?*
'Can't you just leave it (= a radio) there?'
Abel: *Nee.*
'No.'
Adult: *Jawel.*
'Yes.'
Abel: *Nee.*
'No.'
Adult: *Nou…*
'Well...'
Abel: ***Maar** ik moet even daar ee(n) tunnel bouwen.* (Abel, 2;3.23)
'But I just have to build a tunnel there.'
(31) Mother: *En het nijlpaard ging ook onder de douche, hè Abel? Ging 'ie drinken.*
'And that hippo also took a shower, didn't he Abel? He drank.'
Mother: (Two side remarks to the researcher, who is also present.)
Abel: *En [/]* ***en** die nijlpaard moet poetse(n).* (Abel, 2;4.9)
'And the hippo has to brush.'

This implies that my second prediction is not completely borne out, since there is one counterexample I cannot explain.

A third prediction is that – because of different interactions between the basic operation and polarity – there is room for variation in the developmental sequences: after the first positive additive it is possible to encounter either a negative additive or a positive causal. Negative causals, however, should occur last. Nine children acquire the negative additive *maar* before the positive causal *want*, while only three children show the reverse sequence. Considering this fact, the conclusion can be drawn that variation due to different interactions can be borne out. A more detailed analysis of the early occurrences of *maar* reveals that these are all instances of negative additive use. The negative causal use of *maar* (as in (32)) only emerges after the first positive causal (or in the same file, as is the case for Peter). The negative causal nature of (32) can be illustrated by reformulating the relation with the negative causal connective 'although' ("Although the barrier does not open, he can go through").

(32) *Slagboom gaat niet open.* ***Maar** hij kan wel erdoor.* (Peter , 2;8.22)
'Barrier does not open. But he can go through.'

My fourth prediction concerns the interaction between the basic operation and temporal order: purely additive connectives should occur first. Additive relations which also show a temporal order will appear later and causal relations based on temporal order will appear last. The

prediction that pure additives appear before temporal additives is borne out for all twelve children; all children only come up with *toen* after they have produced *en* (compare, for example, Matthijs' ages at the moment of his first production of *en* and *toen*).

(33) ***En*** *eh dat is een schoen.* (Matthijs, 2;4.24)
'And eh that is a shoe.'
(34) *En* ***toen*** *waren dieren wakker en* ***toen*** *waren dieren slaap.* (Matthijs, 3;0.20)
'And then animals were awake and then animals were asleep.'

However, there is variation in the acquisition order of temporal versus causal. Eight children show the developmental sequence *toen* – *want*[16], whereas four children produce *want* before *toen*. Whether these last findings run counter my hypothesis, depends on the type of relation marked by *want*. In cases where the causal relation is not based on a temporal relation, the cumulative complexity predicted for additive – temporal additive – temporal causal does not hold. This seems a plausible explanation: three of the four children who produce *toen* after their first *want*, start with a causal relation that is not based on a temporal relation. For example, the *want*-clause in (35) is a pragmatic one: it provides the reason for asking the question. Hence, the *want*-clause does not function at the semantic level, but at the pragmatic level, where temporal order of the clauses is irrelevant.

(35) *Wil je even m'n haar borstelen?* ***Want*** *ik heb slordig haar.* (Iris, 3;2.11)
'Could you brush my hair? Because my hair is untidy.'

Again, Abel is the remarkable exception. He is the only child who produces a temporal causal *want* before he utters his first creative temporal additive *toen*. In other words, in terms of complexity it seems as if he skips a step in the acquisition process. This is probably too strong a claim, since there are two earlier instances of *toen* (see (36) and (37)), which both occur before the occurrence of Abel's first *want*. These occurrences, which are only semi-creative, indicate that Abel must have some idea of the notion *temporal order* at an earlier age.

(36) Adult: *En toen?*
'And then?'
Abel: *Toen ga voetballen.* (Abel, 2;4.23)
'Then go (and) play football.'
(37) Adult: *En toen?*
'And then?'
Abel: *En toen is eh ijs op.* (Abel, 2;7.29)
'And then there's no ice left.'

It seems as if Abel's acquisition route deviates in more than one respect from those of the other eleven children. Further research is needed to identify whether some other factor triggers his remarkable developmental sequences. The acquisition data of the other children provide ample support for the four hypotheses I put forward in section 10.2.2.

[16] This includes Daan and Josse, who start producing *toen* and *want* at the same time.

10.5 Conclusion and discussion

Returning to the questions asked at the beginning of this chapter, I am able to state that there is an order of acquisition, but that for the Dutch connectives *en, maar, toen*, and *want*, both a fixed and a variable part should be distinguished in the acquisition route. This order can be explained on the basis of the multi-dimensional approach to cognitive complexity, as worked out in section 10.2.2. This approach differs from the one put forward by Bloom and her colleagues in that the conceptual notions are not ordered along one dimension (additive < temporal < causal < adversative), but rather so that each connective is defined on the basis of several cognitive primitives (taken from Sanders et al. 1992). The fixed routes are explained by reference to the relative complexity of different values on the same primitive (e.g. [α causal] vs. [+ causal]), whereas the variation among acquisition routes of Dutch and English children can be explained by reference to the different interactions between the conceptual primitives that characterize each connective. All in all, cumulative conceptual complexity seems to offer a solid explanation for the findings both on Dutch and English connective acquisition. My multi-dimensional approach accounts both for the uniformity and for the diversity in this acquisition.

The multi-dimensional approach needs to be extended in order to give an account for the emergence of other connectives as well. For example, within the temporal domain, further distinctions are needed to account for connectives expressing simultaneity, and within sequentiality between priority and anteriority.

It is likely that conceptual complexity is not the only factor that determines the acquisition process. For example, Brown (1973) and Slobin (1973) have already argued that formal or syntactic complexity plays a role as well. In the current chapter this factor was disregarded, since three of the four selected connectives (*en*, *maar*, and *want*) exhibit a similar syntactic complexity in that they are all coordinators. However, syntactic complexity may be a major determinant in accounting for the relative order of acquisition of the Dutch coordinator *want* versus the subordinator *omdat* (see Chapter 13 for more details on the acquisition of these connectives). In Chapter 12, syntactic complexity will be taken up as an explanation for syntactic developments within the use of each connective.

Another determinant of the acquisition process may be social-pragmatic complexity. For example, Kyratzis et al. (1990) argue along the following lines: the more useful a certain connective is in the life of a child, the easier it is to acquire that connective. This factor will be discussed in more detail in Chapter 13.

In the discussion of factors that might explain the developmental sequence, one important factor has been disregarded so far: the influence of environmental variation or parental input. This factor will be taken up for further investigation in Chapter 11, which also gives a more detailed study of the overall developments of the Dutch connectives treated in this chapter.

Chapter 11

Quantitative developments in children's connective use

The current chapter complements Chapter 10 in two ways. Firstly, Chapter 10 focused on the emergence of the connectives *en*, *maar¸ toen*, and *want*, without providing quantitative data; the current chapter examines the quantitative development of these connectives. To this end, for each connective growth curves per child are presented. Secondly, Chapter 10 introduced cumulative conceptual complexity as the major determinant of the developmental sequence. The current chapter investigates whether the cumulative complexity approach can be maintained in light of a) the quantitative overall development, and b) a first exploration of the role of parental input on the acquisition process.

> *"It has been shown that a complicated growth curve analysis is worthwhile, not only because of its statistical advantages, but also for the reduced number of arbitrary decisions, and therefore the possibilities for interpretation."*
>
> (Van den Bergh & Rijlaarsdam 1996: 231)

11.1 Introduction

Although research into the emergence of connectives is interesting and useful in itself, Chapter 10 does not give a complete picture of the process of connective acquisition (see the methodological discussion in Chapter 9). The current chapter extends the findings in the previous chapter in two ways. First of all, it provides a quantitative analysis of the overall developments of the four connectives *en*, *maar*, *toen*, and *want*. Secondly, it studies the tenability of the cumulative complexity approach. More specifically, I investigate whether environmental variation or parental input should be seen as a serious alternative determinant of the acquisition process. Given limitations of time, the study of parental input is restricted to one connective (*toen*). Hence, this analysis can only be regarded as a primary exploration of this alternative theory. The research question of this chapter is presented in (1).

(1) <u>Research question of this chapter:</u>
Can the cumulative complexity approach be maintained in light of the quantitative overall development of the four connectives?

Section 11.2 gives some theoretical background on the role of parental input in the acquisition process; section 11.3 and 11.4 present the methodology and the results of the quantitative study. The conclusions as well as some points for discussion are given in section 11.5.

11.2 The influence of parental input

In the previous chapter I adhered to the view that cumulative conceptual complexity is a major factor in determining the developmental connective sequences in child language. The current chapter investigates the tenability of the cumulative complexity approach. More specifically, it pursues the idea that parental input may contribute to the way children develop their connective use as well (cf. also section 10.2.1 on Ingram's (1989: 78) environmental

variation). After the analyses of parental frequencies by – among others – Brown et al. (1969), Brown & Hanlon (1970), and Brown (1973), the role of parental input has by and large been neglected in the acquisition literature. For example, generative approaches focused on the LAD (Language Acquisition Device) as the explanatory factor behind early child language acquisition, diminishing the influence of parental input on the acquisition process.[1] However, since the recent introduction of the usage-based approach to child language acquisition, interest in parental input has been renewed (cf. Diessel 2004; Tomasello 2000, 2003).

The influence of parental input has also been investigated in the area of connective acquisition. For example, Diessel (2004: 171-172) is one of the advocates of research into the nature of parental input. His impressive work on the acquisition of English conjoined clauses contains many quantitive analyses of developments within different types of complex clauses. He introduces the ambient language as a determinant for these developments. His very general analysis indicates that parental input indeed seems to play a role: English connectives that frequently appeared in the mother's data emerged earlier, and connectives that were relatively infrequent in the mother's data emerged relatively late (see section 11.3 for some critical remarks on Diessel's methodology).

In a study into the use of markers of causality, McCabe & Peterson (1997) investigated the possibility that parental conversation with children is a source of development. They found that parents predate their children's so-called spontaneous explicit expression of causal links by approximately five months. According to McCabe & Peterson (1997: 151), parents "scaffold their children's emergent causal language by asking questions, repeating or revising such questions, occasionally answering those questions themselves, or abandoning them. Children respond to parental causal questions before they make statements of causal connection without such prompts (…)." This effect of parental scaffolding was prominent: "(…) the more parents stressed causal links in their conversations with their children, the younger the age at which children produced causal links without prompting" (p. 151).

There seems to be good reason, then, to investigate the influence of parental input on the Dutch course of connective developments. In section 11.4.3 I will present a first exploration of the role of parental input on the acquisition of Dutch *toen*.

11.3 Methodology

My first analyses of the overall development of *en*, *maar*, *toen*, and *want* are of a quantitative nature. In order to characterize the development of these connectives after their first emergence I have examined the longitudinal transcriptions of the same children that were studied in the previous chapter (see section 10.4.1). In addition, I analyzed the data of four somewhat older children: Carl, Maike, Sanne, and Tinke (see Table 11.1).[2]

Table 11.1. Data on the Schlichting Corpus (with ages in years;months.days)

Child	Age range	Number of utterances	Corpus
Carl	3;06.01 - 5;03.22	5195	Schlichting
Maike	3;06.26 - 5;04.00	4854	Schlichting
Sanne	3;05.21 - 5;04.01	4640	Schlichting
Tinke	3;05.21 - 5;03.02	5269	Schlichting

[1] This LAD is seen as an innate language capacity that specifies universal linguistic principles.

[2] These data could not be used for the analysis in the previous chapter, since the recordings only started when the children were about three and a half, i.e., some time after the first emergence of the different connectives.

The data of these four children were gathered by Schlichting (1996) with three-month intervals. The transcriptions in these corpora are based on audiotape recordings made at home, in three different settings (telling a story on the basis of four pictures, free conversation, and conversation during an activity like drawing or making a jigsaw puzzle).

All connective utterances (together with several lines from the preceding and following conversation) were collected with the help of the CHILDES program *kwal*. The four connectives were then analyzed in a quantitative way. For each child I determined the connective frequencies per recording. No restrictions were set concerning the completeness or interpretability of the connective clauses. However, false starts and other repaired connectives were excluded from the analysis. According to these criteria, the connective *maar* occurs once in (2), and not at all in (3).[3] The latter clause does exhibit one occurrence of *en*.

(2) *<Maar maar> [/]* ***maar*** *die gaat weglopen.* (Abel, 3;2.11)
'<But but> [/] but that one runs away.'

(3) *<Maar> [//]* ***en*** *dan ga ik naar jouw trein ook.* (Abel, 3;1.7)
'<But> [//] and then I also go to your train.'

In delimiting the relevant class of *en*-utterances, I have chosen not to consider intraclausal relations such as (4) to (6). In the case of *maar*, only the complementizer use was taken into account, disregarding the particle use of which (7) gives an example.

(4) *Daar komen brandweerautootjes* ***en*** *ambulance.* (Peter, 2;5.29)
'There's (the) small fire engines and ambulance coming.'

(5) *Dat moet morgen wel met jou* ***en*** *met mama.* (Niek, 3;9.20)
'Tomorrow that has to (be done) with you and with mommy.'

(6) *Dit gaat zelf rijen* ***en*** *kiepen.* (Thomas, 2;10.26)
'This is gonna ride and tip by itself.'

(7) *Ja, ga* ***maar*** *weg.* (Tomas, 2;6.14)
'Yes, just go away.'

On the basis of the frequency counts a multi-level model per connective was specified and mean growth curves per connective were plotted (see Appendix J for some explanatory remarks on multi-level analyses, and Appendix K for statistical details of the model). The objective of the growth curves was to estimate the occurrence of connectives as a function of the elapsed time (cf. Van den Bergh & Rijlaarsdam 1996).[4] In addition, for each child developmental curves per connective were approximated and plotted. The growth curves in this thesis differ from, for example, the "ordinary" developmental curves presented in Brown (1973: 256), in that they are statistically supported by multi-level models that show significant differences between connectives and between children. A second advantage is that they allow me to derive a mean development, a general pattern from the data per child.

For two children the growth curves of the connective *toen* were related to frequencies in the parental input. This connective was selected because parents frequently prompt their children into producing *toen* by asking *en toen* 'and then' questions. The results of this study are presented in section 11.4.3. For both children proportions of parental connective use in the different recordings were plotted and compared to the estimated probabilities of connective

[3] The signs [//] and [/] respectively indicate false starts with and without a correction.

[4] The actual cumulative frequencies per connective were used to calculate the estimates for the ages for which no recordings were available.

occurrence in the child data. Compared to Diessel's (2004: 172) analysis of the correlation between parental proportions and children's age of first emergence, this analysis is more detailed in two respects. Firstly, Diessel groups the proportions of five parents and calculates a mean proportion of use. This mean proportion is then related to the mean age of first emergence. My analysis, on the other hand, allows for a comparison of the correlations between individual children and their respective parents. I present separate developmental pictures of the two children: for each child I relate the developmental pattern of the child to her mother's proportions of use. This may turn out to be important, since there may be differences between children or parents in their use of connectives. Secondly, I investigate the parental proportions by taking into account the connective proportions in each recording separately. Diessel's proportions of parental use, on the other hand, are the result of adding up the proportions over all recordings. His approach disregards the fact that parental frequencies may change as their children become linguistically more proficient over the three years in which data were gained.

11.4 Results

The current section discusses the results of the quantitative analyses of the four connectives under investigation. Section 11.4.1 discusses children's individual developmental patterns per connective. For each of these connectives a multi-level model was specified and so-called growth curves were plotted (see Appendix K for statistical details), analyzing a) the moment at which children first start to use the connective, and b) the degree of inclination of the growth curves. Section 11.4.2 introduces a picture of the mean development per connective, which allows for a comparison of the four developmental patterns. Section 11.4.3 presents a first exploration of the influence of parental frequencies on the acquisition of Dutch *toen*.

11.4.1 Growth curves per connective

This section presents the individual developmental patterns of the connectives *en*, *maar*, *toen*, and *want* successively. Figure 11.1 was constructed to give an impression of the differences between children in the use of the connective *en*.[5] It only depicts the development of the real connective use of *en*, that is, only those utterances in which the *en*-clause contains both a subject and a verb.[6] Each line in Figure 11.1 represents the development of one child during the period in which the child was recorded. The starting point and ending point of each line mark the beginning and end of the period during which the child was recorded. Since the children were recorded at different age ranges, the lines show variation in the age at which the lines start and stop, as well as in their length.

From the individual curves in this figure it can be inferred that the probability of occurrence increases with time. Most children start their development of the connective use of *en* around their second birthday (± 750 days). All children show a steady increase in their use of *en*. The rises at the right of the developmental curves in Figure 11.1 appear to be artificial (cf. Van den Bergh & Hoeksma 1993: 86-87); they are due to a relatively low number of

[5] In order to construct growth curves that cover the total age range per child, it was necessary to present probabilities of occurrence instead of actual frequencies of occurrence. These probabilities were estimated on the basis of the actual frequencies in the recordings.

[6] *En*-clauses that showed ellipsis were disregarded here, even though correct application of ellipsis results in grammatical utterances. In their early language, children often leave out one or more constituents from their utterances. In clauses without a subject and/or a verb, it is therefore impossible to determine whether they produce an incomplete clause that is ungrammatical from an adult perspective, or a complete, grammatical clause in which ellipsis has been applied.

observations at later ages. Furthermore, the estimated probability of occurrence varies to some extent per child, especially in the curves of the older children.

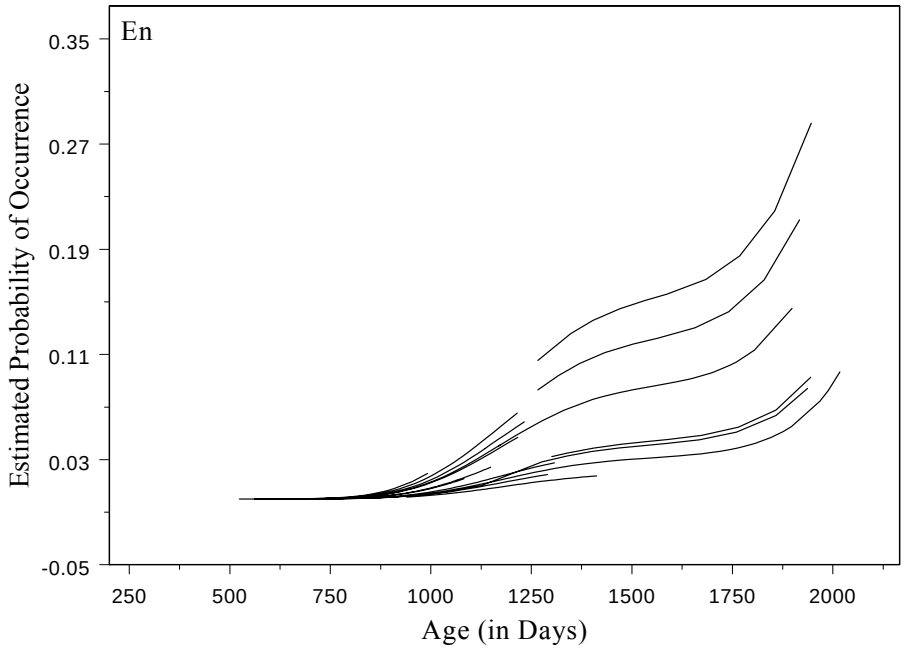


Figure 11.1. *En*: Estimated probabilities of occurrence per child at different ages

Figure 11.2 depicts the individual developmental curves of the children in their use of *maar*. In this figure only the complementizer use was taken into account, disregarding the particle use of *maar* (which occurs far more frequently in child language). The growth curves of *maar* show a less steep increase than the growth curves of *en*. The use of *maar* seems to stabilize at the age of 5. The developmental curves show artificial dips near the end of the age scale. The early and very sharp increase in one child's *maar*-use is an illustration of the individual differences between the children.

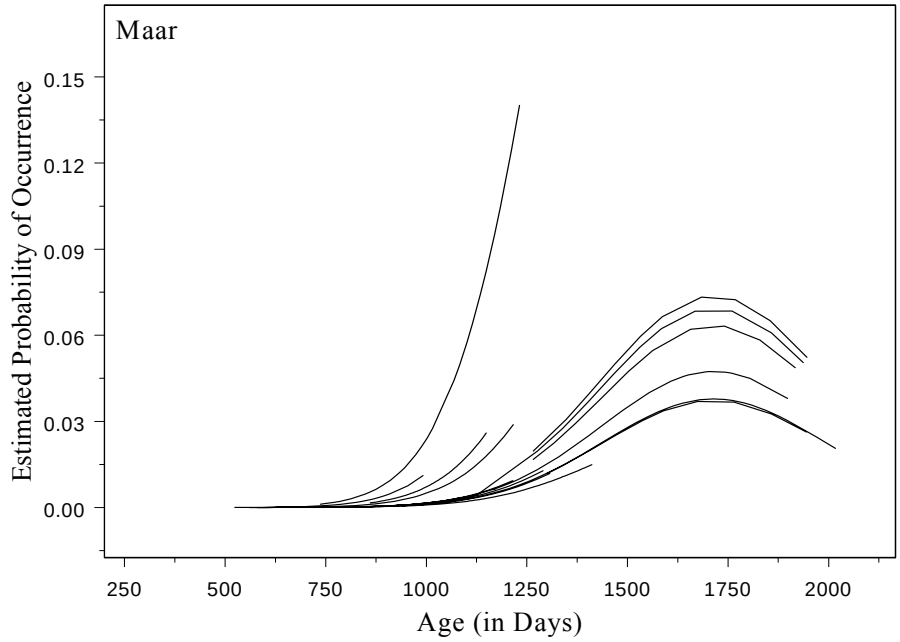


Figure 11.2. *Maar*: Estimated probabilities of occurrence per child at different ages

Figure 11.3 shows the probabilities of occurrence of *toen*. This figure includes both the adverbial and the complementizer use of *toen*. Compared to the developments of *en* and *maar*, the children show more variation at the moment in which they start using *toen*. However, most children do not start using *toen* before about 850 days. Again, the developmental curves show a steady increase, although there is some variation in this area; with age, children become quantitatively more proficient in using *toen*.

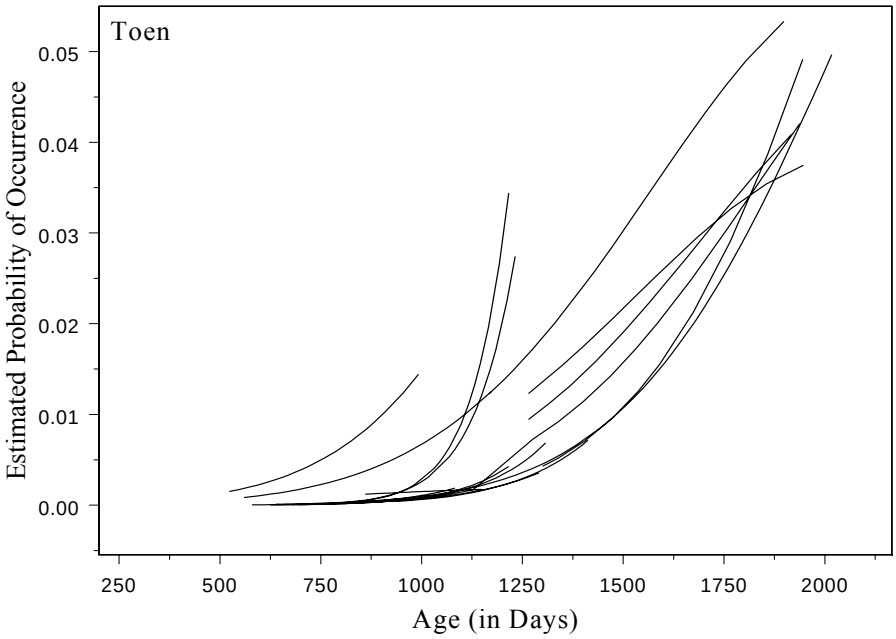


Figure 11.3. *Toen*: Estimated probabilities of occurrence per child at different ages

The developmental curves on *want* in Figure 11.4 reveal that many children start using *want* at a similar age (± 1150 days). This figure furthermore indicates that the older children show remarkable similarities in their acquisition of *want*: all curves show the same steady increase in the use of *want* and about the same probabilities of occurrence.

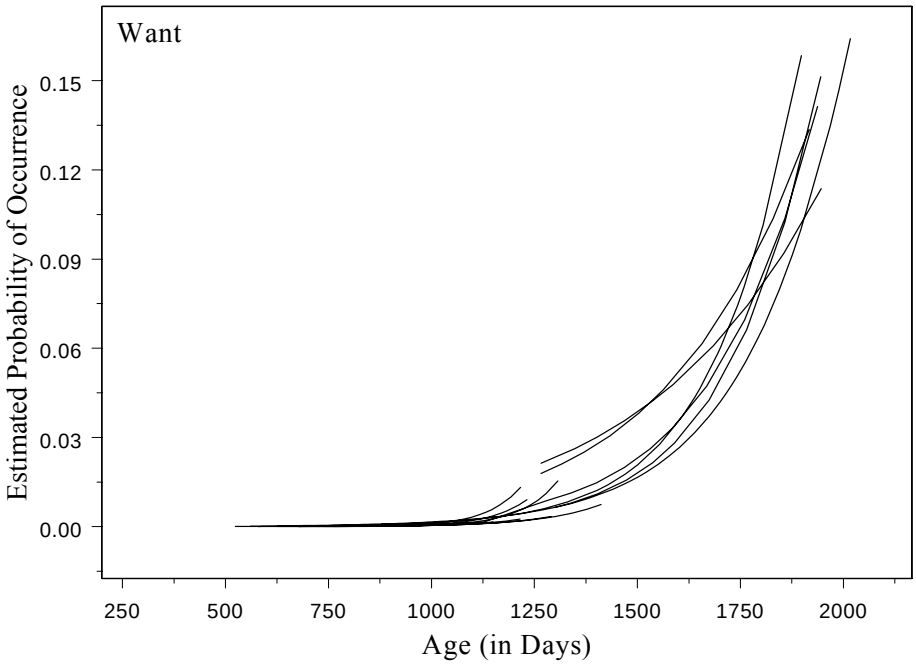


Figure 11.4. *Want*: Estimated probabilities of occurrence per child at different ages

11.4.2 Mean developments

The previous section dealt with the individual developmental patterns per connective. The current section focuses on the mean developmental patterns that could be derived from the data per child. To get a first picture of the quantitative variation per connective, I analyzed the total amount of the connective use per child (see Table 11.2). Roughly speaking, the connectives are ordered according to their frequency of use: *en* is the most frequent connective, whereas *want* is often the least frequent.

Table 11.2. Connective frequencies per child (total of all recordings)

Child	# utterances	en	maar	toen	want
Abel	11883	176	79	52	27
Carl	1071	79	56	20	17
Daan	15229	219	26	9	2
Hein	12781	202	78	16	2
Iris	8771	72	38	2	1
Josse	12651	222	170	52	21
Laura	22323	331	231	94	15
Maike	896	84	22	29	16
Matthijs	19864	141	29	23	41
Niek	15151	99	4	9	0
Peter	8578	39	13	20	3
Sanne	977	84	58	13	17
Sarah	17458	561	183	213	124
Thomas	12670	91	14	9	4
Tinke	962	63	44	22	16
Tomas	9126	9	7	1	1

In addition, for each of the connectives a multi-level model was specified and mean growth curves were plotted. These mean developments, which are portrayed in Figure 11.5, indicate the general patterns that can be derived from the individual data presented in section 11.4.1. From Figure 11.5 it can be concluded for each of the connectives that the probability of occurrence increases as a function of time: the older children get, the more frequently they use the four connectives. *En* is the earliest connective that shows an increase. This is also the connective with the sharpest rise. The development of *maar* starts somewhat later and increases to a lesser extent. The developments of *toen* and *maar* are quite similar to one another: the growth of both connectives starts at about the same time and the curves show a similar increase as well.

The developmental curves in Figure 11.5 seem to reflect the picture that could be obtained from the total frequencies in Table 11.2. The connectives that are the most frequent overall are the ones that are acquired first and show the quickest increase in use. The developmental picture in Figure 11.5 is also in line with the cumulative complexity approach presented in the previous chapter. The connective that is the least complex and that emerged first in the data of eleven children (*en*) is the connective that shows the earliest as well as the largest increase. The observation that children differed in the order of first emergence of *maar*, *toen* and *want*, is in line with the observation that the ages at which these connectives start to show an increase are very close to one another, and that especially the developmental curves of *toen* and *want* are very similar. The fact that *maar* – of these three more complex

connectives – shows the sharpest and largest increase can be related to the observation that more than half of the children produced this connective as their second connective.

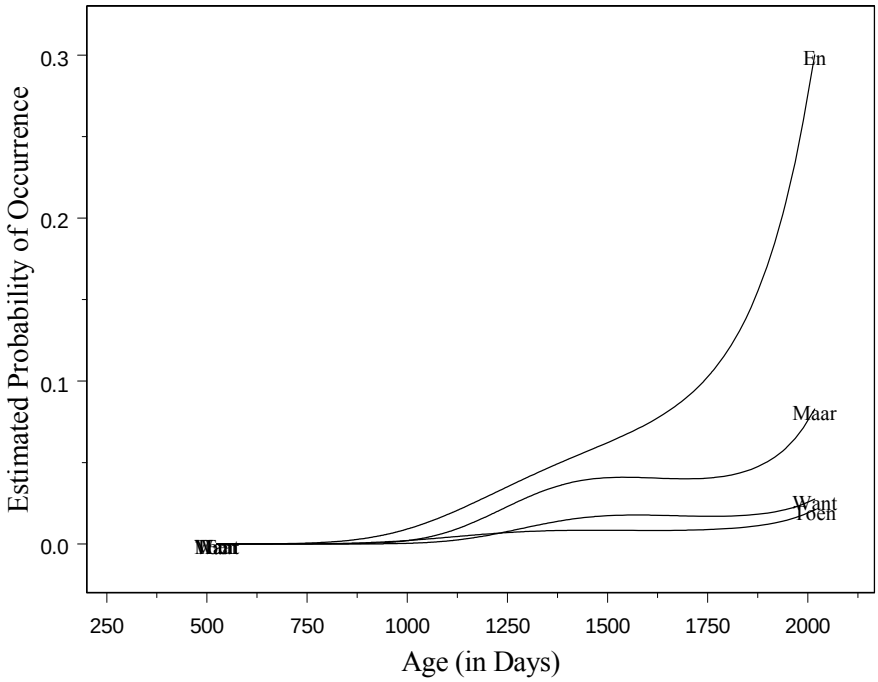


Figure 11.5. Mean development per connective

11.4.3 Parental influence on the acquisition of *toen*

The previous section suggested that the growth curves are in line with the hypotheses of the cumulative complexity approach. The current section presents a first exploration of the role of a second possible determinant in the acquisition of connectives: parental input. In the previous chapter, this type of environmental variation was set aside relatively easily. However, in usage-based approaches it plays a crucial role (cf. Tomasello 2003). In the context of this dissertation it is impossible to include a complete analysis of parental input. In order to shed some light on the chances of parental input to explain for the acquisition process, the data of one connective were analyzed for two children.

I performed a multi-level analysis (see Appendix J for some explanatory remarks) of the frequencies of the connective *toen* in the maternal data of two children: the sisters Laura and Sarah. These parental data are particularly interesting, since they come from the same person. From this analysis Figure 11.6 was constructed. There are two gradually increasing curves in this figure; the lower one depicts the development of Laura, whereas the upper one depicts Sarah's development in her use of *toen*. The other two, irregular, lines depict the maternal frequencies of use as they develop over time.

Several observations can be made from Figure 11.6. First of all, the maternal input shows a large amount of variation in the frequencies of *toen* per recording. An analysis of the recordings that result in the two large peaks in Sarah's input shows that the use of *toen* can be said to be context-dependent. In the recordings that result in the two peaks, Sarah and her mother are reading a story from a book. Many of the events in this story are related in a temporal way, and they are often marked with the connective *toen*. This suggests a genre-specific connection between conversation type (e.g. reading a book versus playing with toys) and type of coherence relation (temporal relations versus other relations) (cf. also Spooren & Sanders 2005).

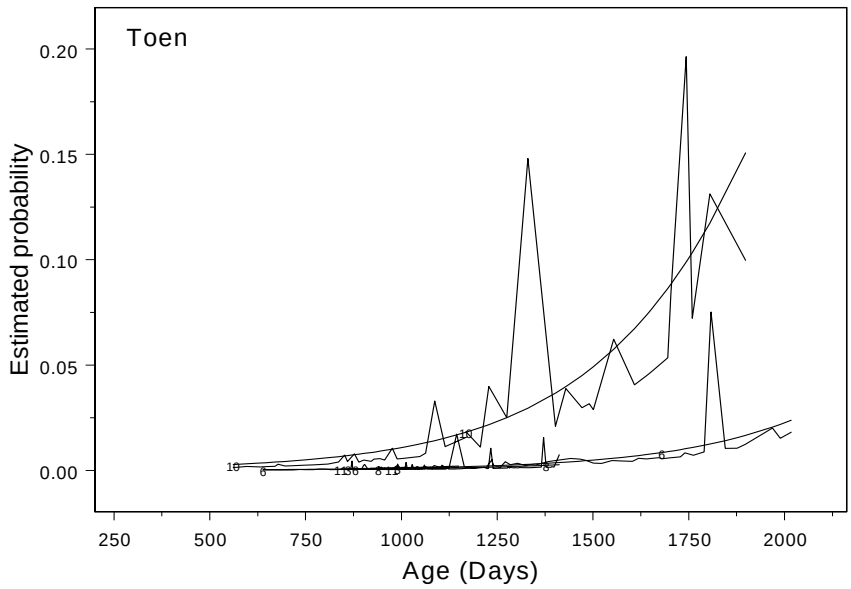


Figure 11.6. *Toen*: children's developmental curves related to parental frequencies

A second observation from Figure 11.6 is that maternal input differs per child and that there is a close relation between maternal input and child output. Sarah, who shows an early and steady increase in her use of *toen*, receives more instances of *toen* in her input language than Laura, who develops at a lower rate. At first sight, this could be taken as evidence in support of parental input as a determinant in the acquisition process: more frequent use of *toen* in the input language results in a faster development of *toen* in the language of the child. However, in this case it *is* remarkable that the parental input varies per child, since the data come from one and the same person: the mother of the sisters Laura and Sarah. What reason could a parent have to produce higher frequencies of *toen* at an early age for one child, and only at a later age for the other? If parental input were actually considered an independent determinant of the acquisition process, it would be expected that parents exhibit similar frequencies of use, irrespective of the child they are talking to.

The differences in maternal input cannot be resolved into different contexts in which the data were gathered: from the description of the corpora, it can be inferred that the data from Laura and Sarah were obtained in highly similar conversational contexts. Howevrer, it might be the case that parents monitor their own language use: they often check whether their children understand what they are saying, and consequently adapt their own language use. This so-called 'audience design' may result in different input frequencies per child. With this in mind, a cautious conclusion from the data in Figure 11.6 can be drawn, namely that parental input cannot be regarded as a completely independent factor: parents are careful in the way in which they talk to their children, determining what their children may or may not understand.

The data in this section are not very promising for the parental input approach, but it is too early to reject this alternative approach completely. Given that the analysis here is only based on the data of two children and their mother, and that it only focuses on parental input in a quantitative way, further research is needed to settle the discussion. For the time being, it can at least be concluded that there is no counterevidence to the cumulative complexity approach. Therefore, cumulative complexity can be maintained as the major determinant of connective acquisition. It sets the pace of children's connective acquisition; parents, in turn, adapt themselves to the abilities of their children. Children have to do with so-called 'audience design'.

11.5 Conclusion and discussion

The quantitative studies presented in sections 11.4.1 and 11.4.2 can be summarized as follows. First of all, the growth curves per connective reveal that for some connectives, children show highly similar developmental patterns (see Figure 11.4 on *want*), whereas in their use of other connectives they show more variation (e.g. see Figure 11.3 on *toen*). The four connectives differ in the times in which they arise in children's language as well as in the rate at which their use increases.

Secondly, the connectives that are the most frequent overall are the ones that are acquired first and show the fastest increase in use. *En*, the connective with the sharpest rise, is the earliest connective that shows an increase. The development of *maar* starts somewhat later and increases to a lesser extent. The growth of *toen* and *maar* starts at about the same time and their curves show a similar increase. These quantitative developments reveal that the picture obtained from the first emergence analysis in Chapter 10 does not rest on accidental early instances of connective usage. Rather, the developmental data presented in the current chapter complement the picture of first emergence.

How do these results relate to the research question of this chapter? Can the cumulative complexity approach be maintained in light of the quantitative overall development of connectives? Firstly, the picture of the mean developments (see Figure 11.5) is in line with the cumulative complexity approach presented in the previous chapter. The connective that is the least complex occurs most frequently, and the more complex connectives appear less frequently.

Secondly, a first exploration of a possible alternative theory suggests that parental input cannot be seen as a completely independent determinant of the acquisition process. Rather, it seems as if cumulative conceptual complexity sets the pace of children's connective acquisition and that parents use principles of 'audience design' when addressing their children: parents adapt themselves to the abilities of their children. If that explanation is on the right track, it would suggest that cumulative complexity can be maintained as the major determinant of connective acquisition.

The study of the parental input discussed in section 11.4.3 can only be regarded as a first exploration of the role of environmental variation. First of all, this study only focused on the connective *toen* in the data of two children and their mother; hence, it needs to be extended to other connectives and data of more children. Secondly, it only provided a quantitative analysis of the parental input. Further research is needed to investigate the qualitative influence of this ambient language on the acquisition of Dutch connectives (compare the English results for parental scaffolding in McCabe & Peterson 1997). My data in section 11.4.3 already indicated that the context in which the language is presented to the child seems to influence the number of connectives that the child may hear.

The quantitative analyses in this chapter only gave a partial picture of the overall connective developments during the recording periods. The next chapter focuses on the qualitative side of these developments.

APPENDICES TO CHAPTER 11

Appendix J – Explanatory remarks on the multi-level analyses

For the longitudinal acquisition data in this chapter, a multi-level model was specified. Because the linguistic capacities of the children continually change during acquisition, the frequency of connectives may change over time. Therefore, a growth model was used. The objective of the growth curve is to estimate the occurrence of connectives as a function of the powers of the elapsed time (Van den Bergh & Rijlaarsdam 1996). From the residual scores per connective curves for each child can be approximated and plotted.

In a multi-level model, means as well as within- and between-child variance are estimated (see Goldstein 1995; Van den Bergh & Rijlaarsdam 1996). Hence, two patterns of connective development can be distinguished. The means indicate the inter-individual patterns, the general patterns that can be derived from the individual data. The within-child variance and the between-child variance denote the intra-individual patterns, children's specific developmental patterns as well as the deviations of each child from the mean pattern.

Multi-level modeling has become a popular technique in analyzing data that are hierarchically ordered. Note that this is the case in the analysis of longitudinal acquisition data: in my multi-level model of connective acquisition the factor *connective frequencies* is nested within the factor *children* (each child has its own connective frequencies). As the contexts in which children are recorded may differ, the occurrence of connectives may differ between children. Multi-level modeling provides a tool to take such non-independence of data into account (cf. also Van den Bergh & Hoeksma 1993: 81-82).

Multi-level models have several advantages over correlative analyses. First of all, they allow for an analysis based on proportions (in which the number of connectives per recording is related to the total number of utterances). There are two reasons why an analysis based on proportions instead of frequencies is to be preferred. First of all, children's acquisition is more adequately characterized by proportions, seeing as though 20 connective clauses out of a total of 200 is quite different from 20 out of 40. Secondly, the total number of utterances is a function of the capacity to talk or fluency. This implies that the frequency with which a connective is encountered depends on verbal fluency: talkative children are likely to have high frequencies in all recordings, and less talkative children have low frequencies in all recordings. Thus, the correlation of frequencies between recordings is confounded with the total number of utterances. Consequently, correlations based on frequencies tend to be overestimated. (Van den Bergh & Rijlaarsdam 1996: 212-213).

A second advantage of the multi-level model concerns the number of observations. In longitudinal acquisition research, the number of children is usually quite low. If the data are aggregated to one observation per child (per recording) the power in the statistical analysis will be quite low (i.e. the a priori chance to reject the null-hypothesis is quite low). Fortunately, there are many observations per child, as each utterance in a recording is a different observation. A multi-level model allows the researcher to make use of this information, resulting in a substantial increase in the power of the analysis (cf. Van den Bergh & Rijlaarsdam 1996: 214).

A third advantage concerns the differences in number of recordings. In longitudinal acquisition research these numbers vary per child, since they are the result of arbitrary decisions of the different researchers who collected the data. A multi-level model poses no restrictions on the number of observations per child; it can handle missing values easily (Van den Bergh & Hoeksma 1993: 88). Another problem concerning the arbitrary number of

recordings is also circumvented here. In a growth model the data are not divided into arbitrary numbers of recordings, but they are placed on an age scale, using time as a continuous variable.

A final advantage is that multi-level models have the impact of individualization. Not only general trends in processes are estimated, but also the individual deviations from the general trend (Van den Bergh & Rijlaarsdam 1996: 232).

To conclude, a growth curve analysis is worthwhile, not only because of its statistical advantages, but also for the reduced number of arbitrary decisions, and therefore the possibilities for interpretation.

The actual multi-level model can be described as follows. Let Y_{ij} be the occurrence of a connective at day t_i of child j. Hence, the relation between the age of the child and the occurrence has to be estimated. Many models have been proposed to estimate such relations (see Van den Bergh & Rijlaarsdam 1996). In general, polynomials are preferred because of their flexibility and the ease of interpretation. That is, Y_{ij} is modeled as a function of powers of t_i of child j. The number of parameters is seen as an empirical question, with two rules of thumb for both the fixed and the random parameters:

1) Parameters are taken into account only if they have a significant contribution;
2) Higher powers are taken into account only if all lower powers reach significance.

A more formalized form of the model can be written as:

$$Logit(Y_{ij}) = \sum_{k=0}^{k=K} \mathrm{B}_k * t_{ij}^k + \sum_{k=0}^{k=K} \mu_{kj} * t_{ij}^k$$

In the equation above, the β's are the fixed parameters, i.e. the general, or mean change with age; the residuals (μ's) denote the deviance for child j of this general trend. The residuals are assumed to be normally distributed with a variance of $S^2_{\mu kj}$. Please note that the model can easily be extended with other parameters, like the gender of the child, or relevant utterances of the mother.

Appendix K – Multi-level analyses of the acquisition of the four connectives

I – Multi-level analysis of *en* (with standard deviation reported between brackets)

Fixed parameters			
$B_0 * t_{ij}^0$	$B_1 * (t_{ij} - t)^1$	$B_2 * (t_{ij} - t)^2$	$B_3 * (t_{ij} - t)^3$
-3.965 (.1765)	0.5710 (.0885)	-0.0952 (.0085)	0.0076 (.0007)
Random parameters			
	$S^2_{\mu 0j} * t_{ij}^0$	$S^2_{\mu 1j} * t_{ij}^1$	
$S^2_{\mu 0j} * t_{ij}^0$	0.4669 (.1709)		
$S^2_{\mu 1j} * t_{ij}^1$	-0.0016 (.0606)	0.1150 (0.0432)	

II – Multi-level analysis of *maar* (with standard deviation reported between brackets)

Fixed parameters			
$B_0 * t_{ij}^0$	$B_1 * (t_{ij} - t)^1$	$B_2 * (t_{ij} - t)^2$	$B_3 * (t_{ij} - t)^3$
-4.796 (.2918)	1.062 (.0439)	-0.1834 (.0117)	0.0107 (.0009)
Random parameters			
	$S^2_{\mu 0j} * t_{ij}^0$		
$S^2_{\mu 0j} * t_{ij}^0$	1.313 (.4731)		

III – Multi-level analysis of *toen* (with standard deviation reported between brackets)

Fixed parameters		
$B_0 * t_{ij}^0$	$B_1 * (t_{ij} - t)^1$	$B_2 * (t_{ij} - t)^2$
-5.718 (.2441)	0.5783 (.0836)	-0.0188 (.0041)
Random parameters		
	$S^2_{\mu 0j} * t_{ij}^0$	$S^2_{\mu 1j} * t_{ij}^1$
$S^2_{\mu 0j} * t_{ij}^0$	0.8625 (.3304)	
$S^2_{\mu 1j} * t_{ij}^1$	0.0161 (.0780)	0.0840 (.0364)

IV – Multi-level analysis of *want* (with standard deviation reported between brackets)

Fixed parameters		
$B_0 * t_{ij}^0$	$B_1 * (t_{ij} - t)^1$	
-6.112 (.2191)	0.7312 (.0872)	
Random parameters		
	$S^2_{\mu 0j} * t_{ij}^0$	$S^2_{\mu 1j} * t_{ij}^1$
$S^2_{\mu 0j} * t_{ij}^0$	0.6927 (.2727)	
$S^2_{\mu 1j} * t_{ij}^1$	-0.1188 (.0796)	0.0951 (.0412)

Chapter 12

The acquisition of connectives and clause integration

Chapter 11 focused on the overall development of four Dutch connectives in a quantitative way. The current chapter shows a qualitative examination of the overall acquisition of the connectives *en*, *maar*, *toen*, *want* and *omdat*. The connective *omdat* has been added to this selection, in order to gain more insight into the degree of clause integration during the acquisition of connectives. This chapter also adds to Chapters 10 and 11 by extending the cumulative conceptual complexity approach with a component of syntactic complexity.

> *"(…) the development of conjoined clauses can be seen as a process of clause integration. Starting from multiple-clause structures that consist of juxtaposed clauses, children gradually learn the use of complex sentences in which two or more clauses are integrated in tightly organized grammatical constructions."*
> (Diessel 2004: 171)

12.1 Introduction

The current chapter extends the findings in Chapter 10 and 11 in two ways. First of all, it adds to the emergence data in Chapter 10 and the quantitative data in Chapter 11 by presenting a qualitative analysis of the overall development of the connectives *en*, *maar*, *toen*, *want* and *omdat*. Secondly, it focuses especially on the role of cumulative syntactic complexity in connective development, whereas the previous chapters dealt with connective acquisition in terms of cumulative conceptual complexity. The five connectives are studied with the following question in mind (see (1)).

(1) <u>Research question of this chapter:</u>
Can the cumulative complexity approach be maintained in the light of the qualitative overall development of the connectives?

The aim of this chapter is to present a more detailed picture of the different conceptual and syntactic uses of each of the selected connectives during the acquisition process. In addition, I try to find out whether these acquisition data can be explained by reference to increasing syntactic complexity.

Section 12.2 discusses the cumulative syntactic complexity approach in more detail. Subsequently, section 12.3 presents the methodology and the results of the qualitative analyses, and section 12.4 provides an answer to the research question of this chapter. The results indicate that cumulative syntactic complexity is especially relevant in accounting for longitudinal variation within the use of each connective.

12.2 Cumulative syntactic complexity

In the current section, I will first discuss the general idea of the cumulative syntactic approach (section 12.2.1), and then introduce Diessel's (2004) theory about different degrees of clause

integration as a variant of this approach that can be used to account for the process of connective acquisition.

12.2.1 Syntactic complexity and processing load

In Chapter 10, the regularities in the order of emergence of connectives have been linked to cumulative semantic or conceptual complexity. However, Brown (1973) and Slobin (1973, 1977) have demonstrated that even when a meaning is potentially accessible to a child, he may be delayed in expressing it, at least in a conventional way, because of complexity in the formal linguistic mechanism used to encode it (cf. Bowerman 1979: 298). Therefore, investigators like Brown (1973) and Slobin (1973) argue that semantic or conceptual complexity must be distinguished from grammatical, formal or syntactic complexity – the complexity of the linguistic devices each language has for the expression of ideas (cf. Clark & Clark 1977: 337-339).

Given this general observation on child language development, a plausible hypothesis would be that the acquisition process of connectives is not only determined by the conceptual complexity of the connective, but that syntactic complexity also plays a role. In order to study the influence of formal, structural, or syntactic complexity, an independent criterion is needed to establish the degree of syntactic complexity. Some investigators have relied on grammars and have used the linguist's account of how certain sentences are related as a basis for predicting an order of acquisition (e.g. Brown & Hanlon 1970). But not all linguists agree on how to describe the relations between sentences of a language, so one grammar might make accurate predictions, while another might even predict the opposite order of acquisition (cf. Clark & Clark 1977: 339 and the references there). Therefore, I will define syntactic complexity in a cumulative way (compare the cumulative conceptual complexity approach in Chapter 10). A sentence Y is considered syntactically more complex than another sentence X, if the production of Y involves all the syntactic abilities that the production of X requires, plus at least one more.

As Diessel & Tomasello (2000, 2001) have argued, the degree of syntactic complexity children can deal with depends on the processing load they can handle. On the basis of data on the acquisition of relative clauses, they observe that children under three years of age tend to avoid relative constructions including two propositions. They argue that this can be related to the *processing complexity* of such complex constructions: "the processing load of these constructions would exceed their processing capacity at this early age" (Diessel & Tomasello 2000: 145). Similarly, they find that finite complement clauses emerge in constructions that are propositionally simple; although the composite structures are formally complex, these structures do not require the children to hold two propositions in short-term memory. According to them, "children's early use of complex sentence constructions might in general include just a single proposition because utterances that are propositionally more complex would exceed their processing capacity" (Diessel & Tomasello 2001: 136). Their view is consistent with recent suggestions by Newport (1990) and Elman (1993) who argue that processing limitations give rise to an incremental development by which the child proceeds from rather simple to more complex constructions.

12.2.2 Developments in the degree of clause integration

In this section I introduce Diessel's (2004) theory about different degrees of clause integration as a concrete variant of the cumulative syntactic complexity approach. I will first present his ideas, and then show how those can be related to syntactic complexity. Diessel's theory is formulated on the basis of English acquisition data, but it is likely that his theory can be used

to account for the development of Dutch connectives as well, because the underlying cognitive notions are not language-specific. Therefore, I will use his approach to account for syntactic variations within the development of the Dutch connectives under investigation.

Diessel (2004: 149) argues that conjoined clauses develop from simple nonembedded sentences. He proposes the following acquisition route for different types of English conjoined clauses:

(2) <u>The development of conjoined clauses (Diessel 2004: 171, Figure 7.5</u>

juxtaposed sentences
↓
sentences that are pragmatically linked by a "discourse marker"
↙ ↘
coordinate sentences | final adverbial clauses
↓
initial adverbial clauses

The earliest multiple-clause utterances consist of juxtaposed clauses, i.e. clauses in which the link between two semantically associated utterances is not expressed overtly by a connective (Diessel 2004: 158). In Braunwald's (1985) terms this first developmental step is "the conjoining of two thoughts in a single context of use" (p. 513). Some typical examples are given in (3) and (4).

(3)	There's the lion. Here kitty.	(Nina, 2;0)	(Diessel 2004: 158)
(4)	Hit ball. Get it.	(Adam, 2;3)	(Diessel 2004: 158)

The earliest conjoined clauses with an explicit connective are pragmatically combined with an utterance in the previous discourse. The majority of these clauses are linked to a clause that constitutes a separate intonation unit.[1] Moreover, they are often associated with an utterance across speaker turns. Some examples are given in (5) and (6).

(5)	Child:	Nina has dolly sleeping.		
	Adult:	The doll is sleeping too?		
	Child:	**And** the man's sleeping on the big bed.	(Nina, 2;2)	(Diessel 2004: 159)
(6)	Adult:	Flipper's on TV yeah.		
	Child:	**And** Shaggy's not on TV.	(Sarah, 3;8)	(Diessel 2004: 159)

Diessel (2004: 159) labels the intonationally unbound use of *and* the "discourse marker" use, in order to distinguish it from its use as a coordinating connective.[2] In coordinate constructions marked with *and*, the conjoined clauses are intonationally bound and linked to a clause within the same utterance.

[1] In the CHILDES format a period indicates the end of an utterance (MacWhinney 1995: 60).

[2] Diessel himself (2004: 159) uses the term "discourse connective" instead of "discourse marker". In order to prevent confusion with the term connective, I have chosen to use a different label, a label that covers the use Diessel is talking about.

While children's early adverbial clauses always follow the associated utterance, their later conjoined clauses can also precede the associated clause. Compare the examples in (7) and (8).

(7) He bite the tongue **while** he was eating. (Nina, 3;0) (Diessel 2004: 167)
(8) **If** he takes all of them I'm gonna beat him up. (Adam, 4;10) (Diessel 2004: 168)

To summarize, Diessel (2004: 171) regards the development of conjoined clauses as a process of clause integration. Starting from multiple-clause structures that consist of juxtaposed clauses, children gradually learn the use of complex structures in which two or more clauses are integrated in tightly organized grammatical constructions. Although Diessel is not explicit about the moment at which combinations of three clauses occur in child language, it is to be expected that such combinations only appear when children are able to produce connective clauses that are intonationally integrated in and linked to a clause within the same utterance.

The developmental pattern observed by Diessel can be regarded as a syntactic extension of the cumulative complexity approach presented in Chapter 10. Firstly, producing two syntactically and intonationally integrated clauses is more complex than producing each clause separately. Secondly, the production of integrated final adverbial clauses can be regarded as more complex than the production of integrated coordinate sentences. In order to produce final adverbial clauses the child needs to order the related clauses in a hierarchical way. Such a hierarchical ordering is not needed in the production of coordinate sentences. Finally, initial adverbial clauses are cognitively more complex – at least in terms of processing cost – than final adverbial clauses: they require children to plan two clauses at a time, instead of one clause after another.

It can be concluded that with age, children become more and more proficient in integrating conceptual ideas in syntactic constructions that are more complex in terms of processing cost or the amount of planning that is needed. In my qualitative study, then, I will investigate whether the development of Dutch connectives follows a route similar to the one proposed by Diessel.

12.3 Methodology and results

The current section discusses the methodology and results of the qualitative analyses of the five connectives. I have examined all connective utterances (together with several lines from the preceding and following conversation) in the longitudinal data of the sixteen children that were mentioned in the previous chapter (see section 11.3). Per connective I a) established in which order specific conceptual functions and syntactic properties were acquired, and b) compared the degrees of clause integration to the acquisition pattern proposed by Diessel (2004). Section 12.3.1 to 12.3.3 successively discuss the results for the connectives *en*, *maar* and *toen*; section 12.3.4 treats the results for *want* and *omdat*. To reduce the amount of work, the occurrence of three-clause combinations has only been investigated for the latter two connectives.

12.3.1 Results for *en*

From the qualitative analyses of *en*, it can be concluded that the clause combining function is not the first function in which *en* occurs. As Table 12.1 shows, eleven of the twelve children begin to use *en* in its function of coordinator of two (or more) constituents like noun phrases or prepositional phrases. In Thomas' recordings the first clause-internal *en* emerges two days

before the first clause combining *en*. However, it is very likely that Thomas has acquired the clause-internal before the age of 2:3.23, his age at the day of his first recording.[3]

Table 12.1. Ages of first correct *en*-usage

Child	**first clause-internal *en***	**first connective *en***
Sarah	1;9.10	1;11.15
Laura	1;10.25	2;2.10
Peter	2;0.7	2;3.7
Thomas	2;3.25	2;3.23
Daan	2;0.22	2;4.0
Abel	2;1.16	2;4.9
Hein	2;4.11	2;4.14
Matthijs	2;2.9	2;4.24
Tomas	2;3.20	2;5.7
Josse	2;1.12	2;8.4
Iris	2;5.12	3;1.0
Niek	3;1.17	3;4.9

The sequences in this table are in line with the data of De Villiers, Tager Flusberg & Hakuta (1976), who found for English *and* that phrasal coordination emerged long before sentential coordination. They therefore propose: "the roots of coordination are found in the conjunction of similar elements. The elements increase in complexity and length until the ability to coordinate complete propositions develops" (1976: 9).

Some early examples of the real connective use of *en* are given in (9) to (11).

(9) Hein: *Dit is van mij.*
'This is mine.'
Mother: *Ja.*
'Yes.'
Hein: ***En** dit is van mij.* (Hein, 2;4.16)
'And this is mine.'

(10) Mother: *Ja in mijn bord zit pap.*
'Yes in my plate there's porridge.'
Thomas: ***En** Loek eet de appelsap.* (Thomas, 2;3.23)
'And Loek eats the apple juice.'

(11) Sarah: *Eigenlijk moge(n) we d'r ook op.*
'Actually we are allowed to get on it too.'
Mother: *Ja.*
'Yes.'
Sarah: ***En** kleine Nijntje gaa(t) ook erop.* (Sarah, 2;8.6)
'And little Nijntje gets on it too.'

[3] The clause-internal *en* only occurs in the very first recording in the data of Hein, who also shows a relatively small number of days between the two moments of emergence. This implies that both the clause-internal and the clause combining use of *en* may have developed before the age of 2;4.11. In all other cases the files with the first emergence of *en* are preceded by at least one file (but often several files) in which the child does not use *en* at all.

As these examples illustrate, many of the two clauses combined with *en* occur in separate intonation units, often even across speaker turns. The first *en*-fragments of each of the children are all instances of such clause combinations across utterances. This is in line with Diessel's (2004) observation that coordination within one single intonation contour (see (12)) occurs later.

(12) *Jij bent kindje* ***en*** *ik ben moeder.* (Laura, 3;4.25)
'You are child and I am mother.'

These results indicate that the children need two abilities to coordinate sentences properly: a) the ability to form syntactic coordinate constructions, which develops from clause-internal coordination to clause-combining coordination, and b) the ability to integrate two ideas into one complex clause, which follows Diessel's route.

12.3.2 Results for *maar*

From a qualitative perspective, it is remarkable that – unlike *en* – *maar* is hardly used to mark clause-internal contrastive relations as in (13) to (15). In fact, many children do not even produce *maar* in clause-internal contrastive relations.

(13) *Niet broodje* ***maar*** *potje he?* (Thomas, 2;5.25)
'Not sandwich, but potty, isn't it?'
(14) *Jij mag de ring lenen van mij* ***maar*** *niet houwe (= houden).* (Maike, 5;1.3)
'You may borrow the ring from me, but not keep (it).'
(15) *Van die niet* ***maar*** *van deze wel.* (Tinke, 5;3.2)
'Not from that one, but from this one.'

This implies that the acquisition of the clause combining function of *maar* can hardly be facilitated by the clause-internal use of *maar*.

The qualitative analysis of clause combining *maar* reveals that often the complementizer use of *maar* cannot be related to a clear negative coherence relation in the sense of Sanders et al. (1992). Instead, examples like (16) and (17) seem to mark 'upcoming disagreement' (cf. Schiffrin 1987), which can be labeled as a discourse marker use of *maar*.

(16) Mother: *Nee, rustig [!] tekenen.*
'No, draw quietly.
Sarah: ***Maa(r)*** *dan zit ik zo laag.* (Sarah, 3;0.19)
'But then I sit so low.'
(17) Adult: *Dat is ook rood, net als jij hebt.*
'That is also red, just like you have.'
Daan: ***Ja maar*** *jij moet [/] jij moet ook rood.* (Daan, 2;10.28)
'Yes but you should (take) read as well.'

The real contrastive *maar* often combines two clauses with a separate intonation contour (as in (18)). Only later on are the two clauses combined within a single intonation contour (see (19)).

(18) *Ik wil geen motor.* ***Maar*** *nou wil ik een politieauto.* (Josse, 2;11.9)
'I don't want a motorbike. But now I want a police car.'

(19) *En deze doet ook net alsof ie even groot is als die* ***maar*** *dat is ie niet.* (Sanne, 5;4.1)
'And this one pretends he is as big as that one but he isn't.'

These results indicate Dutch children follow Diessel's route during their acquisition of *maar*. In contrast to the acquisition of *en*, the acquisition of clause-combining *maar* is hardly facilitated by the clause-internal function of this word. Rather, it is the frequent use of *maar* at the discourse level (as a marker of upcoming disagreement) that triggers the child into combining clauses in a more local contrastive way.

12.3.3 Results for *toen*

The first use of *toen* is restricted to the adverbial use, as is illustrated in (20) and (21).

(20) ***Toen*** *was ik ook mee, hoor.* (Laura, 3;4.21)
'Then I also came along.'
(21) *En* ***toen*** *ging ikke boos worden.* (Niek, 3;8.30)
'And then I got angry.'

As Table 12.2 shows for the younger children in this study, the emergence of adverbial *toen* always precedes the acquisition of complementizer *toen*.

Table 12.2. First emergence of adverbial and complementizer *toen*

Child	Age range	First adverb	First complementizer
Abel	1;10.30 – 3;04.01	2;4.23	2;11.10
Daan	1;08.21 – 3;03.30	3;1.14	3;3.30
Hein	2;04.11 – 3;01.24	2;5.19	after 3;1.24
Iris	2;01.01 – 3;06.15	3;3.23	after 3;6.15
Josse	2;00.07 – 3;04.17	2;2.8	2;11.23
Laura	1;09.04 – 5;06.12	3;4.21	3;11.16
Matthijs	1;10.13 – 3;07.02	3;0.20	3;5.13
Niek	2;07.00 – 3;10.17	3;8.30	3;10.3
Peter	1;05.09 – 2;08.22	2;4.12	after 2;8.22
Sarah	1;06.16 – 5;02.13	2;4.2	3;2.13
Thomas	2;03.22 – 2;11.22	2;7.1	after 2;11.22
Tomas	1;07.05 – 3;01.02	2;7.10	after 3;1.2

In addition, the adverbial use of *toen* remains the more frequent of the two throughout the development (see the total frequencies of use in Table 12.3).

Table 12.3. Specification of the categorical status of *toen*

Child	**# adverbs**			**# complementizers**	**Total**
	# clause-initial (combi *en toen*)	**# clause-initial (without *en*)**	**# clause-medial**		
Abel	14	28	2	8	52
Daan	5	1	2	1	9
Hein	8	8	0	0	16
Iris	1	1	0	0	2
Josse	14	15	7	16	52
Laura	36	49	4	5	94
Matthijs	3	14	3	3	23
Niek	3	4	0	2	9
Peter	0	20	0	0	20
Sarah	93	97	10	13	213
Thomas	5	1	3	0	9
Tomas	0	1	0	0	1
Carl	20	34	2	8	64
Maike	16	23	1	4	44
Sanne	27	32	8	10	77
Tinke	27	48	3	4	82
Total	272	376	45	74	767

Within the adverbial use, *toen* is preferred in a clause-initial topic position, either preceded by *en* or not (see Table 12.3). The clause-medial use of *toen* is relatively rare. Twenty of the 45 clause-medial fragments are questions (cf. (22)).

(22) *Ging de hond het* ***toen*** *opeten?* (Sanne, 3;10.8)
‘Did the dog eat it then?’

The complementizer use of *toen* is illustrated in (23) to (26). The first complementizer *toen*-clauses always occur in a separate clause, functioning as an afterthought or a postmodification (cf. (23) and (24)).[4] Only later on do children really integrate the *toen*-clause into the matrix clause (cf. (25) and (26)). This is in line with Diessel’s observations that intonationally and syntactically unintegrated clauses precede adverbial clauses that are intonationally bound to their matrix clause.

(23) Adult: *Oh, heb je die van Rosie gekregen?*
‘Oh, did you get that one from Rosie?’
Josse: *Ja*.
‘Yes.’
Adult: *Dat is ook een* [*hele mooie*].
‘That’s also a very pretty one.’
Josse: [***Toen*** *we in Sloten*] *waren*. (Josse, 3;1.10)
‘When we were in Sloten.’

[4] The square brackets around the adult’s words *hele mooie* ‘very pretty’ and the child’s words *toen we in Sloten* ‘when we in Sloten’ indicate that these parts of the utterances were produced simultaneously.

(24) Niek: *Ik ben al in het ziekenhuis (ge)legen.*
'I have already been in hospital.'
Father: *Ja.*
'Yes.'
Niek: ***Toen*** *ik nog baby was.* (Niek, 3;10.3)
'When I was still a baby.'
(25) *Maar deze was kapot* ***toen*** *ik hem vond.* (Sanne, 5;4.1)
'But this one was broken when I found him.'
(26) ***Toen*** *ik jarig was dan komt zwarte piet.* (Laura, 4;0.20)
'When I had my birthday 'zwarte piet' came.'

The acquisition of complementizer *toen*-clauses in postposition precedes the acquisition of *toen*-clauses in preposed position in the recordings of six of the seven children who acquire the complementizer *toen* during the period in which they are recorded.[5] This is in line with Diessel's observation that the acquisition of final adverbial clauses precedes the acquisition of initial clauses.

Furthermore, the postposed *toen*-clauses outnumber the preposed ones. Only 24 of the 74 fragments with complementizer *toen* have the *toen*-clause in preposed position. In 21 of the 24 combined clauses with a preposed *toen*-clause, *toen* also occurs as a resumptive element within the postposed matrix clause (cf. (27) and (28)).

(27) ***Toen*** *ík klein was* ***toen*** *kan 'k nog niet opruimen*
en toen hadden we daar naartoe gereden. (Maike, 4;0.13)
'When I was little then I couldn't clean up and then we drove there.'
(28) (About a dog that wants to eat pancakes)
Maar ***toen*** *ie het op wou eten* ***toen*** *viel het op zijn neus*
en toen kwam de mama d'raan. (Sanne, 3;10.8)
'But when he wanted to eat it then it fell on his nose and then the mommy came.'

From the qualitative data on the acquisition of *toen*, it can be concluded that children start with the adverbial use of *toen*, which is used mainly in topic position, and only later acquire the complementizer use of *toen*.[6] The developmental patterns of the complementizer use of *toen* are in line with Diessel's acquisition route.

12.3.4 Results for *want* and *omdat*

The qualitative analyses reveal that only four of the twelve younger children used *omdat* creatively during the period in which they were recorded, whereas all these children, except for Niek, used *want* (see also Chapter 10). The four older children – Carl, Maike, Sanne and

[5] Sarah's developmental sequence is the exceptional one: her first *toen*-clause in which *toen* functions as a complementizer is in preposed position (age 3;2.13). Her first postposed *toen*-clause occurs two months later, at the age of 3;4.13.

[6] It is likely that the complementizer use of *toen* is triggered by the adverbial use of *toen*. Adverbial *toen* in topic position (e.g. *en toen…* 'and then…') marks a shift along the time line, without tying the event mentioned in the *toen*-clause to a specific point in time. *Toen* only marks the temporal ordering of the two events. In its use as a complementizer, *toen* also marks the temporal relation between the events mentioned in the two combined clauses. In addition, *toen* ties the events to a specific point in time (the one mentioned in the *toen*-clause). In this sense, the complementizer use is more complex than the adverbial use of *toen*.

Tinke – all produced both connectives. The qualitative analysis of *want* shows that the earliest *want*-clauses appeared in separate utterances (cf. (29)). Only later on did integrated use of *want* develop (cf. (30)), but none of the children produced combinations of three clauses.

(29) Adult: *Past die er niet in?*
'Doesn't that one fit in it?
Abel: *Nee, die past er niet in.* ***Want*** *die te groot voor.* (Abel, 3;3.8)
'No, that one does not fit in it. Because that one (is) too big for (it).'
(30) *Ze kunnen ook niet praten* ***want*** *het zijn geen mensen.* (Carl, 5;1.3)
'They also cannot talk because they are not human.'

A quantitative analysis of the *want*-clauses in the final two recordings of the four older children shows that 16 *want*-clauses were integrated intonationally into their matrix clause, whereas 40 *want*-clauses were produced in independent utterances. Even around their fifth birthday, children still preferred to produce *want*-clauses separately. As with the other connectives, the increasing degree of clause integration in the use of *want* is in line with Diessel's findings.

The developmental data on *omdat* reflect Diessel's acquisition pattern as well. The earliest *omdat*-clauses appeared in separate utterances. In the data of the younger children, 16 of the 21 interpretable *omdat*-clauses that were produced separately were responses to *why*-questions from the parents (cf. (31)). Only five interpretable *omdat*-clauses were intonationally integrated into their matrix clause. These five utterances came from Laura's data: three with a postposed *omdat*-clause (cf. (32)), and two with a preposed *omdat*-clause (see (33)). None of the integrated connective clauses involve combinations of three clauses.

(31) Adult: *Waarom wil jij mij niet helpen?*
'Why do you not want to help me?'
Hein: ***Omdat*** *ik niet zin heb.* (Hein, 3;0.11)
'Because I don't feel like it.'
(32) *Die hebben allemaal dekens gepakt,* ***omdat*** *ze [//] ik hun bedje moet maken.*
(Laura, 5;2.21)
'They all took blankets, because I have to make their beds'
(33) ***Omdat*** *je niet zoveel gedrinkt heb, moet je nog een keer (…)* (Laura, 5;2.21)
'Because you did not drink so much, you have to go one more time (…)'

The number of integrated *omdat*-clauses hardly increases with age. A quantitative analysis of the *omdat*-clauses in the final two recordings of the four older children shows that only five of the 29 interpretable *omdat*-clauses were integrated intonationally into their matrix clause. It can be concluded that children aged five still prefer to produce *omdat*-clauses separate from the matrix clause to which they relate.

12.4 Conclusion and discussion

Summarizing the qualitative data in this chapter, it can be stated that the clause combining function of the selected words is not always the first function in which these words appear in child language. In the case of *en* and *toen*, children start with clause-internal functions, and only later on produce this word in their clause-combining function. The connective *maar*, on the other hand, seems to develop from a more global discourse marker, which marks upcoming disagreement, to a connective.

The Dutch connective developments are similar to one another in that the connective clauses are all first used in intonationally unbound clauses and only later occur in syntactically and intonationally integrated clause combinations. That is, in their use of the five connectives, children adhere to the developmental pattern observed by Diessel. These findings imply that the cognitive complexity approach can be maintained, but that it needs to be extended with a syntactic complexity component. For example, the development from clause-internal coordination to clause-combining coordination can be explained on the basis of the cumulative syntactic complexity approach: clause-combining coordination is syntactically more complex than clause-internal coordination because the child needs to deal with larger chunks of text (clauses instead of phrases). A similar story holds for the development from connective clauses with a separate intonation contour to connective clauses that are intonationally integrated into their matrix clause. With age, children become more and more proficient in integrating conceptual ideas in syntactic constructions that are more complex in terms of processing cost or the amount of planning that is needed.

Data on combinations of three clauses would have been interesting, given the closure theory presented in Chapter 5. However, the data on *want* and *omdat* revealed that children did not produce such three-clause combinations during the recording periods. The reason for this is not clear. A first reason might be that such combinations are too complex. Given their relatively limited memory span, children may not yet able to take three clauses into account at the same time. A second reason might be that combinations of three clauses hardly occur in oral language. Diessel (2004: 159-160) observes that the use of unintegrated connective clauses is not a specific trait of child language. He suggests that this unintegrated use of conjoined clauses is a common feature of spoken adult language. From this perspective, it would be surprising to find combinations of three clauses in spoken language. In order to test the tenability of this second option, further research is needed to find out whether adults produce three-clause combinations in their oral language.

The current chapter already discussed the qualitative development of *want* and *omdat*. However, up till now, a conceptual analysis based on domains has been lacking. In the next chapter, then, I will investigate the development of four causal connectives in more detail. Chapter 13 also returns to the main question of this thesis.

Chapter 13

The acquisition of connectives and domains of use

Chapter 10, 11 and 12 discussed the early emergence and further development of several Dutch connectives. Chapter 10 focused mainly on the development in terms of the cognitive primitives *basic operation* and *polarity*. The current chapter provides a detailed examination of the acquisition of four causal connectives in terms of a third primitive: *source of coherence*, or more specifically, *domains of use*. This chapter also returns to the main question of this thesis. An analysis of both the conceptual and syntactic properties of children's development in their use of *want*, *omdat*, *dus*, and *daarom*, should result in a higher insight into the relation between domains of use and word order, and the interaction between domains of use and positioning of connectives.

> *"Adults' competence with causal interclause relations begins with 2 to 3-year-old children's intense interest in the cause and consequences of their own and other people's behavior and the reason why events in their environment occur."*
>
> (Braunwald 1997: 136)

13.1 Introduction

The first aim of this chapter is to present a more detailed picture of children's connective development based on domains. The results on these domain developments are required for the second half of this chapter, in which I return to the form-function relations put forward in Chapters 5 to 7. Two of the hypotheses about form-function relations involve the domains of use. In order to answer the main question of this thesis, I try to find out whether the data on the acquisition of causal connectives can be used to put the hypotheses about the interaction between word order and domains, and the interaction between word order and position to the test. The following questions are taken up for investigation (see (1)).

(1) <u>Research questions of this chapter:</u>
   a. What does children's connective development look like in terms of domains?
   b. What do the acquisition data reveal about the form-function relations put forward in Chapters 5 to 7?

Section 13.2 introduces two conflicting theories about the order in which the domains emerge in child language. Subsequently, section 13.3 presents the results of an experiment on the use of epistemic and speech-act causals by Dutch children aged four and five. Section 13.4 shows the results of a longitudinal study into the development of four Dutch causal connectives (*want*, *omdat*, *dus*, and *daarom*). Subsequently, in section 13.5 I return to the main question of this thesis, and discuss what these developmental data reveal about the interaction between domains of use and syntactic characteristics of the connective clause. The conclusions and some points for discussion can be found in section 13.6.

13.2 Domain acquisition according to the literature

Is there an order in which children acquire the different uses of causal connectives? At least two studies in existing literature on this topic claim there is. Surprisingly, however, these studies provide different answers to the question. Kyratzis, Guo & Ervin-Tripp (1990) claim that speech act is acquired first (see section 13.2.1), whereas Spooren & Sanders (2005) support the theory that content emerges first (see section 13.2.2). Section 13.2.3 will present several factors that may account for the clash between the acquisition orders in the two studies. These factors indicate (see section 13.2.4) that additional research into domain acquisition is needed to settle the debate.

13.2.1 Speech act first…

Kyratzis, Guo & Ervin-Tripp (1990) analyzed data from 21 English children aged 2;4 to 12;0 in the 1976 Ervin-Tripp family corpus. This corpus is composed of natural family interactions recorded in the home involving activities such as having lunch and playing tea party. Kyratzis et al. focused on the acquisition of the causal connectives *because* and *so*. On the basis of the quantitative data in Table 13.1, Kyratzis et al. concluded that very young children only used *because* and *so* to mark speech-act relations, and that later on the content use of these connectives increased. Epistemic use was very infrequent, and remained so throughout the age period studied here.

Table 13.1. Domain frequencies by age according to Kyratzis et al. (1990: 209)[1]

Age	Content	Epistemic	Speech act	Incomplete	Total
2;4-3;6	0	0	4	14	18
3;7-6;6	30	9	110	48	197
6;7-12;0	16	1	40	10	67

The claim that speech act occurred first is based on a small number of observations. The majority of causals produced by the youngest age group and one third of the causals of the intermediate age group were incomplete. These incomplete utterances included constructions that lacked either an action or a reason clause, or that were a response to a *why* question (Kyratzis et al. 1990: 210).

Kyratzis et al. (1990: 210) conclude: "It appears that children first learn causals as realizing one particular pragmatic function, that of justifying speech acts, and only later, with development, acquire a broader range of pragmatic functions for their causal constructions – their function of explaining events described in propositions and their function of justifying conclusions." They explain their order of acquisition by reference to the fact that children want to achieve what they want and that speech-act relations are the most convenient way to do so. This explanation can be regarded as a social-pragmatic complexity approach: the more useful a certain connective use is in the life of a child, the easier it is to acquire that connective use.

13.2.2 … or content first?

The English findings of Kyratzis et al. (1990) are in contrast with the Dutch results reported in Spooren & Sanders (2005). Spooren & Sanders performed a corpus analytic study under experimental control using Dutch children as informants. Their research was carried out with

[1] The numbers in Table 13.1 are inferred from the percentages mentioned in Kyratzis, Guo & Ervin-Tripp (1990: 209, see their Table 1).

elementary school children from different age groups: six children from grade 1 (age 6-7) and six from grade 6 (age 11-12). The children were given two tasks, both aiming at eliciting clause combinations that expressed a causal coherence relation. In the description task, which biased for semantic or content relations, children had to describe pictures that showed large numbers of causally related events. In the conversation task, which biased for the two types of pragmatic relations (i.e. epistemic and speech-act relations), the children had to formulate their opinion on a number of controversial topics, such as vetoed TV programs and the amount of children's pocket money (Spooren & Sanders 2005: 17).

The results of this study were as follows (see Spooren & Sanders 2005: 20). Summed over both tasks, the younger children produced more semantic content relations than the older children. Within the pragmatic relations, the speech-act use predominated. However, the two age groups did not differ in their numbers of epistemic and speech-act relations. The data showed a significant influence of context on the occurrence of different domains: in the conversation task children produced more epistemic and speech-act relations. In the description task the content domain prevailed.

Spooren and Sanders take their results as an indication that semantic content relations are acquired before both types of pragmatic relations.[2] They claim that it is the cognitive or conceptual complexity that determines this order in the acquisition of different coherence relations and the connectives expressing them. According to them (see Sanders et al. 1993), argumentation (i.e. pragmatic use) often involves reasoning on the basis of semantic (i.e. real-world) relations, and therefore pragmatic use is more complex than semantic use.

13.2.3 Search for an explanation

The results of Kyratzis et al. (1990) and Spooren & Sanders (2005) are in contrast to one another. Several factors may be responsible for this clash. First of all, the two studies differ in their objects of analysis: Spooren and Sanders focused on Dutch coherence relations with and without connectives, whereas Kyratzis and colleagues only took English connective clauses into account.

Secondly, Spooren and Sanders only studied children aged six or older, whereas Kyratzis et al. analyzed the causal language of children as young as 2;4. The youngest children in the Dutch study had already acquired all three domain types. As Spooren & Sanders (2005: 30) suggested themselves, "the utterances of younger should be studied to find the kind of differences under discussion."

A third factor that can account for the different acquisition orders is that Kyratzis et al. only took into account causal utterances in which the child produced both the antecedent and the consequent clause. This selection criterion may have deflated the number of causal utterances as a whole. This is especially relevant for the youngest age group, since my analyses in the previous chapter revealed that very young children only explicate causal links across utterance boundaries, and often also across speaker turns (cf. also Diessel 2004: 158).[3] More importantly, it is likely that this selection criterion had two special effects on the number of content relations in the child data. Firstly, the use of the criterion deflated the

---

[2] Spooren & Sanders (2005) use the technique of *implicational scaling* here: the rank orders of the domains (in terms of frequency) are equated to the acquisition sequence of these domains. See Klein-Gunnewiek (1999: 52) for some critical remarks on this technique.

[3] Spooren & Sanders (2005) apply the same criterion in their analyses. However, the impact of this criterion is less drastic in their study, since the production of causal utterances by children aged six and older is less dependent on the occurrence of *why*-questions.

overall number of content relations, because it led to the exclusion of children's answers to *why*-questions, questions that are known to frequently trigger content relations. Secondly, the criterion especially deflated the number of content relations in the data of the youngest children. Other studies into the acquisition of English causal connectives (cf. Braunwald 1997; Diessel 2004) have shown that children's earliest instances of *because* are often restricted to occurrence in responses to *why*-questions, and that only later on do children develop the ability to produce *because*-clauses in more spontaneous ways. Kyratzis' relatively high number of incomplete causal expressions, which included the responses to *why*-questions, is in line with this developmental pattern. It appears that the selection criterion set by Kyratzis and colleagues resulted in an underestimation of young children's abilities in terms of domains.

The most important factor that can account for the clash between the results is the context in which the acquisition data were gathered. According to Spooren and Sanders, the type of relations children produce depends on the communicative task.[4] The results of their experiment support this claim: in the conversation task children produced more pragmatic relations, and in the description task the content domain prevailed. Translating this claim to Kyratzis' findings, it can be stated that "the number of speech acts is inflated due to the context in which the data were collected" (Spooren & Sanders 2005: 11). The bias for speech-act relations in conversations during plays and during dinner can be inferred from the following remark in Kyratzis et al. (1990: 210): "a preliminary analysis of the adults' uses of causals in this corpus revealed that a vast majority were also Speech Act-Level Causals."

13.2.4 Conclusions from the literature

The two studies presented in section 13.2.1 and 13.2.2 resulted in different orders of domain acquisition. Kyratzis et al. (1990) found that speech acts appeared first in child language, but they collected their data in a conversational context that was biased towards speech acts. In addition, they excluded certain causal utterances from their analysis, which resulted in a lower number of content relations in the data of the younger children. Spooren & Sanders (2005), on the other hand, claimed that content relations emerge first, but their data were collected from children aged six or older, and therefore provide only indirect evidence for this acquisition order.

As section 13.2.3 mentioned, there are several factors that can account (at least partially) for the clash between the two acquisition studies. These factors indicate that additional research into domain acquisition is needed to settle the debate. In the next two sections, I will present the results of two studies that were carried out to shed more light on the acquisition of domains. Section 13.3 discusses an experiment that was conducted to find out whether Dutch children as young as four are able to produce epistemic relations. Section 13.4 introduces a domains analysis of the longitudinal data of twelve Dutch children in the CHILDES-database. This analysis was carried out to find out the order in which the three domains emerge in Dutch child language and to give a more detailed picture of the domains acquisition of the different causal connectives. Such a picture is needed in order to investigate in more detail the

[4] The influence of task type is also reported in Shapiro & Hudson (1997: 42): "The level of narrative skill that children demonstrate can be increased or decreased by providing them with more or less difficult tasks. Moreover, the structure of the story that they are telling (i.e. the degree of causal complexity) will affect the types of interclausal devices they employ, independent of their general ability to use these markers."

interaction between domains and word order, and the interaction between domains and positioning (see section 13.5).

13.3 An experiment on Dutch: epistemic late?

The data reported in Kyratzis et al. (1990) and Spooren & Sanders (2005) both indicated that epistemic relations are acquired relatively late. The younger children in the English study hardly produced any epistemic causal relations, and the Dutch study revealed that six-year olds were able to come up with epistemic relations, but that neither six-year olds nor twelve-year olds produced them frequently. However, it can be argued for both acquisition studies that the numbers of epistemic relations were probably deflated due to independent reasons (see section 13.3.1). Hence, the Dutch and English data do not present a proper picture of the number of epistemic relations in early child language. In order to find out whether young children are indeed hardly able to produce epistemic relations, I conducted an experiment among Dutch children aged four and five.[5] The materials, participants and analytical method of this study are accounted for in section 13.3.2; the results are presented in section 13.3.3.

13.3.1 Deflated numbers of epistemic relations

In the English study, the number of epistemic relations was deflated due to the conversational contexts in which these data were gathered. These contexts showed a bias for speech-act use and not for epistemic use.

In the Dutch study, the number of epistemic relations was deflated due to another reason: Spooren and Sanders restricted their epistemic category to relations involving the generation or inference of new knowledge (see their Appendix 3). Spooren & Sanders classified the causal relation in (2) as epistemic, because the speaker concludes "here and now" that a possible explanation for the mummy walking around is that someone opened the coffin. In contrast, the relation in (3) is not categorized as an epistemic relation, but as a speech-act relation, since it conveys existing instead of new knowledge (both examples are taken from Appendix 3 in Spooren & Sanders 2005: 41-42).

(2) [Chantal points to a mummy walking around; in the corner of the picture there is an open coffin.]
Daar komt een mummie. Misschien hebben ze de kist opengedaan. (Chantal, 12;6)
'There comes a mummy. Maybe they opened the coffin.'

(3) *En Sesamstraat vind ik niet meer zo leuk. Want daar ben ik nu te groot voor.* (Rose, 12;3)
'And I don't like Sesame Street very much anymore. Because I'm too old for that now.'

The "here and now" restriction in the definition of the epistemic category resulted in a relatively low number of epistemic relations and a relatively high number of speech acts. This relatively high number of speech acts must have been due to this restriction indeed, because the two conversational tasks hardly caused the child to produce motivated interrogatives or imperatives. Kyratzis and colleagues, on the other hand, only included justified interrogatives, direct or indirect imperatives, and responses in their category of speech acts. In their study, this category contained hardly any matrix clauses with the form of an assertion or a declarative (cf. Kyratzis et al. 1990: 208).

---

[5] The experiment was designed in close collaboration with Ted Sanders, Arvid van Maaren, and Marijke van Middendorp. It was carried out by Arvid van Maaren and Marijke van Middendorp (cf. Van Middendorp & Van Maaren 2001).

Because the Dutch and English data did not present a proper picture of the number of epistemic relations in early child language, I conducted an experiment in which children were prompted to produce both epistemic relations and non-declarative speech act relations.

13.3.2 Materials, participants and analytical method

Each child was given two ten-minute tasks: an argumentation task with a bias for epistemic relations, and a directive task that was biased for relations in the speech-act domain. In the argumentation task children had to select their favorite teacher, friend or sports man out of a set of four pictures and then convince a hand puppet why this was the right choice. In the directive task children were told that they were the teacher of a hand puppet and that they had to instruct the hand puppet where to put certain stickers on a picture. Because this hand puppet was very naughty, they had to express the motivation for why the puppet should put it there and not somewhere else on the picture.

The investigation was carried out with five children from *groep 1* 'pre-kindergarten' and five children from *groep 2* 'kindergarten' of an elementary school in a medium-sized town in the southwest of the Netherlands. The ages of the children in the first group varied from 4;7 to 4;11. The ages of the children in the second group varied from 5;5 to 6;5. All children had Dutch as their first language and were interviewed individually in a separate room of the elementary school they attended.

The children's utterances were recorded on a cassette recorder and transcribed. The unit of transcription and analysis was the clause. In addition to complete clauses (containing a verb), I took into account answers to *yes/no*-questions as well as contracted clauses that would be considered grammatical in adult Dutch. Of each clause it was decided whether it could be linked to another one, and if so, whether it should be linked in an additive or a causal way (cf. the basic operation of Sanders et al. 1992). Using a paraphrase test, I categorized all the causal utterances in terms of domains. For ease of comparison with the Dutch study among children aged six and eleven, the definitions of the epistemic and the speech-act category were the same as in Spooren & Sanders (2005).[6] Per task and per group, the frequencies of use in the epistemic and the speech-act domain were calculated. These frequencies were analyzed statistically using a logit analysis (see Appendix L).

13.3.3 Experimental results and conclusions

The children produced 156 spontaneous causal coherence relations. These included four causally related utterances in which the first clause was not expressed by the child, but by the investigator. The frequencies of the domains produced by the two age groups are presented in Table 13.2.

Table 13.2. Frequencies of use of the two groups in the directive and the argumentative task

	Directive task		**Argumentative task**		**Total**
	Group 1	**Group 2**	**Group 1**	**Group 2**	
Speech act	31	19	12	6	68
Epistemic	10	2	31	45	88
Total	41	21	43	51	156

[6] The choice of these definitions instead of Kyratzis' ones did not affect the results in a significant way. The children mainly produced speech-act relations that included imperatives. There were hardly any declaratives in the speech-act category, because the argumentative task forced the children to draw new conclusions instead of reporting old ones.

This table shows that children aged four and five are able to relate clauses in an epistemic way. In the argumentative task, the children produced more epistemic than speech-act relations (χ^2 (1) = 63.5; p < .001). The argumentative task, then, often resulted in utterances like (4), whereas the directive task mainly triggered utterances like (5).

(4) *Deze is een leuke meester, omdat 'ie leuke schoenen aan heeft.* (Danny, 5;11)
'This is a nice teacher, because he is wearing nice shoes.'

(5) *Pak de wolk 'es. Leg 'm 'es in de lucht, omdat de wolken altijd in de lucht horen.* (Hugo, 5;11)
'Grab the cloud. Put it in the sky, because clouds always belong in the sky.'

Furthermore, group 2 produced more epistemic relations in the argumentative task than group 1 (χ^2 (1) = 5.6; p < .03), which indicates that the ability to produce epistemic relations increases with age.

The experimental data confirm the influence of context on the occurrence of domain types. In addition, they reveal that children's development in terms of domains takes place before the age of four. Children as young as four are capable of producing epistemic relations when they have the communicative goal of persuading someone, which is the case in an argumentative context. This implies (a) that the debate on the acquisition order of the domains should not be restricted to the content and the speech-act domain, and (b) that children below the age of four should be studied in order to decide which domain of use emerges first. In section 13.4, then, I will present a longitudinal study into the domain acquisition of causal connectives by Dutch children aged two and older.

13.4 Longitudinal study of Dutch

In order to settle the debate on the acquisition order of the three domains, I examined longitudinal data from twelve very young Dutch children in the corpus of the CHILDES Project (MacWhinney 2000). Most children were recorded from the age of two till the age of 3;6 (see section 10.4.1 for more details on the Dutch longitudinal data), comparable to the ages of the youngest group in Kyratzis' study.

Section 13.4.1 presents an overall picture of the domain developments. The three following sections present characteristics of the domain developments per causal connective: section 13.4.2 focuses on *want*, section 13.4.3 on *omdat*, and section 13.4.4 on *dus* and *daarom*. Section 13.4.5 provides a summary of the results on domain acquisition.

13.4.1 General picture of the domain developments

All utterances with a causal connective (i.e. with *want*, *omdat*, *dus* or *daarom*) were selected, and on the basis of a paraphrase test, classified as content, epistemic or speech act. This time, the definitions were highly similar to the ones in the English study (and also in line with the domain definitions used in Chapters 4-8). The analysis included connective clauses that were causally linked to utterances from the children themselves or to utterances from their parents, including *why*-questions. Only direct imitations and other non-creative utterances were excluded from the calculations. Using the paraphrase test, I established for each child the distribution of the connective fragments over the domains.

As a first start, I compared the Dutch data to the English data mentioned in Kyratzis, Guo & Ervin-Tripp (1990). Since Kyratzis et al. did not provide individual developmental patterns, but presented their domain data per age group, I also grouped the data of the twelve Dutch children. To be more precise, I determined the total frequencies of use in the different

domains till the age of 3;6, the upper age boundary of the youngest age group in the English study. Table 13.3 presents a comparison of my distribution results for Dutch and the English frequencies (see also the data in Table 13.1, taken from Kyratzis et al. 1990: 209).

Table 13.3. Distribution of Dutch and English connective utterances over the domains

Age range	Content	Epistemic	Speech act	Not interpretable	Total
Dutch 1;6-3;6	29	23	41	63	156
English 2;4-3;6	0	0	4	18	22

First of all, this table shows that the Dutch children produced a large number of uninterpretable utterances. This category of uninterpretable fragments included incomplete utterances, utterances that contained inaudible parts of speech, and fragments in which the two related clauses could not be linked in a sensible way.

Secondly, Table 13.3 shows that Dutch children as young as three were able to produce connectives in all three domains. This contrasts with the English study, in which no content or epistemic fragments were found. It can be concluded, therefore, that the grouped English data did not present a clear picture of the order in which the three domains emerged in child language.

Because the grouped data in Table 13.3 did not result in a clear acquisition order of the domains, I determined the acquisition order per child. An overview of these individual orders of emergence can be found in Table 13.4.

Table 13.4. Orders of emergence of the three domains in connective utterances

# children showing	Order of emergence			
this order	1	2	3	not acquired
3	content			speech act, epistemic
2	content	speech act	epistemic	
1	content	epistemic	speech act	
1	content, epistemic	speech act		
2	speech act			content, epistemic
2	speech act	epistemic	content	
1	speech act	content	epistemic	

A quick count shows that seven children produced content first, and five children produced speech act first. Only one child, Abel, came up with two domains at the same time: The first file in which he used *want* contained both a content-fragment and an epistemic fragment. Although some children had a very infrequent use of connectives, these individual data imply that the relative order of content and speech act cannot be established on the basis of these findings. The only clear conclusion is that epistemic is never acquired first.

13.4.2 Domains in the acquisition of *want*

What does the domains development of *want* look like? The analyses revealed that five children were able to produce *want* in all three domains. Some child examples are given in (6)-(8). The content example in (6) gives a description of a real-world causal relation: It shows Patricia's motivation for her act of going to Spain. The epistemic example in (7) is not a description of the real world, but it gives an argument for the child's claim that the train

goes the wrong way. The reason clause in (8) supports the illocutionary force of the first segment, in other words, it gives an argument for performing the imperative speech act.

(6) *Patricia gaat soms naar Spanje,* ***want*** *ze papa woont daar.* (Laura, 5;5.14)
'Patricia sometimes goes to Spain, because her (lit. his) father lives there.'

(7) Child: *Ja, hij* (= de trein) *gaat dan fout, hè.* (Matthijs, 3;6.3)
'Yes, it (= the train, lit. he) goes wrong then, doesn't it.'
Mother: *Ja?*
'Yes?'
Child: ***want*** *eigenlijk hoort ie zo te rijden.*
'because actually it (lit. he) should drive this way.'

(8) *Niet zo hard,* ***want*** *anders kan ik niet werken.* (Laura, 4;9.10)
'Not so loud, because otherwise I cannot work.'

Table 13.5 shows the children's distributions of *want* over the three domains of use.

Table 13.5. *Want*: Distribution over the three domains of use

Child	Content	Epistemic	Speech act	Not interpretable	Total
Abel	6	7	8	6	27
Daan	0	1	1	0	2
Hein	0	1	1	0	2
Iris	0	0	1	0	1
Josse	2	6	4	9	21
Laura	5	2	4	4	15
Matthijs	4	3	13	21	41
Niek	0	0	0	0	0
Peter	1	0	0	2	3
Sarah	22	40	32	29	123
Thomas	0	1	1	2	4
Tomas	1	0	0	0	1

Table 13.5 shows that Daan, Hein, Iris, Niek, Peter, Thomas, and Tomas use *want* very infrequently. For this reason, the data of these seven children are not suitable to draw conclusions on the development of the three domains of use. For the other five children, who produce *want* more frequently, an overview like the following one for Laura was constructed (see Table 13.6).

Table 13.6 reveals that Laura has a clear order in her acquisition of the three domains: the content use of *want* emerges first, subsequently the speech-act use appears, and finally the epistemic use of *want* emerges. Laura is remarkable in this respect, since the other four children do not show a clear preference for one of the domains during their first recordings. The earliest *want*-fragments produced by Abel, Josse, Matthijs, and Sarah provide examples from all three domains. These children differ from each other in that Josse and Sarah have a slight preference for the epistemic domain, whereas Matthijs prefers to use *want* in speech-act fragments, and Abel has no domain preference at all (compare the frequencies in Table 13.5). All in all, for *want* it is not possible to infer a general order in which the three domains emerge in child language.

Table 13.6. Frequencies of *want* and *omdat* in Laura's data[7]

Age	Content	Epistemic	Speech act	Not interpretable
3;4.25	1 want	-	-	-
4;1.11	1 want	-	-	-
4;5.9	-	-	-	1 want
4;7.31	1 omdat	-	-	1 want
4;8.28	-	-	1 want	1 want
4;9.10	-	-	1 want	-
4;10.2	1 want	-	-	-
5;0.20	-	-	1 want	-
5;1.21	2 want 1 omdat	-	-	1 want
5;2.12	4 omdat	2 want 2 omdat	1 want 3 omdat	-
5;6.12	2 daaromdat	-	-	
Total	13	4	7	4

13.4.3 Domains in the acquisition of *omdat*

The domains analyses of *omdat* showed that only four children used *omdat* in a correct and creative way. Two of these children, Hein and Josse, only used *omdat* when they answered *why*-questions. These *why*-questions prompted mostly content relations (as in (9)) or speech-act relations (cf. (10)).

(9) Mother: *Waarom wil jij mij niet helpen?*
'Why don't you want to help me?'
Child: ***Omdat** ik niet zin heb.* (Hein, 3;0.11)
'Because I don't feel like it.'

(10) Child: *Je mag niet meer zingen.*
'You may not sing anymore.'
Adult: *Nee, waarom niet?*
'No, why not?'
Child: ***Omdat** ik niet leuk vind.* (Josse, 3;2.15)
'Because I don't like it.'

The other two children, Laura and Sarah, produced *omdat* spontaneously most of the time (see (11) for an example). Their *omdat*-use was remarkable in another respect, since they came up with the same combination of *daarom* and *omdat*: *daar-om-dat*. Laura produced this form two times, but Sarah consistently used *daar-om-dat* instead of *omdat*. This form also occurred once in the data of Abel (see (15)).

(11) *Dit is xxx natte kleren, **omdat** ik in 't zwembad gezeten heb.* (Laura, 5;2.12)
'This is xxx wet clothes, because I have been in the swimming pool.'

(12) (Laura is talking to herself during her play.)
*Waarom heb je mij spullen gepikt! **Daar-om-dat** jij zo veel had.* (Laura, 5;6.12)
'Why did you steal my things? Because you had so many of them.'

[7] In this table, the recording dates without *want*-fragments are disregarded.

(13) ***Daar-om-dat*** *ik vier geworden ben, daarom kan ik deze.* (Sarah, 4;0.11)
'Because I became four, that's why I can make this one.'

(14) Mother: *Eh, waarom wil je chocomelk?*
'Eh, why do you want chocolate milk?'
Child: ***Daar-om-dat*** *ik veel dorst heb.* (Sarah, 4;0.11)
'Because I am very thirsty.'

(15) (Abel just recovered from an infection in his finger.)
Adult: *Maar hoe kwam dat dan, die infectie?*
'But how did you get that infection?'
Child: ***Daar-om-dat*** *in de auto komt.* (Abel, 2:10.14)
'Because it came in the car.'

Table 13.7 shows the distributions of *omdat* over the three domains.

Table 13.7. *Omdat*: Age of first emergence and distribution over the three domains of use

Child	First emergence	Content	Epistemic	Speech act	Not interpretable
Abel	(2;10.14)	1 daaromdat	0	0	2
Hein	2;10.15	4	1	2	0
Josse	2;11.23	2	1	4	3
Laura	5;1.21	5 +2 daaromdat	2	3	10
Sarah	3;3.21	9 daaromdat	0	1	1

What do these developmental data reveal with respect to the acquisition of the three domains? For eight children (including Abel), no conclusions can be formulated on the use of *omdat*, because these children do not use this connective at all. The data of the remaining four children show variation. Sarah has a clear preference for (*daar)omdat* in the content domain, although her very first *omdat*-clause marks a speech-act relation. Hein, Josse, and Laura do not show such a clear preference for one of the domains. Their *omdat*-clauses occur in all three domain types. To conclude, for *omdat* it is not possible to infer a general order in which the three domains emerge in child language, which is partly due to the very low number of *omdat*-clauses in the data of these young children.

13.4.4 Domains in the acquisition of *dus* and *daarom*

The analysis of the longitudinal data reveals that both *dus* and *daarom* occur very infrequently in child language. Table 13.8 presents an overview of the ages of first emergence of the four causal connectives. This table shows that only two children use *dus* as a causal connective, whereas four children produce the causal connective *daarom* during their recording periods.

Table 13.8. Overview of ages of first emergence of the four causal connectives

Child	want	omdat	daarom	dus
Abel	2;10.0	-	-	-
Daan	3;1.14	-	-	-
Hein	2;8.28	2;10.15	2;10.24	-
Iris	3;2.11	-	-	-
Josse	3;0.20	2;11.23	2;11.23	-
Laura	3;4.25	5;1.21	-	4;0.20
Matthijs	2;11.19	-	-	
Niek	-	-	3;9.6	
Peter	2;8.22	-	-	
Sarah	2;9.7	3;3.21	4;0.11	4;0.11
Thomas	2;10.19	-	-	-
Tomas	2;10.10	-	-	-

The word *dus* only appears as a causal connective in the data of Laura and Sarah. The lexical item does occur in the data of some of the other children as well (e.g. in the data of Abel, Matthijs, and Thomas), but in these cases the *dus*-clause is either interpretable, or *dus* seems to be used as a kind of filler. In the latter case, *dus* either occurs as a single word proposition, or at the end of a clause, preceded by a short pause (as in (16)).

(16) *Daar moet de chauffeur in. Hier is de dierentuin,* ***dus.*** (Matthijs 3;6.3)
'A driver has to go in there. Here is a zoo, so.'

Laura and Sarah, the only two children who use *dus* as a causal connective, have a clear positioning preference: as a connective marking new conclusions, it occurs in clause-initial position (cf. (17) and (18)), but as a discourse marker or as a causal connective marking accessible conclusions, it occurs in clause-medial or clause-final position (see (19) and (20)). All connective fragments are instances of epistemic relations.

(17) *Het (= Laura's schoenen) is groot,* ***dus*** *ik ben al zoveel jaar.* (Laura, 4;0.20)
'It (= Laura's shoes) is big, so I am that many years.'

(18) Mother: *Ja, het zijn een hele hoop druiven.*
'Yes, there are a lot of grapes.'
Child: ***Dus*** *die krijg je niet allemaal op.* (Sarah 4;5.29)
'So you cannot eat all of them.'

(19) Child: *Ik heb er veel.*
'I've got many of them.'
Mother: *Tweeëntwintig.*
'Twenty-two.'
Child: *Ja, ik heb dr +... Dat is tw [//]* ***dus*** *veel, hè.* (Laura 4;8.14)
'Yes, I've got… That is tw [//] "dus" many, isn't is?'

(20) *Oh ja, ik was een meisje. Ik ben geen prins* ***dus.*** (Sarah 4;5.29)
'Oh yes, I was a girl. I am not a prince "dus".'

The lexical item *daarom* is very infrequent in child language: Hein and Josse use it only once, Niek produces it twice, and Sarah comes up with an interpretable *daarom* eight times. There

are more children who use the lexical item *daarom*, but they either use it in incomprehensible clauses or they use it as a bromide (as in (21)). This bromide function of *daarom* occurs in the data of Abel, Daan, Laura, Matthijs, Niek, Sarah and Tomas.

(21) Child: *Niet doen!*
'Don't do that!'
Adult: *Waarom niet?*
'Why not?'
Child: ***Daarom** niet.* (Tomas, 3;1.2)
'Just because.'

The majority of the *daarom*-clauses are instances of a causal relation in the content domain (cf. (22)). There is one example in which *daarom* marks a speech-act relation (see (23)).

(22) Child: *Want jij heb em (= een puzzel) nooit meer gedaan, hè.*
'Because you didn't do it (= a puzzle) for a long time, did you?'
Mother: *Nee.*
'No.'
Child: ***Daarom** weet je em niet zo goed.* (Sarah, 4;9.29)
'That's why you cannot do it (lit. him) so well.'
(23) *Ik ga eh alleen (pro)beren.* ***Daarom** mag Pepijn niet kijken.* (Hein, 2;10.24)
'I'm gonna try it on my own. That's why Pepijn is not allowed to watch.'

The children show hardly any variation in their positioning of *daarom*: it has a relatively fixed position at the head of the clause (either preceded by a complementizer like *en* or not). There is one exception (see (24)): Sarah produces one instance of *daarom* in clause-final position. In this position, *daarom* refers back to the reason mentioned in the *want*-clause. It stresses that the reason in the *want*-clause is the reason why her mother is not allowed to go on.

(24) Mother: *Ik mag niet verder gaan?*
'I may not go on?'
Child: *Nee, nee, met dit xxx (…) want ik ben nog niet klaar,* ***daarom.*** (Sarah, 4;1.11)
'No, no, with this xxx (…) because I'm not ready yet, that's why.'

Although *daarom* is relatively rare in child language, it can be concluded that children have a syntactic preference for *daarom* in clause-initial positions, and a conceptual preference for *daarom* as a marker of content relations.

13.4.5 Summary of the domain results

The grouped data showed that Dutch children as young as three were able to produce connectives in all three domains. This contrasts with the English study by Kyratzis et al. (1990), in which neither content nor epistemic fragments were found for the youngest age group.

The orders of emergence per child showed that the relative acquisition order of content and speech act could not be established on the basis of these recordings. Although some children showed very infrequent use of connectives, a clear conclusion is that epistemic was never acquired first.

The data on the development of the individual connectives did not result in further insight into the order of emergence of the domains. For *want* it was not possible to infer a general order in which the three domains emerged in child language, because different children showed different domain preferences, whereas others used *want* too infrequently. A similar story holds for *omdat*, which only four children used in a correct and creative way. Three of these children did not show a clear preference for any one of the domains. Only *dus* and *daarom* showed a nice division of labor: *dus* was only used in epistemic relations, whereas the majority of the *daarom*-clauses were instances of content relations.

13.5 Connective acquisition in relation to the interaction hypotheses

What do the acquisition data on the causal connectives reveal with respect to the interaction hypotheses put forward in Chapter 3, and tested against the diachronic data in Chapters 5 to 7?[8] In this section, I will discuss each hypothesis in relation to the acquisition data.

The first interaction hypothesis concerned the level of attachment of the connective clause. In Chapter 3 I put forward Verhagen's hypothesis about the interaction between word order and closure: connective clauses showing V2 could either attach to the preceding clause, or to a combination of two preceding clauses; connective clauses showing V-late, on the other hand, only had the former option. In Chapter 5 this hypothesis was subject to a first test on the basis of data on the diachronic development of V2 *want* and V-late *omdat*. This diachronic study revealed that Verhagen's closure hypothesis was on the right track, but that it should be restricted to postposed connective clauses, since preposed V-late clauses proved not to be restricted to late-closure interpretations.

The closure rule only becomes relevant when children are able to distinguish between the V2 and the V-late word order and when they are able to produce combinations of three clauses. From the analyses in Chapter 12, it can be concluded that the acquisition data on *want* and *omdat* do not provide suitable conclusive evidence in favor of or against the modified closure hypothesis. Eight children only used *want* during the recording periods, which implies that they did not have a choice in marking causal relations either with a V2 word order or with a V-late word order. The other eight children, who did use both *want* and *omdat*, did not produce combinations of three clauses. Therefore, there were no constructions in which they had to apply the closure rule.

The second hypothesis concerned the interaction between domains of use and word order. The diachronic data on *want* and *omdat* in Chapter 6 revealed that the hypothesis about the interaction between domains and word order was on the right track. It appeared that only V-late connectives in the bound subordination construction were restricted to the content domain, whereas V2 clauses in free coordination as well as V-late clauses in other constructions were not restricted in this respect. Translating this hypothesis to acquisition terms, this could imply that the content use of *want* decreases as soon as children acquire *omdat*. However, the acquisition data do not provide conclusive evidence in this area. First of all, eight children did not use *omdat* at all during their recording periods, which implies that they did not have a deliberate choice in the word order within the causal connective clause.

---

[8] The subjectification hypothesis is disregarded here, although an increase in subjectivity is conceivable in child language as well. For example, Braunwald (1997) observes a development that can be characterized as such: "The meaning of children's early language is embedded in the immediate context of the "here and now". With development, language becomes freed from the "here and now" or decontextualized. That is to say, language bears the burden of conveying shared meaning in a relevant exchange of thoughts and feelings" (p. 124).

Furthermore, it can be concluded that the acquisition of *omdat* did not seem to affect the way children used *want*, since the children used *omdat* too infrequently and *want* continued to be used in all three domains. Only Sarah showed a relatively clear division of labor between *want* and *(daar)omdat*: she only used *(daar)omdat* in the content domain, whereas *want* showed up in all three domains. However, her particular use of *omdat* did not result in a change in her use of *want*.

The third hypothesis concerned the interaction between domains of use and positioning. The diachronic analyses on *dus* and *daarom* in Chapter 7 revealed that the hypothesis about the interaction between domains and positioning should be rejected. The acquisition data in the current chapter seem to point in the same direction. The causal connective *dus* was only used in the epistemic domain, whereas the vast majority of the *daarom*-clauses occurred in the content domain. Given that both connectives only occur in one domain, it is not possible to draw conclusions on the domains-positioning interaction for each connective separately. However, given that both connectives (with their different domain preferences) only appeared in clause-initial position, there is no indication that positioning is related to usage in different domains.

The fourth hypothesis concerned the interaction between different functions of a word and positioning. The diachronic data in Chapter 7 revealed that connective *dus* had a different positioning preference than discourse marker *dus*. Remarkably enough, children as young as four show the same division of labor between connective use and discourse marker use of *dus*.

13.6 Conclusion and discussion

From the longitudinal study I can only conclude that the epistemic domain seems to be acquired last. For content versus speech act I did not find a clear picture. However, the most important conclusion from this acquisition study is that context plays a crucial role in the production of domain types. It was not possible to draw firm conclusions on the sequence of development here, because the selection of child data did not systematically control for different contexts of use.

From the crucial role of context it follows that a complete account of the acquisition route can only be based on data from different contexts in which children use language. Such a picture can be obtained by using the methods advocated by, for example, Brown (1973) and Tomasello (2003), namely using longitudinal studies of interactions between children and their parents. In addition, I have shown that specific hypotheses about the use of certain types of relations in a specific communicative context can successfully be tested in experiments with young children (see section 13.3). Supplementing the study of longitudinal corpora with experimental data therefore seems a fruitful methodology.

With respect to the interaction hypotheses, it can be concluded that the acquisition data were not always suitable in providing additional support or counterevidence. Given the absence of combinations of three clauses and given the major influence of the contexts in which the data were gathered, no conclusions could be drawn with respect to the modified closure hypothesis or to the interaction between word order and domains.

The acquisition data did provide additional counterevidence to the hypothesis about the interaction between positioning and domains of use. In addition, the data on *dus* revealed that children as young as four were able to support the conceptual distinction between the connective use and the discourse marker use of *dus* with specific positioning preference (clause-initial versus clause-medial). This is in line with the diachronic findings in Chapter 7 and with the hypothesis that positioning interacts with conceptual functions other than the domains of use.

APPENDIX TO CHAPTER 13

L – Logit analysis of the experimental acquisition data

I – Remarks

- In the logit analysis, the frequencies per domain were related to the total number of utterances per age group (448 for group 1, and 518 for group 2).
- This analysis included all causal coherence relations that were uttered spontaneously, including four causal utterances in which the first clause was not expressed by the child, but by the investigator.

II – Data: Distribution of the causal coherence relations over groups and tasks

	Directive task		Argumentative task		Total
	Group 1	Group 2	Group 1	Group 2	
Speech act	31	19	12	6	68
Epistemic	10	2	31	45	88
Total	41	21	43	51	156

III – Results logit analysis

Logit (fr)	χ^2 model	df	p model	χ^2 factor	df	p factor
1. constant	89.43	7	< .001	-	-	-
+ 2. age group	85.79	6	< .001	3.63	1	< .1
+ 3. task	78.90	5	< .001	6.89	1	< .01
+ 4. domain	76.21	4	< .001	2.69	1	< .1
+ 5. task x domain	12.71	3	< .01	63.50	1	< .001
+ 6. group x task	6.15	2	< .05	6.56	1	< .025
+ 7. group x domain	5.59	1	< .025	0.56	1	< .9
+ 8. group x task x domain	0	0	1	5.59	1	< .025

IV – Parameter estimates for model 8

Parameter	Estimate	s.e.	z-score	p
constant	-2.60	0.19	-13.96	< .001
group: group 2	-0.67	0.30	-2.24	0.03
task: argumentative	-0.99	0.35	-2.87	0.004
domain: epistemic	-1.18	0.37	-3.19	0.001
group x task: group 2 argumentative	-0.18	0.59	-0.32	0.75
group x domain: group 2 epistemic	-1.10	0.83	-1.33	0.18
task x domain: argumentative epistemic	2.17	0.51	4.29	< .001
group x task x domain: group 2 argumentative epistemic	2.21	1.00	2.20	0.03

Part IV – Conclusion

Chapter 14

Conclusion and discussion

The current chapter provides an answer to the main question of this thesis, on the basis of the diachronic results and the acquisition results presented in Part II and Part III. In addition, it concludes this thesis with several general points for discussion.

> *"The meanings of syntactic structures are, in fact, less apparent than those of words, which are expected to have some meaning."* (Radden 1992: 515)

14.1 Introduction

In Chapter 1, I introduced the main research question of this thesis (cf. (1)).

(1) Main research question of this thesis:
What is the relationship between the text-linguistic and the sentence-linguistic properties of connectives?

In Chapter 3, specific hypotheses about form-function relations pertaining connectives were formulated, which were then tested on the basis of diachronic data (Chapters 5-8) and acquisition data (Chapters 10-13). Section 14.2 presents a summary of the main results that came out of these connective studies. Results pertaining to specific form-function relations will be discussed in section 14.3, which provides an answer to the main research question of this thesis Section 14.4 ends this thesis with several points for discussion.

14.2 Summary of the main results

In this thesis, I studied the development of several Dutch connectives from the perspective of language change as well as from the perspective of language acquisition. The major findings of both types of connective study will be discussed in section 14.2.1 and 14.2.2 respectively.

14.2.1 Highlights from the diachronic studies

The diachronic analyses of the four causal connectives showed that many of the present-day properties of these connectives were already there during the 13th century. These relatively stable properties and the major changes in the use of the four causal connectives are summarized in Table 14.1. The four connectives also showed relatively constant profiles of use over a period of 800 years in terms of SOC type and in terms of domains.

Table 14.1. Stable properties and major changes in the history of the four causal connectives

	want	**omdat**	**dus**	**daarom**
Conceptual function	causal connective	causal connective	causal connective	causal connective
Preferred domain	epistemic domain	content domain	epistemic domain	content domain
Word order	V2	V-late	-	-
Categorical status	complementizer	complementizer	adverb	adverb
Conceptual changes	loss of temporal connective use	loss of finalistic ("in order that") use	- loss of anaphoric functions - gain of discourse marker function	loss of anaphoric functions
Syntactic changes	loss of V-late word orders	reanalysis of preposition *om* + relativum *dat* into complementizer *omdat*	gain of the complementizer function	-

Perhaps the remarkable constant profile of the four words should not surprise us too much, since the specific profile of a certain connective guarantees its right to exist. Drastic change might result in a complete overlap with the profile of another causal connective, which could subsequently lead to the disappearance of one of the two connectives.

Still, Table 14.1 shows that certain conceptual and syntactic shifts were found. From these changes it can be concluded that specialization of connectives indeed plays a role. For example, the disappearance of V-late *want* resulted in a nice division of labor between V2 *want* and V-late *omdat*: V2 *want* is used mainly to express epistemic relations, whereas V-late *omdat* occurs typically in the content domain. A second case of specialization was shown in the divergence of *aldus* and *dus*. In Middle Dutch these words could be exchanged freely, but in modern Dutch, *aldus* can only fulfill the anaphoric function, whereas *dus* can only be used as a causal connective or as a discourse marker. A third instance of specialization has occurred with *opdat* and *omdat*. In Middle Dutch both connectives could explicate causal relations with a desirable, but not yet realized consequent; in modern Dutch *omdat* has lost this possibility. As a result, *opdat* now has a specific profile that is clearly distinct from that of *omdat*. Even a fourth change, the objectification of *daarom* may be the result of specialization, it is possible that *daarom* has become more objective because the connective *dus* came to be used in the relatively subjective areas more often (cf. Stukker 2005). These results suggests that future studies into connective change should not restrict their object of research to isolated connectives, but that they should also take linguistic alternatives into account.

On the basis of my results, some remarks can be made on the range of the subjectification theory as proposed by Traugott (1995). It can be stated that subjectification only takes place at changes across functions, e.g. at the transition of lexical to functional (from deictic or anaphoric use to connective use), or at the transition from one text-linguistic function to another (from connective to discourse marker use). Contrary to Traugott's (1995) findings on *while*, subjectification does not occur between different uses within one connective function (like the domains within the use as causal connective). On the contrary, within the connective function, changes can even go in the opposite direction.

A first explanation for the absence of subjectification within the connective function has already been presented. It appears that connectives change when their conceptual profile is

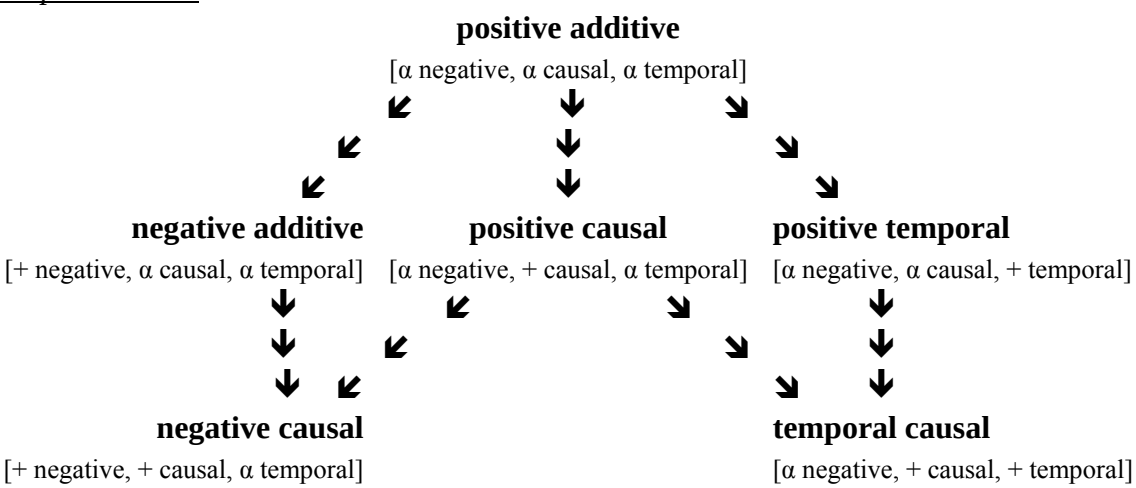
very close to the profile of another connective expressing a similar relation. The change resulting from this "specialization process" may either be a case of subjectification, or of objectification. The second explanation is that none of the connectives under discussion showed real changes in their domains of use in the sense that they came to be used in a domain in which they could not occur earlier. The four causal connectives could express all three domains from the 13th century on. It may be the case that subjectification only occurs when lexical items gain new meanings, and not when lexical items show a shift in the distribution over the meanings they can already express.

14.2.2 Highlights from the acquisition studies

In Chapters 10 to 12, I presented data on the acquisition of several Dutch connectives. As an explanation, I put forward cumulative complexity as the major determinant of the process of connective acquisition. This complexity approach consists of at least two components, both defined in a cumulative way: one component that establishes the conceptual complexity of connectives, and another that establishes the syntactic complexity of the clause combination. My multi-dimensional approach to cumulative conceptual complexity, presented in Chapter 10, differs from the one put forward by Bloom and colleagues in that the conceptual notions are not ordered along one dimension (additive < temporal < causal < adversative), but that each connective is defined on the basis of several cognitive primitives (taken from Sanders et al. 1992).

The cumulative complexity approach seemed to offer a solid explanation for the findings on Dutch and English connective acquisition. For both languages, a fixed and a variable part should be distinguished in the acquisition route of connectives. The fixed acquisition routes were explained by reference to the relative complexity of different values on the same conceptual primitive (e.g. [α causal] vs. [+ causal]), whereas the variation among acquisition routes of Dutch and English children could be explained on the basis of different interactions between the conceptual primitives that characterize each connective. This developmental picture can be schematically represented as in (2).

(2) <u>Acquisition orders based on the interaction between *basic operation*, *polarity*, and *temporal order*:</u>



The findings in Chapter 12 indicated that the cumulative complexity approach should be extended with a syntactic complexity component (cf. also Slobin 1973 on the relevance of formal complexity). For example, the development from clause-internal coordination to clause-combining coordination could be explained on the basis of cumulative syntactic complexity:

clause-combining coordination is syntactically more complex than clause-internal coordination because the child needs to deal with larger chunks of text (clauses instead of phrases). All in all, cumulative complexity can account both for the uniformity and for the diversity in the acquisition of connectives.

In order to give a full account of the emergence of all types of connectives, additional conceptual and syntactic primitives need to be incorporated in this multi-dimensional model. For example, within the temporal domain, further distinctions are needed to account for connectives expressing simultaneity. Within sequentiality, further distinctions such as priority and anteriority are required. Chapter 13 presented the first results for the cognitive primitive *source of coherence*. From the longitudinal study in that chapter, I can only conclude that epistemic relations seem to be acquired last. For content versus speech act I did not find a clear picture. In addition, it appeared that all three domains are acquired very early, that is, before the age of three. The most important conclusion from the acquisition study in Chapter 13 is that context plays a crucial role in the production of domain types. It was not possible to draw firm conclusions on the sequence of development there, because the collectors of the longitudinal child data did not systematically control for different contexts of use.

The exact range of the cumulative complexity approach needs to be tested by studying the influence of other factors in more detail. For example, the role of parental input could not be ruled out entirely on the basis of the analyses in Chapter 11. This first exploration of parental input to explain the data on the acquisition order only contained quantitative analyses of two children and their mother in their use of one connective, lacking a qualitative analysis of parental input. Still, a cautious conclusion from the data in Chapter 11 was that parental input cannot be regarded as a completely independent factor. Rather, it seems as if complexity sets the pace of children's connective acquisition and that parents apply principles of 'audience design': they adapt themselves to the abilities of their children. In this view, parents strengthen the effects of cumulative complexity on the acquisition process. Hence, it seems attractive to combine the cumulative complexity approach with usage-based approaches of language acquisition.

A complete account of the acquisition route can only be based on data from different contexts in which children acquire connectives. Such a picture can be obtained by using the methods advocated by Tomasello (2003) and his colleagues, e.g. using longitudinal studies of interactions between children and their parents. In addition, experimental studies can be used. For example, I have shown that specific hypotheses about the use of certain types of relations in a specific communicative context can successfully be tested in experiments with young children (see section 13.3).

14.3 Results pertaining to the form-function hypotheses

The first aim of this thesis was to gain more insight into the diachronic development and the acquisition of Dutch connectives. The second goal of my research was to develop a more precise relationship between the two disciplines of text and sentence linguistics in the field of connectives. In Chapter 2 I presented four primitives that can be used to characterize connectives in a syntactic way (see (3)).

(3) a. Linearization of the connective clause
b. Word order (i.e. positioning of the finite verb) within the matrix clause
c. Word order (i.e. positioning of the finite verb) within the connective clause
d. Positioning of the connective itself

In my discussion of existing literature on form-function relations (see Chapter 3), I argued that the first primitive in (3), linearization, can be related to the organization of discourse: initial adverbial clauses create a ground, restricting the way in which the upcoming clause(s) should be interpreted. Final adverbial clauses on the other hand, only modify the preceding clause. Hence, a clause in post-position is not merely a positional variant of a preposed one (and thereby always freely interchangeable with it) but a tool 'designed' for specific contexts of usage. The second primitive in (3), word order within a matrix clause that follows the adverbial connective clause, can be related to the text-linguistic interpretation of the coherence relation based on domains, although this mapping is not one-to-one.

The remaining two syntactic primitives in (3), word order within the connective clause and positioning of the connective itself, were used to formulate hypotheses about form-function relations pertaining connectives. These hypotheses were then tested on the basis of data on the diachronic development (Chapters 5-7) and the acquisition (Chapters 10-13) of several Dutch connectives. In addition, a specific hypothesis was formulated on form-function relations in diachronic changes involving subjectification. This hypothesis was tested in Chapter 8.

In this section, I discuss the results for the form-function hypotheses first presented in Chapter 3. Section 14.3.1 focuses on the word order within the connective clause. Section 14.3.2 looks at positioning of the connective itself. Finally, section 14.3.3 deals with form-function relations concerning subjectification. Section 14.3.4 presents a general conclusion.

14.3.1 Word order within the connective clause

The word order within the connective clause has been related to two different text-linguistic properties of adverbial clauses: (a) their level of attachment to the preceding discourse, and (b) their interpretation based on domains of use or source of coherence.

The first interaction hypothesis concerned the level of attachment of the connective clause. In Chapter 3 I put forward Verhagen's hypothesis about the interaction between word order and closure: connective clauses showing V2 could either attach to the preceding clause (resulting in the hierarchical structure [A-[B-C]]), or to a combination of two preceding clauses (resulting in the structure [[A-B]-C]); connective clauses showing V-late, on the other hand, only had the former option. In Chapter 5 this hypothesis was subject to a first test on the basis of data on the diachronic development of V2 *want* and V-late *omdat*. This diachronic study revealed that Verhagen's closure hypothesis was on the right track, but that it should be restricted to postposed connective clauses, since preposed V-late clauses proved not to be restricted to late-closure interpretations. For example, the preposed *omdat*-clause (4) is not only related to the B-clause, but to the combination of the two clauses following the *omdat*-clause.

(4) *Mar* [**C** ***v^{o}mb dat*** *hi* (= de leeuw) *naturlik ku^{o}nheit an hu^{o}me heuet.*] [**B** *so scaimt hi v^{o}me angst te hebene.*] [**A** *Jnde lopt den man v^{o}p. als hastelik als di^{e} man dru^{o}p siit.*]
(NM, 1270-1290)
'But because he (= the lion) has a natural courage in him, he is ashamed of having fear and attacks the man as soon as he looks at him.'

Given the limited availability of combinations of three-clause combinations in my samples, further research into Modern Dutch is imperative in order to find support either in favor of or against the modified version of Verhagen's hypothesis. The analyses in Chapter 12 revealed that the acquisition data on *want* and *omdat* were not suitable to provide such conclusive

evidence either. Eight children only used *want* during the recording periods, which implies that they did not have a choice in marking causal relations either with a V2 word order or with a V-late word order. The other eight children, who did use both *want* and *omdat*, did not produce combinations of three clauses. Therefore, there were no constructions in which they had to apply the closure rule.

The second hypothesis concerned the interaction between word order within the connective clause and domains of use. The diachronic data on *want* and *omdat* in Chapter 6 revealed that the hypothesis about the interaction between domains and word order was on the right track. It appeared that only V-late connectives in the bound subordination construction were restricted to the content domain, whereas V2 clauses in free coordination as well as V-late clauses in other constructions were not restricted in this respect. Further domain restrictions in the use of connectives seem to be triggered by connective-specific properties. The acquisition data did not provide conclusive evidence in this area. First of all, eight children did not use *omdat* during their recording periods, which implies that one of the word order variants was not available in their use of causal connectives. Furthermore, it could be concluded that the acquisition of *omdat* did not seem to affect the way children used *want*, since the children used *omdat* too infrequently, while *want* continued to be used in all three domains.

14.3.2 Positioning of the connective

The positioning of lexical items used as connectives has been related to two different text-linguistic properties of adverbial clauses, (a) their interpretation based on domains of use, and (b) their use in both connective and non-connective functions.

The diachronic analyses of *dus* and *daarom* in Chapter 7 revealed that the hypothesis about the interaction between domains and positioning should be rejected, since no significant interaction between the two was found. The acquisition data in Chapter 13 pointed in the same direction. Children only used the causal connective *dus* in the epistemic domain, whereas the vast majority of the *daarom*-clauses occurred in the content domain. Given that both connectives (with their different domain preferences) only appeared in clause-initial position, there is no indication that positioning is related to usage in different domains.

The second hypothesis concerned the interaction between different functions of a word and positioning. It appeared that the positioning of lexical items can indeed single out the connective function from other text-linguistic functions these words may have. The diachronic data in Chapter 7 revealed that positioning is used to discriminate between the connective function and the anaphoric function: in the 13th and 16th centuries, the connective function had a preference for the clause-initial position (see the example in (5)), whereas the non-causal anaphoric function had a preference for the clause-medial position (cf. (6)).

(5) *En nu die sommige out en cout / sy hebben niet waer mee*
dat sij haer hongerige buyck / sullen versaeden
***dus** toont aan haer u lieft* (Spel van sinnen, 1597)
'And now some old and cold ones, they don't have anything with which they can satisfy their hungry stomachs, so show your love to them.'

(6) *Als die Fransoysen die coninc **dus** hoirden roepen (…)*
(Historie van den vier heemskinderen, 1508)
'When the French heard the king call this way (…)'

In the 20th century, only the connective function remained, which could then be expressed either in the clause-initial or in the clause-medial function.

In the 20th century, the positioning of *dus* could be related to the accessibility of the information in the *dus*-clause. The discourse marker use only differs in positioning preferences from connective fragments that do not contain accessible conclusions (compare the positionings of discourse marker *dus* in (7) to the positioning of *dus* in the non-accessible conclusion in (8)). In those 20th-century connective fragments in which *dus* marks the information as accessible, *dus* shows a behavior similar to *dus* in discourse marker fragments: it has a preference for the clause-medial position (cf. (9)).

(7) [Context: this is a report of the opening of a swimming pool. The information that the opening was done by Hilde Zoer had been mentioned earlier in the text.]
Op het moment dat de eerste burger zich openlijk afvroeg wie de openingshandeling wilde verrichten, ging een groot aantal vingers de lucht in. Het werd ***dus*** *Hilde Zoer.* (Meppeler Courant, 1995)
'At the moment the mayor explicitly asked who wanted to do the opening ceremony, many people raised their hands. Hilde Zoer, then, became the person to do it.'

(8) *Patrick woonde in een zijstraat van de laan van Nieuw Guinea,* ***dus*** *Hendriks* <u>*kwam*</u> *er vlak langs.* (De kunstrijder, 1989)
'Patrick lived in a road off the New Guinea's avenue, so Hendriks came right past it.'

(9) [Topic is the number of Sundays in a year; in 1995 there were 53 instead of 52 Sundays.]
Als ik het goed heb uitgerekend, herhaalt zich deze situatie in het veelbesproken jaar 2000, want dan valt nieuwjaarsdag op een zaterdag en de eerste zondag ***dus*** *op 2 januari.* (MC, 1995)
'If I calculated it correctly, this situation repeats itself in the much-discussed year 2000, because then New Year's Day will fall on a Saturday and thus the first Sunday on the second of January.'

The acquisition data on *dus* provided additional support for this interaction: these data revealed that the language use of children as young as four supported the conceptual distinction between the connective use and the discourse marker use of *dus* with specific positioning preference (clause-initial versus clause-medial).

It can be concluded, then, that syntactic characteristics based on positioning only support one difference at the text-linguistic level at one time. In the case of *daarom* and *dus*, positioning is used to support the distinction between connective and non-connective use, and not to support the distinction based on domains of use.

14.3.3 Subjectification

A final interaction hypothesis concerned the interaction between subjectification and grammaticalization. More precisely, I investigated whether syntactic changes were a necessary prerequisite for subjectification to occur. Chapter 8 showed that subjectification was only found at changes from or to the conceptual function of a causal connective, but not within that connective function. Examples are the decrease of the anaphoric use of *dus* and *daarom* and the increase in the new discourse marker use of *dus*. Both developments led to an overall subjectification in the use of these words. The diachronic data on *dus* and *daarom* revealed that subjectification should not be tied to syntactic changes. It appeared that changes in the positioning or categorical status of lexical items were not a necessary condition for their subjectification.

14.3.4 Conclusion

In Table 14.2, I present an overview of the various form-function relations that have been proposed, and the results that were reported in the literature or were obtained from the diachronic studies and acquisition studies in this thesis.

Table 14.2. Overview of claims and results on form-function interactions

Syntactic property	Interaction with conceptual properties?
Linearization of connective clause	Yes: preposed clauses serve a grounding function (cf. Thompson 1985; Ramsay 1987, Degand 2001)
Word order within matrix clause	Yes: in the case of conditional clauses, non-integrative word order is only possible with adverbial clauses that are separately assertable (cf. König & Van der Auwera 1988; Van Belle 1997)
Word order within connective clause	Yes: postposed V-late clauses force a "late closure" interpretation (cf. Verhagen 2001 and the results in Chapter 5)
Word order within connective clause	Yes: V-late clauses in the bound subordination construction force a content interpretation (cf. Verstraete 2000 and the results in Chapter 6)
Positioning of the connective	Yes: positioning discriminates between connective and non-connective functions of a word (cf. the results in Chapter 7)
Positioning of the connective	No: positioning does not discriminate between domains of use within the connective function (cf. the results in Chapter 7 and 13)
Positioning of the connective	Yes: positioning discriminates between accessible and non-accessible information (cf. Ariel 1988 and the results in Chapter 7 and 13)
Categorical status	No: subjectification is not restricted to changes in categorical status (Chapter 8)

The majority of these results suggest that the syntactic properties and the text-linguistic properties of connectives are indeed related. Different syntactic structures, which are equally grammatical from a traditional grammatical perspective, can be used to support different functions at the text-linguistic level. This is in line with the approach in Bates & MacWhinney (1989: 3), who state that "the forms of natural language are (…) used in the service of communicative functions." However, it has been observed that many of the form-function mappings are not one-to-one. Often, only one of the syntactic options is constrained in its text-linguistic interpretations, whereas the remaining syntactic options are not restricted in this respect.

14.4 Discussion

In the remainder of this chapter, three points will be taken up for discussion. Section 14.4.1 discusses the scope and limitations of this thesis. Section 14.4.2 deals with the link between developments in diachronic change and in child language acquisition . Section 14.4.3 concludes this thesis with some remarks on the psychological status of the form-function relations.

14.4.1 Scope and limitations of this thesis

A final point I want to mention here is the methodology of this study. An important goal of this research was to develop more objective and quantifiable methods to test the hypotheses about form-function interactions. In order to test my hypotheses, I made use of both qualitative and quantitative methods. These methods indeed proved to be complementary to one another. The qualitative analyses provided insight in the different usage possibilities of lexical items at particular point in time, both in the diachronic study and in the acquisition study. The quantitative analyses facilitated tracking and explaining linguistic change, because they resulted in insights that cannot be obtained from the study of exemplary fragments. In addition, they allowed for statistical testing of specific hypotheses. Although I restricted the number of centuries and fragments that were analyzed, the combination of qualitative and quantitative approaches seems fruitful. With the increasing possibilities of analyzing data with the help of computer programs, it should be relatively easy to study larger corpora.

My analysis of existing data on the acquisition of connectives revealed the importance of formulating clear definitions. The concept of 'acquisition' was defined in different ways in research concerning connective acquisition, and these differences also led to different results. This variation in results could be traced back to differences in the definition of the term 'acquisition'. This notion covers three types of 'acquisition': the emergence of connectives, the process of further development, and finally, full mastery of connectives. The discussion in section 9.3 has shown that several methods are suitable to establish connective acquisition in all its varieties. A full picture from longitudinal data on connective acquisition can be gained by (a) looking at the first correct and creative use of connectives; (b) analyzing intra- and interindividual patterns of acquisition using developmental curves; and (c) analyzing the connective utterances produced at different ages in a qualitative way.

An additional methodological consideration is that a developmental study should not be restricted to a particular lexical item or a specific syntactic construction. On the contrary, a full picture of both connective change and connective acquisition can only be obtained by taking other connectives or other syntactic options with related functions into account. My diachronic analysis revealed that connectives can change as a result of competition with another connective. In addition, children may change their use of a certain connective because they acquire other connectives.

Furthermore, serious claims about diachronic developments and child language developments in the use of text-linguistic items can only be obtained by taking into account the different contexts in which these items are used. In Chapter 13 on the acquisition of domains, I argued that it was not possible to draw firm conclusions on the sequence of development, because the selection of child data did not systematically control for different contexts of use. However, the role of context or conversation type is also relevant for diachronic research, since the interpretation of text-linguistic relations is partially dependent on text genres (cf. Sanders 1997). In order to control for genre-effects, it is possible to study translations of one book in several periods, thus keeping the genre constant. For example, Schoonenboom (2000) investigated diachronic changes in the use of relative pronouns by studying biblical fragments from five different periods. Alternatively, it is possible to control for genre-effects by incorporating different genres into the sample. This is the approach that was used in this thesis.

14.4.2 On the link between diachrony and acquisition

In this thesis, I have chosen to study both connective change and connective acquisition. This choice does not imply that I expect that children's acquisition of connectives necessarily

follows the same route as the diachronic development these connectives have gone through. As has been mentioned repeatedly, it is regularly the case that "the developmental paths are parallel in diachrony and ontogeny" (Slobin 1994: 120), but this does not imply that the child always recapitulates the history of the connective under discussion. Formulating it in terms of the famous slogan: "ontogeny need not recapitulate phylogeny" (cf., among others, Traugott & Dasher 2002: 42-44). Children are only occasionally confronted with an ambiguity from the historical residue (Nishigauchi & Roeper 1987: 96). In Slobin's (1994: 128) words: "the parallels are, in a sense, illusory."

The acquisition data in this thesis indeed show that children do not necessarily follow the historical path of development. For instance, young Dutch children do not use *want* with a V-late word order, despite the fact that the V-late word order did occur in the use of 13th-century *want*. Their use of *want* is restricted to the V2 word order, the only grammatical option in Modern Dutch. Similarly, Dutch children do not use *dus* in an anaphoric way, despite the fact that *dus* started out as an anaphor centuries ago. On the contrary, it seems as if children start using *dus* as a discourse marker, the function that appeared last in the diachronic development of this lexical item. The absence of V-late *want* and anaphoric *dus* need not surprise us, since these particular usages of the lexical items are not attested in Modern Dutch parental input at all.

An accidental parallel between the two developmental paths can be found in the acquisition of words that can be used both as an adverb and as a complementizer. Historically, this ambiguity arises from adverbs that gain a complementizer function. Some Dutch examples are found in the diachronic development of *dus* (see Chapter 7) and the development of *toen* (see Van Es 1954, 1955). The data on the acquisition of Dutch *toen* revealed that children also start out with the adverbial use, and only later acquire the complementizer use of *toen*. Nishigauchi & Roeper (1987: 96-97) argue that it is quite plausible that the history of linguistic elements might be recapitulated if it entailed natural acquisition steps. Similarly, Slobin (1994: 131, footnote 7) observes that "the ontogenetic course is based on developmental psychological factors." Certain types of connective use are cognitively difficult for children, as becomes apparent from later acquisition of these uses. In contrast, new meanings of grammatical forms arise in adult language use on the basis of pragmatic inferences drawn from existing referential and propositional meanings. Although the diachronic and ontogenetic developments can sometimes be parallel, they seem to result from different processes (Slobin 1994: 129-130).

Thus, a full account of Dutch acquisition must be based on a detailed unfolding of cognitive capacities, while a historical account would rely on inferences made by cognitively mature adults as to the possible meanings of forms in actual usage (compare the recommendations in Slobin 1994: 131).

14.4.3 On the psychological status of the interactions

The psychological plausibility of the form-function interactions can be shown from the manipulative use or rhetorical use of certain connectives (cf. Van den Hoven 1997). In the case of *dus*, language users sometimes choose to make "mistakes" in the form-function relations in order to reach a specific effect. Normally, *dus* is placed in clause-medial positions if it marks information that is accessible. It appears in clause-initial if it marks new conclusions. However, in the rhetorical use of *dus*, the language user neglects this "rule". By placing the connective *dus* in a clause-medial position, despite the fact that the *dus*-clause does not contain an accessible conclusion, the language user presents his conclusion as obvious. This rhetorical use of *dus* may come in handy if he wants to rule out discussion

beforehand. More importantly, this rhetorical use can only exist because language users understand and use the general 'positioning-function rule'.

The results in this thesis revealed that the mapping between form and functions is often not one-to-one. This raises the question what the psychological status of these interaction "rules" is. For example, what happens if language users make mistakes in the syntactic positioning of discourse marker *dus*? It appears that such mistakes are less drastic than mistakes in, for example, the word order within a subordinate clause. The latter mistakes result in ungrammaticality of the connective clause, whereas the former will probably 'only' raise some eyebrows with respect to the intended text-linguistic interpretation (or they can have a special effect). A cautious conclusion is that – as far as these systematic relations between text-linguistic and sentence-linguistic properties can be seen as rules – these interaction rules are probabilistic in nature, as other regularities in language use are. Clearly, this conclusion does not imply that language users cannot rely on such rules; after all, it has been shown that the regularities for connective usage clearly differed from chance. Therefore, the further linguistic challenge is to understand when these regularities do, or do not show up, what are the contexts, which the restrictions?

A similarly important question is: Why are the form-function mappings often not one-to-one? One possibility is that "grammars can only be viewed as a class of partial solutions to the problem of mapping nonlinear meanings onto a highly constrained linear medium whose only devices are word order, lexical marking, and suprasegmentals" (Bates & MacWhinney 1989: 8). According to Bates & MacWhinney (1989: 13), there is a rich array of contextual factors governing the formal device in question, and hence, "probabilistic mappings do not reflect a failure on the part of the linguist (or psycholinguist)." This line of reasoning does not imply that linguists should stop looking for and accounting for specific form-function relations. To the contrary, it urges linguists to look for other factors that disturb the one-to-one mapping. One option in this area is this that there are other form-function mappings that interfere (cf. the optimality approach in Prince & Smolensky 2004). For example, the topic position of an adverbial clause can host adverbial connectives like *dus* and *daarom* if they occur in clause-initial position. However, this topic position can also be used to stress one of the constituents within that same adverbial clause. Since the adverbial connective and the stressed constituent cannot be placed in topic position at the same time, the language user may choose to put the connective in clause-medial position, despite the fact that this positioning would not support the intended text-linguistic function. Hence, the original mapping between form and function cannot be reached.

The findings in this thesis suggest that more research is needed to gain more insight into the interaction between different form-function relations. My empirical research based on specific operationalizations shows how (a) both quantitative and qualitative analyses, (b) both diachronic and acquisition data, and (c) both text-linguistic and sentence-linguistic approaches can contribute to the development of a theory about form-function relations in the use of connectives and other text-linguistic elements.

References

Abraham, W. (1993). Grammatikalisierung und Reanalyse: einander ausschlieβende oder ergänzende Begriffe? 'Grammaticalization and reanalysis: Mutually exclusive or supplementary concepts?' *Folia Linguistica Historica* 13/1-2, 7-26.

Andersen, R.W. (1978). An implicational model for second language research. *Language Learning* 28, 221-282.

Ariel, M. (1988). Retrieving propositions from context: why and how. *Journal of Pragmatics* 12/5-6, 567-600.

Ariel, M. (1990). *Accessing noun-phrase antecedents.* London: Routledge.

Ariel, M. (1999). Mapping so-called "pragmatic" phenomena according to a "linguistic-extralinguistic" distinction; the case of propositions marked "accessible". In: M. Darnell, E. Moravesik, F. Newmeyer, M. Noonan & K. Wheatly (1999). *Functionalism and formalism in linguistics*. *Volume II: Case studies*. Amsterdam/Philadelphia: John Benjamins, 11-38.

Auer, P. (2000). Pre- and post-positioning of *wenn*-clauses in spoken and written German. *InList* 15, March 2000. URL: <http://www.uni-potsdam.de/u/inlist/issues/15/index.htm>.

Auer, P. & S. Günthner (2003). Die Entstehung von Diskursmarkern im Deutschen – ein Fall von Grammatikalisierung? *InList* 38, December 2003. URL: <http://www.uni-potsdam.de/u/inlist/issues/38/index.htm>.

Banfield, A. (1973). Narrative style and the grammar of direct and indirect speech. *Foundations of language* 10, 1-39.

Bates, E. & B. MacWhinney (1989). Functionalism and the competition model. In: B. MacWhinney & E. Bates (eds.). *The crosslinguistic study of sentence processing*. Cambridge: Cambridge University Press, 3-73.

Bateman, J.A. & K.J. Rondhuis (1997). Coherence relations: Towards a general specification. *Discourse Processes* 24, 3-49.

Berko, J. (1958). The child's learning of English morphology. *Word* 14, 150-177.

Berman, R.A. (1996). Form and function in developing narrative abilities. In: D.I. Slobin, J. Gerhardt, A. Kyratzis & J. Guo (eds.), *Social interaction, social context, and language*. Mahwah, New Jersey: Lawrence Erlbaum, 343-367.

Blom, C. (2002). Word order in Middle Dutch: The interpretation of different types of data. In: H. Broekhuis & P. Fikkert (eds.), *Linguistics in the Netherlands 2002*. Amsterdam/Philadelphia: John Benjamins, 13-24.

Bloom, L. (1991). *Language development from two to three*. Cambridge: Cambridge University Press.

Bloom, L., M. Lahey, L. Hood, K. Lifter & K. Fiess (1980). Complex sentences: acquisition of syntactic connectives and the semantic relations they encode. *Journal of Child Language* 7, 235-261. Reprinted in: L. Bloom (1991), *Language development from two to three*. Cambridge: Cambridge University Press, 261-289.

Bloom, L., P. Lightbown & L. Hood (1975). Structure and variation in child language. *Monographs of the Society for research in child development* 40/2 (serial no. 160).

Bol, G.W. (1996). Optional subjects in Dutch child language: In: C. Koster & F. Wijnen (eds.), *Proceedings of the Groningen assembly on language acquisition held at the University of Groningen, 7-9 September 1995*, Groningen: Center for language and cognition Groningen, 125-135.

Bol, G. & F. Kuiken (1988). *Grammaticale analyse van taalontwikkelingsstoornissen*. Ph.D. dissertation University of Amsterdam, Utrecht: Elinkwijk.

Bouman, A.C. (1918). *Bijdrage tot de syntaxis der "dat"-zinnen in het Germaansch.* Utrecht: Electrische Drukkerij L.E. Bosch & Zoon.

Bowerman, M. (1975). Commentary. In: Bloom, Lightbown & Hood (1975: 80-90).

Bowerman, M. (1979). The acquisition of complex sentences. In: P. Fletcher & M. Garman (eds.), *Language acquisition: Studies in first language development*. Cambridge: Cambridge University Press, 285-305.

Braunwald, S.R. (1985). The development of connectives. *Journal of Pragmatics* 9, 513-525.

Braunwald, S.R. (1997). The development of *because* and *so*: Connecting language, thought, and social understanding. In: J. Costermans & M. Fayol (eds.), *Processing interclausal relationships: Studies in the production and comprehension of text.* New Jersey: Lawrence Erlbaum Associates, 121-137.

Britton, B.K. (1994). Understanding expository text: Building mental structures to induce insights. In: M.A. Gernsbacher (ed.), *Handbook of psycholinguistics*. San Diego, etc.: Academic Press, 641-674.

Brown, R. (1973). *A first language*. Cambridge, Mass.: Harvard University Press.

Brown, R., C. Cazden & U. Bellugi-Klima (1969). The child's grammar from I to III. In: R. Brown (ed.) (1970), *Psycholinguistics: Selected papers by Roger Brown*. New York: Free Press, 100-154. Reprinted from J.P. Hill (ed.), *Minnesota Symposia on Child Language, Volume 2*. Minneapolis: University of Minnesota Press, 28-73.

Brown, R. & C. Hanlon (1970). Derivational complexity and order of acquisition in child speech. In: R. Brown (ed.), *Psycholinguistics: Selected papers by Roger Brown*. New York: Free Press, 155-207. Reprinted from: J.R. Hayes (ed.), *Cognition and the development of language*. New York, etc.: Wiley.

Burridge, K. (1993). *Syntactic change in Germanic: Aspects of language change in Germanic with particular reference to Middle Dutch*. Amsterdam/Philadelphia: John Benjamins.

Bybee, J., R. Perkins & W. Pagliuca (1994). *The evolution of grammar; Tense, aspect, and modality in the languages of the world.* Chicago/Londen: The University of Chicago Press.

Byrnes, J.P. & S.A. Gelman (1991). Perspectives on thought and language: Traditional and contemporary views. In: S.A. Gelman & J.P. Byrnes (eds.), *Perspectives on language and thought: Interrelations in development*. New York: Cambridge University Press, 3-27.

Chafe, W. (1984). How people use adverbial clauses. In: C. Brugman & M. Macaulay (eds.), *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society, 437-449.

Cinque, G. (1999). *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: Oxford University Press.

Clancy, P., T. Jacobsen & M. Silva (1976). The acquisition of conjunction: A cross-linguistic study. *Stanford Papers and Reports on Child Language Development* 12, 71-80.

Clark, E.V. (1970). How young children describe events in time. In: G.B. Flores D'Arcais & W.J.M. Levelt (eds.), *Advances in psycholinguistics*. Amsterdam: North-Holland, 275-284.

Clark, E.V. (1973). How children describe time and order. In: C.A. Ferguson & D.I. Slobin (eds.), *Studies of child language development*. New York: Holt, Rinehart and Winston, 585-606.

Clark, H.H. & E.V. Clark (1977). *Psychology and language; An introduction to psycholinguistics*. New York etc.: Harcourt Brace Jovanovich.

Couper-Kuhlen, E. (1996). Intonation and clause combining in discourse: the case of *because*. *Pragmatics* 6/3, 389-426.

Cozijn, R. (2000). *Integration and inference in understanding causal sentences*. Ph.D. dissertation, Tilburg University.

Crevels, M. (2000). Concessives on different semantic levels: A typological perspective. In: E. Couper-Kuhlen & B. Kortmann (eds.), *Cause, condition, concession, contrast; Cognitive and discourse perspectives*. Berlin: Mouton de Gruyter, 313-339.

Cuenca, M.J. (1997). Form-use mappings for tag questions. In: W.A. Liebert, G. Redeker & L. Waugh (eds.), *Discourse and Perspective in Cognitive Linguistics*. Amsterdam: Benjamins, 3-19.

Dancygier, B. & E. Sweetser (1996). Conditionals, distancing, and alternative spaces. In: A.E. Goldberg (ed.), *Conceptual structure, language and discourse*. Stanford: CSLI Publications, 83-98.

Dancygier, B. & E. Sweetser (2000). Constructions with *if*, *since*, and *because*: Causality, epistemic stance, and clause order. In: E. Couper-Kuhlen & B. Kortmann (eds.), *Cause, condition, concession, contrast; Cognitive and discourse perspectives*. Berlin: Mouton de Gruyter, 111-142.

Dasher, R. (1995). *Grammaticalization in the system of Japanese predicate honorifics*. Ph.D. Dissertation, Stanford University.

Degand, L. (1996). *A situation-based approach to causation in Dutch with some implications for text generation.* Ph.D. dissertation, Université Catholique de Louvain.

Degand, L. (2000). Contextual constraints on causal sequencing in informational texts. *Functions of language* 7/2, 173-201.

Degand, L. (2001). *Form and function of causation; A theoretical and empirical investigation of causal constructions in Dutch*. Leuven etc.: Peeters. Studies op het gebied van de Nederlandse taalkunde 5.

Degand, L. & T. Sanders (1999). Causal connectives in language use. Theoretical and methodological aspects of the classificiation of coherence relations and connectives. *Working Notes workshop Levels of representation in discourse, Edinburgh, July 7-9 1999*, 3-11.

Degand, L. & T. Sanders (2002). The impact of relational markers on expository text comprehension in L1 and L2. *Reading and Writing* 15 (7-8), 739-757.

De Haan, G. (1987). A theory-bound approach to the acquisition of verb placement in Dutch. In: G. de Haan & W. Zonneveld (eds.), *Formal parameters of generative grammar; OTS Yearbook III*. Dordrecht: ICG Publications, 15-30.

De Haan, G.J. (2001). More is going on upstairs than downstairs: Embedded root phenomena in West Frisian. *Journal of Comparative Germanic Linguistics* 4, 3-38.

De Rooij, J. (1982). Omdat en doordat in het Nederlands. *De Nieuwe Taalgids* 75, 329-342.

De Villiers, J.G. & P.A. de Villiers (1973). A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research* 2, 267-278.

De Vries, J.W. (1971). Want en omdat. *De Nieuwe Taalgids* 64, 414-420.

De Vries, M., L.A. te Winkel et al. (eds.) (1882-1998). *Woordenboek der Nederlandsche Taal* [Dictionary of the Dutch language]. 's-Gravenhage/Leiden etc.: M. Nijhoff/A.W. Sijthoff etc. Also available on CD-rom: *Het Woordenboek der Nederlandsche Taal op CD-Rom* (second release, 2000). Rotterdam: AND.

Diessel, H. (2004). *The acquisition of complex sentences*. Cambridge: Cambridge University Press.

Diessel, H. & M. Tomasello (2000). The development of relative clauses in spontaneous child speech. *Cognitive Linguistics* 11-1/2, 131-151.

Diessel, H. & M. Tomasello (2001). The acquisition of finite complement clauses in English: A corpus-based analysis. *Cognitive Linguistics* 12/2, 97-141.

Dik, S.C. (1968). *Coordination: its implications for the theory of general linguistics*. Amsterdam: North Holland.

Dik, S., K. Hengeveld, E. Vester & C. Vet (1990). The hierarchical structure of the clause and the typology of adverbial satellites. In: J. Nuyts, M. Bolkestein & C. Vet (eds.), *Layers and levels of representation in language theory*. Amsterdam: Benjamins, 25-70.

Dubinsky, S. & K. Williams (1995). Recategorization of prerpositions as complementizers: The case of temporal prepositions in English. *Linguistic Inquiry* 26/1, 125-137.

Ehlich, K. (1994). Funktionale Etymologie. In: G. Brünner & G. Graefen (eds.), *Texte und Diskurse; Methoden und Forschungsergebnisse der Funktionalen Pragmatik*. Opladen: Westdeutscher Verlag, 68-82.

Eisenberg, A.R. (1980). A syntactic, semantic and pragmatic analysis of conjunction. *Stanford Papers and Reports on Child Language Development* 19, 129-138.

Elffers, E. (1992). Wat betekent *toch* toch? In: E.C. Schermer-Vermeer e.a. (eds.), *De kunst van de grammatica; Artikelen aangeboden aan Frida Balk-Smit Duyzentkunst bij haar afscheid als hoogleraar Taalkunde van het hedendaags Nederlands aan de Universiteit van Amsterdam.* Amsterdam: Vakgroep Nederlandse Taalkunde, Universiteit van Amsterdam, 63-80.

Elman, J.L. (1993). Learning and development in neural networks: The importance to start small. *Cognition* 48, 71-99.

Ernst, T.B. (1984). *Toward an integrated theory of adverb position in English*. Indiana University Linguistics Club.

Evers, A. & J. van Kampen (2001). *E-language, I-language and the order of parameter setting.* UiL OTS Working Papers 00105-S-S. Available: <http://www.let.uu.nl/~Jacqueline.vanKampen/personal/downloadables/Syntax.pdf>.

Evers-Vermeul, J. (2000). De complexiteit van connectiefverwerving [The complexity of connective acquisition]. *Nederlandse Taalkunde* 5/3, 251-271.

Evers-Vermeul, J. & N. Stukker (2003). Subjectificatie in de ontwikkeling van causale connectieven? De diachronie van *daarom*, *dus*, *want* en *omdat*. *Gramma/TTT* 9-2/3, 111-139.

Fauconnier, G. (1985). *Mental spaces*. Cambridge: Cambridge University Press.

Fienberg. S.E. (1977). *The analysis of cross-classified categorical data.* Cambridge, MA: MIT Press.

Fischer, O., A. van Kemenade, W. Koopman & W. van der Wurff (2000). *The syntax of Early English*. Cambridge: Cambridge University Press.

Ford, C.E. (1993). *Grammar in interaction. Adverbial clauses in American English conversations*. Cambridge: Cambridge University Press.

Ford, C.E. & S.A. Thompson (1986). Conditions in discourse: a text-based study from English. In: E.C. Traugott et al. (eds.), *On conditionals*. Cambridge: Cambridge University Press, 353-372.

Fraser, B. (1996). Pragmatic markers. *Pragmatics* 6(2), 167-190.

Genetti, C. (1991). From postposition to subordinator in Newari. In: E.C. Traugott & B. Heine (eds.), *Approaches to grammaticalization*, vol. II. Amsterdam/Philadelphia: Benjamins, 227-255.

Gerritsen, M. (1982). Word order change in Dutch imperative clauses: the interaction between contextual and syntactic factors. In: A. Ahlqvist (ed.), *Papers presented from the 5th International conference on historical linguistics*. Amsterdam: Benjamins, 62-73.

Gerritsen, M. (1987). *Syntaktische verandering in kontrolezinnen; een sociolinguïstische studie van het Brugs van de 13e tot de 17e eeuw* [Syntactic change in infinitive constructions. A sociolinguistic study of the dialect of Bruges between the 13th and 17th century]. Dordrecht: ICG Printing.

Givón, T. (1987). Beyond foreground and background. In: R.S. Tomlin (ed.), *Coherence and grounding in discourse*. Amsterdam: John Benjamins, 175-188.

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

Gohl, C. & S. Günthner (1999). Grammatikalisierung von 'weil' als Diskursmarker in der gesprochene Sprache. *Zeitschrift für Sprachwissenschaft* 18, 39-75.

Goldstein, H. (1995). *Multilevel statistical models. Second edition*. London: Deward Arnold.

Goorhuis-Brouwer, S.M. (1997). *Het wonder van de taalverwerving; basisboek voor opvoeders van jonge kinderen*. Utrecht: De Tijdstroom.

Greenbaum, S. (1969). *Studies in English adverbial usage*. London: Longman.

Günthner, S. (1993). "...weil - man kann es ja wissenschaftlich untersuchen" - Diskurspragmatische Aspekte der Wortstellung in WEIL-Sätzen. *Linguistische Berichte* 143, 37-59.

Günthner, S. (1996). From subordination to coordination? Verb-second in German clausal and concessive constructions. *Pragmatics* 6/3, 323-356.

Günthner, S. (2000). From concessive connector to discourse marker: The use of *obwohl* in everyday German interaction. In: E. Couper-Kuhlen & B. Kortmann (eds.), *Cause, condition, concession, contrast; Cognitive and discourse perspectives*. Berlin: Mouton de Gruyter, 439-468.

Haegeman, L. (1994). Verb raising as verb projection raising: Some empirical problems. *Linguistic Inquiry* 25/3, 509-522.

Haegeman, L. (2001). *Anchoring to speaker and the structure of CP*. Ms. Université Charles de Gaulle, Lille III.

Haegeman, L. (2003). Conditional clauses: external and internal syntax. *Mind & Language* 18/4, 317-339.

Haegeman, L. & H. van Riemsdijk (1986). Verb projection raising, scope, and the typology of rules affecting verbs. *Linguistic Inquiry* 17/3, 417-466.

Haesereyn, W., K. Romijn, G. Geerts, J. de Rooij & M.C. van den Toorn (eds.) (1997). *Algemene Nederlandse Spraakkunst (ANS)*. Groningen: Martinus Nijhoff.

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

Harris, A.C. & L. Campbell (1995). *Historical syntax in cross-linguistic perspective.* Cambridge: Cambridge University Press.

Heersche, J.P.G. (1991). *Syntactische verschijnselen in het Vroegmiddelnederlands; een onderzoek naar de bouw van begin- en eindgroep in Vroegmiddelnederlands ambtelijk proza*. Ph.D. dissertation University of Amsterdam.

Heine, B., U. Claudi & F. Hünnemeyer (1991). *Grammaticalization; A conceptual framework*. Chicago/London: University of Chicago Press.

Hengeveld, K. (1997). Cohesion in functional grammar. In: J.H. Connolly, R.M. Vismans, C.S. Butler & R.A. Gatward (eds.), *Discourse and pragmatics in functional grammar*. Berlin/New York: Mouton de Gruyter, 1-16.

Hobbs, J. (1979). Coherence and coreference. *Cognitive Science* 3, 67-90.

Hooper, J. & S. Thompson (1973). On the applicability of root transformations. *Linguistic Inquiry* 4, 465-497.

Hopper, P.J. & E.C. Traugott (1993). *Grammaticalization*. Cambridge: Cambridge University Press.

Hovy, E.H. (1990). Parsimonious and profligate approaches to the question of discourse structure relations. *Proceedings of the 5th International Workshop on Natural Language Generation*. Pittsburg, PA.

Huijsinga, A. (1953). Dan ook – immers. *De Nieuwe Taalgids* 46, 147-153.

Ingram, D. (1981). Early patterns of grammatical development. In: R. Stark (ed.), *Language behavior in infancy and early* childhood. New York: Elsevier North-Holland, 327-352.

Ingram, D. (1989). *First language acquisition: Method, description, and explanation*. Cambridge: Cambridge University Press.

Jackendoff, R. (1972). *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

Kehler, A. (2002). *Coherence, reference, and the theory of grammar*. Stanford, CA: CSLI Publications.

Keller, R. (1995). The epistemic *weil*. In: Stein, D. & S. Wright (eds.), *Subjectivity and subjectivisation: linguistic perspectives*. Cambridge: Cambridge University Press, 16-30.

Klein Gunnewiek, L. (1999). Acquisition sequence of German: A comparison of cross-sectional versus longitudinal data. In: J. Don & T. Sanders (eds.), *Utrecht Institute of Linguistics OTS Yearbook 1998-1999*. Utrecht: Utrecht Institute of Linguistics OTS, 48-57.

Knott, A. & R. Dale (1994). Using linguistic phenomena to motivate a set of coherence relations. *Discourse Processes* 18: 35-62.

Knott, A. & T. Sanders (1998). The classification of coherence relations and their linguistic markers: An exploration of two languages. *Journal of Pragmatics* 30, 135-175.

Koelmans, L. (1979). *Inleiding tot de historische taalkunde van het Nederlands*. Utrecht: Bohn, Scheltema & Holkema.

König, E. & J. van der Auwera (1988). Clause integration in German and Dutch conditionals, concessive conditionals, and concessives. In: J. Haiman & S.A. Thompson (eds.), *Clause combining in grammar and discourse*. Amsterdam/Philadelphia: John Benjamins, 101-133.

Krikhaar, E. (1992). *Voegwoordloze bijzinnen in kindertaal*. Unpublished MA-thesis Utrecht University.

Kruyt, J.G. & M.W.F. Dutilh (1997). A 38 million words Dutch text corpus and its users, *Lexikos* 7, 229-244.

Kuno, S. (1987). *Functional syntax. Anaphora, discourse and empathy*. Chicago: University Press.

Kyratzis, A., J. Guo & S. Ervin-Tripp (1990). Pragmatic conventions influencing children's use of causal constructions in natural discourse. In: K. Hall et al. (ed.) *Proceedings of the sixteenth annual meeting of the Berkeley Linguistics Society*, 205-214.

Lagerwerf, L. (1998). *Causal connectives have presuppositions; Effects on coherence and discourse structure*. Ph.D. dissertation, Katholieke Universiteit Brabant, Den Haag: Holland Academic Graphics.

Lambrecht, K. (1988). Presentational cleft constructions in spoken French. In: J. Haiman & S.A. Thompson (eds.), *Clause combining in grammar and discourse*. Amsterdam/Philadelphia: John Benjamins, 135-179.

Lamiroy, B. & W. van Belle (1995). Connectives of contrast and concession in Dutch and French. *Leuvense Bijdragen* 84/3, 397-418.

Lang, E. (2000). Adversative connectors on distinct levels of discourse: A re-examination of Eve Sweetser's three-level approach. In: E. Couper-Kuhlen & B. Kortmann (eds.), *Cause, condition, concession, contrast; Cognitive and discourse perspectives*. Berlin: Mouton de Gruyter, 235-256.

Langacker, R.W. (1977). Syntactic reanalysis. In: C.N. Li (ed.), *Mechanisms of syntactic change*. Austin/London: University of Texas Press, 57-139.

Langacker, R.W. (1985). Observations ans speculations on subjectivity. In: J. Haiman (ed.), *Iconicity in syntax*. Amsterdam/Philadelphia: John Benjamins, 109-150.

Langacker, R.W. (1990). Subjectification. *Cognitive Linguistics* 1: 5-38.

Leech, G. (1974). *Semantics*. Hardmondsworth: Penguin.

Lehmann, W.P. (1972). Proto-Germanic syntax. In: F. van Coetsem & H.L. Kufner (eds.), *Toward a grammar of Proto-Germanic*. Tübingen: Niemeyer, 239-268.

Lehmann, C. (1988). Towards a typology of clause linkage. In: J. Haiman & S.A. Thompson (eds.), *Clause combining in grammar and discourse*. Amsterdam/Philadelphia: John Benjamins, 181-226.

Lenk, U. (1998). *Marking discourse coherence. Functions of discourse markers in spoken English*. Tübingen: Narr.

Lust, B. & C.A. Mervis (1980). Development of coordination in the natural speech of young children. *Journal of Child Language* 3, 309-318.

MacWhinney, B. (1995). *The CHILDES project: Tools for analyzing talk.* Hillsdale, NJ: Lawrence Erlbaum Associates.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk, Third edition.* Mahwah, NJ: Lawrence Erlbaum Associates.

Mann, W.C. & S.A. Thompson (1986). Relational propositions in discourse. *Discourse Processes* 9, 57-90.

Mann, W.C. & S.A. Thompson (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text: an Interdisciplinary Journal for the Study of Discourse* 8, 243-281.

Martin, J.R. (1992). *English text: system and structure*. Amsterdam: John Benjamins.

Matthiessen, C. & S.A. Thompson (1988). The structure of discourse and `subordination'. In: J. Haiman & S.A. Thompson (eds.), *Clause combining in grammar and discourse*. Amsterdam/Philadelphia: John Benjamins, 275-329.

McCabe A.E., S. Evely, R. Abramovitch, C.M. Corter & D.J. Pepler (1983). Conditional statements in young children's spontaneous speech. *Journal of Child Language* 10, 253-258.

McCabe, A. & C. Peterson (1997). Meaningful "mistakes": The systematicity of children's connectives in narrative discourse and the social origins of this usage about the past. In: J. Costermans & M. Fayol (eds.), *Processing interclausal relationships: Studies in the production and comprehension of text*. New Jersey: Lawrence Erlbaum Associates, 139-154.

McConnell-Ginet, S. (1982). Adverbs and logical form: A linguistically realistic theory. *Language* 58, 144-187.

Michels, L.C. (1949). Twee gevallen van *dus lang*. *De Nieuwe Taalgids* 42, 212.

Millis, K.K. & M.A. Just (1994). The influence of connectives on sentence comprehension. *Journal of Memory and Language* 33, 128-147.

Nelson, K. (1991). The matter of time: Interdependencies between language and thought in development. In: S.A. Gelman & J.P. Byrnes (eds.), *Perspectives on language and thought: Interrelations in development*. Cambridge: Cambridge University Press, 278-318.

Newport, E.L. (1990). Maturation constraints on language learning. *Cognitive Science* 14, 11-28.

Nishigauchi, T. & T. Roeper (1987). Deductive parameters and the growth of empty categories. In: T. Roeper & E. Williams (eds.), *Parameter setting.* Dordrecht etc.: D. Reidel Publishing Company, 91-121.

Noach, B.M. (1952). Anticiperende gevolgtrekkingen. *Nieuwe Taalgids* 45, 342-343.

Noordman, L. & W. van Rijswijk (1997). De functie van het voegwoord `hoewel' voor de samenhang van tekst [The function of the conjunction 'although' for the connectedness of text]. *Taalbeheersing* 3, 252-264.

Noordman, L.G.M. & W. Vonk (1998). Memory-based processing in understanding causal information. *Discourse Processes* 26, 191-212.

Östmann, J.-O. (1981). ***You know****: A discourse-functional approach*. Amsterdam: Benjamins.

Oversteegen, L.E. (1997). On the pragmatic nature of causal and contrastive connectives. *Discourse Processes* 24/1, 51-86.

Pander Maat, H. (1994). *Tekstanalyse; een pragmatische benadering*. Groningen: Martinus Nijhoff.

Pander Maat, H. (1994a). Coherentierelaties hebben zowel semantische als pragmatische aspecten. In: Maes, A., P. van Hauwermeiren & L. van Waes (red.), *Perspectieven in taalbeheersingsonderzoek*. Dordrecht: ICG Publications, 118-129.

Pander Maat, H. (1998). The classification of negative coherence relations and connectives. *Journal of Pragmatics* 30, 177-204.

Pander Maat, H. & L. Degand (2001). Scaling causal relations and connectives in terms of speaker involvement. *Cognitive Linguistics* 12/3, 211-245.

Pander Maat, H. & T. Sanders (1995). Nederlandse causale connectieven en het onderscheid tussen inhoudelijke en epistemische coherentie-relaties. *Leuvense Bijdragen* 84/3, 349-374.

Pander Maat, H. & T. Sanders (1996). Perspectief in coherentie-relaties en connectieven? Over het gebruik van *dus, daarom* en *daardoor*. *Gramma/TTT* 5-3, 191-207.

Pander Maat, H. & T. Sanders (2000). Domains of use or subjectivity? The distribution of three Dutch causal connectives explained. In: E. Couper-Kühlen & B. Kortmann (eds.), *Cause, condition, concession and contrast: Cognitive and discourse perspectives*. Berlin: Mouton de Gruyter, 57-81.

Pander Maat, H. & T. Sanders (2001). Subjectivity in causal connectives: An empirical study in language use. *Cognitive Linguistics* 12/3, 247-273.

Pasch, R. (1997). Weil mit Hauptsatz – Kuckucksei im Denn-Nest. *Deutsche Sprache* 3, 252-271.

Peterson, C. & A. McCabe (1985). Understanding *because*: How important is the task? *Journal of Psycholinguistic Research* 14, 199-218.

Piaget, J. (1969). *Judgement and reasoning in the child*. London: Routledge & Kegan Paul. [Translation of: J. Piaget (1924). *Le jugement et le raisonnement chez l'enfant*. Neuchatel etc.: Delachaux et Niestlé.]

Pit, M. (2003). *How to express yourself with a causal connective. Subjectivity and causal connectives in Dutch, German and French.* Ph.D. dissertation, Utrecht University. Amsterdam/New York: Rodopi.

Pit, M., H. Pander Maat & T. Sanders (1997). `Doordat', `omdat' en `want'; perspectieven op hun gebruik. *Taalbeheersing* 3, 238-251.

Polanyi, L. (1988). A formal model of the structure of discourse. *Journal of Pragmatics* 12, 601-638.

Polanyi, L. & R. Scha (1983). The syntax of discourse. *Text* 3/3, 261-270.

Prideaux, G.D. (1989). Text data as evidence for language processing principles: The grammar of ordered events. *Language Sciences* 11, 27-42.

Prideaux, G.D. (1993). Subordination and information distribution in oral and written narratives. *Pragmatics and Cognition* 1, 51-69.

Prince, A.S. & P. Smolensky (2004). *Optimality theory: constraint interaction in generative grammar*. Malden, etc.: Blackwell.

Quirk, R., S. Greenbaum, G. Leech & J. Svartvik (1985). *A grammar of contemporary English*. London: Longman.

Radden, G. (1992). The cognitive approach to natural language. In: M. Pütz (ed.), *Thirty years of linguistic evolution*. Amsterdam/Philadelphia: John Benjamins, 513-541.

Ramsay, V. (1987). The functional distribution of preposed and postposed "if" and "when" clauses in written discourse. In: R.S. Tomlin (ed.), *Coherence and Grounding in discourse.* Amsterdam/Philadelphia: John Benjamins, 383-408.

Redeker, G. (1990). Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics* 14, 367-381.

Redeker, G. (1991). Linguistic markers of discourse structure. *Linguistics* 29, 1139-1172.

Renkema, J. (1996). Cohesion analysis and information flow: the case of 'Because' versus 'because'. In: C. Cremers & M. den Dikken (eds.), *Linguistics in the Netherlands* 1996. Amsterdam/Philadelphia: John Benjamins, 233-244.

Rizzi, L. (1997). The fine structure of the left periphery. In: L. Haegeman (ed.), *Elements of grammar*. Dordrecht: Kluwer, 289-330.

Roelofs, M. (1998). *"Hoe bedoel je?" De verwerving van pragmatische vaardigheden*. Ph.D. dissertation University of Amsterdam, Den Haag: Holland Academic Graphics.

Rossari, C. & J. Jayez (1996). *Donc* et les consécutifs des systèmes de contraintes différentiels. *Lingvisticæ Investigationes* XX/1, 117-143.

Rutherford, W.E. (1970). Some observations concerning subordinate clauses in English. *Language* 46/1, 97-115.

Sanders, T.J.M. (1992). *Discourse structure and discourse coherence: Aspects of a cognitive theory of discourse representation*. Ph.D. dissertation, Katholieke Universiteit Brabant, Tilburg.

Sanders, T. (1997). Semantic and pragmatic sources of coherence: On the categorization of coherence relations in context. *Discourse Processes* 24, 119-147.

Sanders, T.J.M. & L.G.M. Noordman (2000). The role of coherence relations and their linguistic markers in text processing. *Discourse Processes* 29/1, 37-60.

Sanders, J. & W. Spooren (1997). Perspective, subjectivity, and modality from a cognitive linguistic point of view. In: W. Liebert, G. Redeker & L. Waugh (eds.), *Discourse and perspective in cognitive linguistics*. Amsterdam/Philadelphia: John Benjamins, 85-112.

Sanders, T. & W. Spooren (1999). Communicative intentions and coherence relations. In: W. Bublitz, U. Lenk & E. Ventola (eds.), *Coherence in spoken and written discourse*, Amsterdam/Philadelphia: John Benjamins, 235-250.

Sanders, T. & W. Spooren (2005). Discourse and text structure. In: D. Geeraerts & H. Cuykens (eds.), *Handbook of Cognitive Linguistics*. Oxford: Oxford University Press (to appear).

Sanders, T.J.M., W.P.M. Spooren & L.G.M. Noordman (1992). Toward a taxonomy of coherence relations. *Discourse Processes* 15, 1-35.

Sanders, T.J.M., W.P.M. Spooren & L.G.M. Noordman (1993). Coherence relations in a cognitive theory of discourse representation. *Cognitive Linguistics* 4/2, 93-133.

Sanders, T. & C. van Wijk (1996). PISA - A procedure for analyzing the structure of explanatory texts. *Text* 16, 91-132.

Schaerlakens, A.M. & S. Gillis (1987). *De taalverwerving van het kind; een hernieuwde oriëntatie in het Nederlandstalige onderzoek*. Groningen: Wolters-Noordhoff.

Schiffrin, D. (1987). *Discourse markers*. Cambridge: Cambridge University Press.

Schiffrin, D. (2001). Discourse markers: Language, meaning, and context. In: D. Schiffrin, D. Tannen & H. Hamilton (eds.), *The Handbook of Discourse Analysis*. Malden, MA etc.: Blackwell, 54-75.

Schilperoord, J. & A. Verhagen (1998). Conceptual dependency and the clausal structure of discourse. In: J.-P. König (ed.), *Discourse and cognition: Bridging the gap*. Stanford, CA: CSLI Publications, 141-163.

Schlichting, J.E.P.T. (1996). *Discovering syntax; An empirical study in Dutch language acquisition.* Ph.D. dissertation Katholieke Universiteit Nijmegen. Nijmegen: Nijmegen University Press.

Schoonenboom, J. (2000). *Analyse, norm en gebruik als factoren van taalverandering: een studie naar veranderingen in het Nederlands onzijdig relativum.* Ph.D. dissertation Universiteit van Amsterdam.

Shapiro, L.R. & J.A. Hudson (1997). Coherence and cohesion in children's stories. In: J. Costermans & M. Fayol (eds.), *Processing interclausal relationships: Studies in the production and comprehension of text*. New Jersey: Lawrence Erlbaum Associates, 23-48.

Slobin, D.I. (1973). Cognitive prerequisites for the development of grammar. In: C.A. Ferguson & D.I. Slobin (eds.), *Studies of child language development*, New York: Holt, Rinehart & Winston, 175-208.

Slobin, D.I. (1977). Language change in childhood and in history. In: J. Macnamara (ed.), *Language learning and thought*. New York: Academic Press, 185-214.

Slobin, D.I. (1994). Talking perfectly; Discourse origins of the present perfect. In: W. Pagliuca (ed.), *Perspectives on grammaticalization.* Amsterdam/Philadelpia: John Benjamins, 119-133.

Smessaert, H. & J. Beeken (1995). The syntax of Dutch connectives: some preliminary observations. *Leuvense Bijdragen* 84/3, 375-396.

Spooren, W. (1997). The processing of underspecified coherence relations. *Discourse Processes* 24, 149-168.

Spooren, W. & T. Sanders (2005). The acquisition of coherence relations: On cognitive complexity in discourse (submitted for publication).

Spooren, W., T. Sanders & J. Visser (1994). Taxonomie van coherentierelaties: Evidentie uit taalverwervingsonderszoek. *Gramma/TTT* 3/1, 33-54.

Spooren, W., H. Tates & T. Sanders (1996). Taalverwerving en de classificatie van coherentierelaties [Language acquisition and the classification of coherence relations]. *Nederlandse Taalkunde* 1, 26-52.

Stoett, F.A. (1977). *Middelnederlandsche spraakkunst; syntaxis*. Third, revised edition. 's-Gravenhage: Nijhoff.

Stukker, N. (2005). *Causality marking across levels of language structure: A cognitive linguistic approach*. Ph.D. dissertation, Utrecht University.

Stukker, N., T. Sanders & A. Verhagen (1999). Waar een wil is, is geen wet. De categorisering van causale relaties binnen en tussen zinnen. *Gramma/TTT* 1, 66-86.

Sweetser, E.E. (1990). *From etymology to pragmatics. Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.

Tabor, W. & E.C. Traugott (1998). Structural scope expansion and grammaticalization. In: A. Giacalone Ramat & P. Hopper (eds.), *The limits of grammaticalization*. Amsterdam/Philadelphia: John Benjamins, 229-272.

Thompson, S.A. (1985). Grammar and written discourse: Initial vs. final purpose clauses in English. *Text* 5, 55-84.

Thompson, S.A. & R.E. Longacre (1985). Adverbial clauses. In: T. Shopen (ed.), *Language typology and syntactic description. Volume II*. Cambridge: Cambridge University Press, 171-234.

Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics* 11-1/2, 61-82.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard: Harvard University Press.

Traugott, E.C. (1995). Subjectification in grammaticalisation. In: D. Stein & S. Wright (eds.), *Subjectivity and subjectivisation: linguistic perspectives*. Cambridge: Cambridge University Press, 31-54.

Traugott, E.C. (1995a). *The role of the development of discourse markers in a theory of grammaticalization*. Paper presented at ICHL XII, Manchester, 1995. Available: <http://www.stanford.edu/~traugott/ect-papersonline.html>.

Traugott, E.C. (2001). *Legitimate counterexamples to unidirectionality*. Paper presented at Freiburg University, October 17th 2001. Available: <http://www.stanford.edu/~traugott/ect-papersonline.html>.

Traugott, E.C. & R.B. Dasher (2002). *Regularity in semantic change*. Cambridge: Cambridge University Press.

Traugott, E.C. & E. König (1991). The semantics-pragmatics of grammaticalization revisited. In: E.C. Traugott & B. Heine (eds.), *Approaches to grammaticalization*, vol. I. Amsterdam/Philadelphia: Benjamins, 189-218.

Uhmann, S. (1998). Verbstellungsvarianten in *weil*-Sätzen: Lexicalische Differenzierung mit grammatischen Folgen. *Zeitschrift für Sprachwissenschaft* 17/1, 92-139.

Uit den Boogaart, P.C. (ed.) (1975). *Woordfrequenties in geschreven en gesproken Nederlands*. Utrecht: Oosthoek, Scheltemat & Holkema.

Van Belle, W. (1989). Want, omdat en aangezien; een argumentatieve analyse. *Leuvense Bijdragen* 78, 435-456.

Van Belle, W. (1997). Conditionele zinnen: een continuüm van een dialogisch over een argumentatief naar een logisch verband. *Taalbeheersing* 3, 215-224.

Van den Bergh, H., W. Herrlitz & L. Klein Gunnewiek (1999). Cognitive co-ordination in foreign language learning: A sketch of a research paradigm. In: J. Don & T. Sanders (eds.), *Utrecht Institute of Linguistics OTS Yearbook 1998-1999*. Utrecht: Utrecht Institute of Linguistics OTS, 20-33.

Van den Bergh, H. & J.B. Hoeksma (1993). Modelling development in education: a three-level model. In: J.H.L. Oud & R.A.W. van Blokland-Vogelesang (eds.), *Advances in longitudinal and multivariate analysis in the behavioral sciences*. Nijmegen: ITS (Instituut voor Toegepaste Sociale Wetenschappen), 81-90.

Van den Bergh, H. & G. Rijlaarsdam (1996). The dynamics of composing: Modeling writing process data. In: C.M. Levy & S. Randall (eds.), *The science of writing: Theories,*

methods and individual differences and applications. Mahwah, New Jersey: Lawrence Erlbaum Associates, 207-232.

Van den Hoven, P. (1997). Niet wij doen het, maar het recht! *Tijdschrift voor taalbeheersing* 9/3, 207-214.

Van der Heijden, E.M.R. (1999). *Tussen nevenschikking en onderschikking: een onderzoek naar verschillende vormen van verbinding in het Nederlands*. Ph.D. dissertation Katholieke Universiteit Nijmegen. Den Haag: Holland Academic Graphics.

Van der Wal, M.J. (1986). *Passiefproblemen in oudere taalfasen. Middelnederlands sijn/werden + participium praeteriti en de pendanten in het Gothisch, het Engels en het Duits*. Dordrecht: ICG Printing.

Van Dijk, T. (1979). Pragmatic connectives. *Journal of Pragmatics* 3, 447-456.

Van Es, G.A. (1954). Voegwoordelijke verbindingen voor de aspectische functies der simultaniteit in het Middelnederlands. *Tijdschrift voor Nederlandse Taal- en Letterkunde* 72, 241-284.

Van Es, G.A. (1955). Voegwoordelijke verbindingen met *doe* en *als* ter uitdrukking van de aspectische functie der progressiviteit in het Middelnederlands. *Tijdschrift voor Nederlandse Taal- en Letterkunde* 73, 16-67.

Van Gestel, F., J. Nijen Twilhaar, T. Rinkel & F. Weerman (1992). *Oude zinnen; grammaticale analyse van het Nederlands*. Leiden/Antwerpen: Martinus Nijhoff.

Van Hell, J.G., L. Verhoeven & L. Wengelin (1999). *Narrative and L1 acquisition: Coordinating and subordinating conjunctions*. Paper presented at the 12th World congres of Applied Linguistics, AILA '99, August 1-6, 1999, Tokyo, Japan.

Van Kampen, J. (1997). *First steps in wh-movement*. Ph.D. dissertation Utrecht University.

Van Kampen, J. (2001). Bootstraps at two for lexicon and discourse. In: *Proceedings of ELA (Early Lexicon Acquisition)*, December 2001, Lyon. Available: <http://www.let.uu.nl/~Jacqueline.vanKampen/personal/downloadables/Bootstrapping-ELA.pdf>.

Van Kemenade, A. (1999). Functional categories, morphosyntactic change, grammaticalization. *Linguistics* 37/6, 997-1010.

Van Megen, N. (2002). *Dan* en *want*: hun functie en betekenis in zeventiende-eeuws taalgebruik. *Neerlandistiek.nl 02.02*, 8 april 2002. URL: <http://www.neerlandistiek.nl/02/02/index.html>.

Van Middendorp, M. & A. van Maaren (2001). *"Ik kies nog steeds voor die, omdat ´ie mij het best voetballen kan leren"; Onderzoek naar het gebruik van coherentierelaties door jonge kinderen.* Unpublished manuscript, Utrecht University.

Verhagen, A. (2000). "The girl that promised to become something": An exploration into diachronic subjectification in Dutch. In: T.F. Shannon & J.P. Snapper (eds.), *The Berkeley Conference on Dutch Linguistics 1997: the Dutch Language at the Millennium*. Lanham, MD: University Press of America, 197-208.

Verhagen, A. (2001). Terug naar *want* en *omdat*. In: B. Dongelmans, J. Lalleman & O. Praamstra (eds.), *Kerven in een rots. Opstellen over Nederlandse taalkunde, letterkunde en cultuur, aangeboden aan Jan W. de Vries bij zijn afscheid als hoogleraar Dutch Studies aan de Universiteit Leiden*. Leiden: Stichting Neerlandistiek Leiden, 107-119.

Verhagen, A. (2005). *Constructions of Intersubjectivity. Discourse, Syntax and Cognition.* Oxford: Oxford University Press.

Verstraete, J.-C. (1998). A semiotic model for the description of levels in conjunction: external, internal-modal and internal-speech functional. *Functions of Language* 5/2, 179-211.

Verstraete, J.-C. (2000). Interpersonal grammar and the typology of clause combining mechanisms in English. *Preprints of the Department of Linguistics* 171. Leuven: Katholieke Universiteit Leuven.

Verwijs, E. & J. Verdam (1885-1952). *Middelnederlandsch Woordenboek (MNW)* 'Middle Dutch dictionary'. ´s-Gravenhage: Nijhoff. Also available on CD-rom: *CD-Rom Middelnederlands: Woordenboek en teksten* (1998). Den Haag/Antwerpen: Sdu.

Vismans, R. (1994). *Modal particles in Dutch directives: a study in functional grammar*. Amsterdam: IFOTT. Ph.D. dissertation Free University of Amsterdam.

Wagner, L. (1998). *The semantics and acquisition of time in language*. Ph.D. dissertation University of Pennsylvania.

Weerman, F. (1988). Moet kunnen: Middelnederlandse zinnen zonder subject. *De Nieuwe Taalgids* 81-4, 289-310.

Weerman. F. (1989). *The V2 conspiracy: A synchronic and a diachronic analysis of verbal positions in Germanic languages*. Dordrecht: Foris.

Wegener, H. (1993). *Weil* – das hat shon seinen Grund. Zur Verbstellung in Kausalsätzen mit WEIL im gegenwärtigen Deutsch. *Deutsche Sprache* 4, 289-305.

Werth, R.N. (1970). The problem of a germanic sentence prototype. *Lingua* 26, 25-34.

Wijnen, F. (1997). Functionele categorieën in Nederlandse kindertaal. *Nederlandse taalkunde* 2/3, 178-198.

Wijnen. F. & L. Elbers (1993). Effort, production skill, and language learning. In: C. Ferguson, L. Menn & C. Stoel-Gammon (eds.), *Phonological development*. Timonium, MD: York.

Wijnen, F. & M. Verrips (1998). The acquisition of Dutch syntax. In: S. Gillis & A. de Houwer (eds.), *The acquisition of Dutch*. Amsterdam/Baltimore: Benjamins.

Wilson-Birnie, J. (2002). *"…daarom dat ik dat zeg"; De rol van complexiteit van connectieven in de verwervingsvolgorde van connectieven bij jonge kinderen*. Unpublished MA-thesis, Utrecht University.

Wing, C.S. & E.K. Scholnick (1981). Children's comprehension of pragmatic concepts expressed in 'because', 'although', 'if' and 'unless'. *Journal of Child Language* 8, 347-365.

Zaalberg, C.A. (1973). *Dus* als onderschikkend voegwoord. *De Nieuwe Taalgids* 66-2, 146-147.

Zobl, H. & J. Liceras (1994). Functional categories and acquisition orders. *Language Learning* 44, 159-180.

Samenvatting

Summary in Dutch

Deze dissertatie richt zich op Nederlandse connectieven, woorden zoals *want* en *omdat* die de relatie tussen zinnen expliciet maken. In dit onderzoek staat de vraag in (1) centraal.

(1) Hoofdvraag van dit proefschrift:
Wat is de relatie tussen de tekstlinguïstische en de zinslinguïstische eigenschappen van connectieven?

Het antwoord op deze hoofdvraag kan zicht bieden op eventuele vorm-functierelaties in het gebruik van Nederlandse connectieven.

Ik heb de relatie tussen tekstlinguïstische/conceptuele eigenschappen en zinslinguïstische/syntactische eigenschappen bestudeerd vanuit een ontwikkelingsperspectief. De reden voor deze keuze is als volgt: als er een interactie bestaat tussen conceptuele eigenschap A en syntactische eigenschap B, en als conceptuele eigenschap A een bepaalde ontwikkeling vertoont, dan is het aannemelijk dat er ook iets zal veranderen met syntactische eigenschap B. Concreet heb ik gekeken naar de vier typen connectiefontwikkeling in Tabel 1.

Tabel 1. Vier typen connectiefontwikkeling die in deze dissertatie bestudeerd worden

	Connectiefverandering in de loop der eeuwen	**Connectiefverwerving door kinderen**
Conceptueel	verandering in betekenis/functie	acquisitie van een betekenis/functie
Syntactisch	verandering in vorm/syntaxis	acquisitie van een vorm/betekenis

In hoofdstuk 2 heb ik ‘connectieven’ gedefinieerd als de talige elementen die lezers en luisteraars instrueren hoe ze een nieuw tekstsegment (minimaal een clause of deelzin) op lokaal niveau met de vorige moeten verbinden. Mijn onderzoek richt zich op de connectieven die in Tabel 2 genoemd zijn.

Tabel 2. Bestudeerde connectieven

Onderzoeksgebied	**Connectief**	**Conceptuele classificatie**	**Syntactische classificatie**
Taalverandering en taalverwerving	daarom	causaal	bijwoord
	dus	causaal	bijwoord, nevenschikkend voegwoord
	omdat	causaal	onderschikkend voegwoord
	want	causaal	nevenschikkend voegwoord
Taalverwerving	en	additief	nevenschikkend voegwoord
	maar	contrastief	nevenschikkend voegwoord
	toen	temporeel	bijwoord, onderschikkend voegwoord

In hoofdstuk 2 bespreek ik vier conceptuele en vier syntactische criteria waarmee connectieven geclassificeerd kunnen worden. In hoofdstuk 3 bespreek ik bestaande literatuur over vorm-functierelaties, waarmee ik laat zien dat het idee van een interactie tussen vorm en functie van connectieven niet onaannemelijk is.

Deel II van dit boek – hoofdstuk 4 tot en met 8 – richt zich op de historische ontwikkeling van vier Nederlandse causale connectieven. In dit deel volg ik een top-downbenadering: de bespreking in ieder hoofdstuk staat direct in relatie tot de hoofdvraag van dit proefschrift, aangezien er steeds een bepaalde hypothese over een specifieke vorm-functierelatie wordt getoetst. Deel III (hoofdstuk 9-13) gaat in op de acquisitie van connectieven. In dit deel heb ik gekozen voor een bottom-up-benadering, aangezien er tot nu toe relatief weinig bekend is over de verwerving van Nederlandse connectieven. In hoofdstuk 10 tot en met 13 ga ik dan ook uitgebreid in op verschillende aspecten van het acquisitieproces en probeer ik daar een verklaring voor te geven in termen van cumulatieve complexiteit. Pas in hoofdstuk 13 wordt een duidelijke link gelegd met de hoofdvraag van deze dissertatie. Hieronder bespreek ik eerst de belangrijkste resultaten in relatie tot de interactiehypotheses. Daarna ga ik in op de hoogtepunten uit diachronie en acquisitie.

Deel II begint in hoofdstuk 4 met een bespreking van de methodologie van de diachrone studies. Voor de connectieven *want*, *omdat*, *dus* en *daarom* zijn steeds steekproeven van 150 fragmenten samengesteld. Deze fragmenten zijn als volgt verdeeld over eeuwen en tekstsoorten (zie Tabel 3).

Tabel 3. Aantal en soort historische fragmenten per connectief

Periode	**aantal rijmende en/of literaire fragmenten**	**aantal niet-rijmende en/of niet-literaire fragmenten**	**Totaal**
13^{e} eeuw	25	25	50
16^{e} eeuw	25	25	50
20^{e} eeuw	25	25	50
Totaal	75	75	150

In hoofdstuk 5 toets ik een hypothese over de interactie tussen de woordvolgorde binnen de connectiefzin – in het bijzonder de plaatsing van het finiete werkwoord – en de hiërarchische structuur van tekstsegmenten. Verhagen (2001) beweert namelijk dat de woordvolgorde van connectiefzinnen in combinaties van drie zinnen [A-B-C] aangeeft of de C-zin alleen aan de B-zin gehecht moet worden of aan de combinatie A-B. Volgens hem dwingen connectiefzinnen met het finiete werkwoord achteraan (V-laat) aanhechting aan de B-zin af (resulterend in de tekststructuur [A-[B-C]]). Bij connectiefzinnen met het finiete werkwoord op de tweede plek (V2) is ook aanhechting aan de combinatie A-B mogelijk ([[A-B]-C]). De diachrone studie van *want* en *omdat* in hoofdstuk 5 laat zien dat Verhagens hypothese op het juiste spoor zit, maar dat er een modificatie nodig is. De interactie moet beperkt worden tot achteropgeplaatste connectiefzinnen, aangezien vooropgeplaatste zinnen met V-laat beide aanhechtingsmogelijkheden hebben. Zo is de vooropgeplaatste *omdat*-zin in (2) niet alleen gerelateerd aan de B-zin, maar aan de B-A-zinscombinatie als geheel.

(2) *Mar* [**C** ***v^{o}mb dat*** *hi* (= de leeuw) *naturlik ku^{o}nheit an hu^{o}me heuet.*] [**B** *so scaimt hi v^{o}me angst te hebene.*] [**A** *Jnde lopt den man v^{o}p. als hastelik als di^{e} man dru^{o}p siit.*]
(NM, 1270-1290)
'Maar omdat hij een natuurlijke moed heeft, schaamt hij zich ervoor om angst te hebben en valt hij de man aan zodra deze naar hem kijkt.'

De kindertaaldata waren niet geschikt om aanvullende evidentie rondom deze hypothese te bieden. In hoofdstuk 12 bleek namelijk dat kinderen óf nog geen combinaties van drie zinnen

produceerden, óf nog niet de woordvolgorde in hun causale relaties konden variëren omdat ze alleen het V2-connectief *want* hadden verworven en nog niet *omdat*.

In hoofdstuk 6 toets ik een hypothese over de interactie tussen de woordvolgorde binnen de connectiefzin en de interpretatie in termen van drie zogenoemde 'domeinen' (cf. Sweetser 1990): content, epistemisch en speech act. De content-relatie in (3) beschrijft een causale relatie tussen twee gebeurtenissen in de werkelijkheid. In de epistemische relatie in (4) is geen sprake van een oorzaak-gevolgrelatie, maar van een conclusie-argumentrelatie. Fragment (5) illustreert een relatie in het speech-act domein. De zin *de zon schijnt* geeft een motivatie voor de taalhandeling: het voorstel doen om in de tuin te gaan eten.

(3) We gingen in de tuin zitten omdat de zon scheen.
(4) De temperatuur zal waarschijnlijk stijgen, want de zon schijnt.
(5) Laten we in de tuin gaan eten, want de zon schijnt.

In hoofdstuk 6 betoog ik dat woordvolgordes onderdeel uitmaken van bepaalde syntactische constructies, die op hun beurt weer variatie vertonen in hun mogelijke domeininterpretaties. Uit de diachrone analyse van *want* en *omdat* bleek dat alleen V-laat connectieven in de "bound subordination construction" beperkt waren tot interpretaties in het content-domein, terwijl V2-zinnen in "free coordination" en V-laat-zinnen in andere constructies niet op een dergelijke manier beperkt waren. Eventuele verdere beperkingen in de mogelijke domein-interpretaties lijken per connectief te verschillen. De acquisitiedata waren opnieuw niet geschikt om de geformuleerde hypothese verder te toetsen. Het leek erop dat de verwerving van *omdat* het domeinengebruik van het eerder verworven *want* nauwelijks beïnvloedde: *omdat* bleef infrequent en *want* werd in alle drie domeinen gebruikt.

In hoofdstuk 7 bestudeer ik de interactie tussen de betekenis en de positie van connectieven aan de hand van een diachrone analyse van *dus* en *daarom*. Hieruit, en uit de acquisitiedata in hoofdstuk 13, blijkt dat de positie van deze woorden niet samenhangt met hun interpretatie in termen van domeinen. Wel is er evidentie dat positionering gebruikt wordt om de connectieffunctie apart te zetten van eventuele andere functies die een woord kan vervullen. De diachrone data toonden aan dat de positionering van *dus* en *daarom* gebruikt wordt om de connectieffunctie en de anaforische functie van elkaar te onderscheiden: in de dertiende en de zestiende eeuw had de connectieffunctie een voorkeur voor de zinsinitiële positie (zoals in (6)), terwijl de niet-causale anaforische functie een voorkeur had voor de een zinsinterne positie (zoals in (7)).

(6) *En nu die sommige out en cout / sy hebben niet waer mee*
dat sij haer hongerige buyck / sullen versaeden
***dus** toont aan haer u lieft* (Spel van sinnen, 1597)
'En nu, sommige ouden en kouden hebben niets waarmee zij hun hongerige buik kunnen verzadigen. Dus betoon hen uw liefde.'
(7) *Als die Fransoysen die coninc **dus** hoirden roepen (…)*
(Historie van den vier heemskinderen, 1508)
'Toen de Fransen de koning zo hoorden roepen (…)'

In de twintigste eeuw kan de positionering van *dus* gerelateerd worden aan de toegankelijkheid van de informatie in de *dus*-zin. Wanneer *dus* een "bekende" conclusie introduceert (zoals in (8)), heeft het een voorkeur voor een zinsinterne positie, net als wanneer *dus* de functie van discourse marker vervult (zoals in (9)). In dat laatste geval markeert *dus*

geen causale relatie meer, maar geeft het slechts aan dat de informatie in de bijbehorende zin bekend verondersteld wordt. Bij niet-bekende conclusies zoals in (10) heeft *dus* een voorkeur voor de zinsinitiële positie.

(8) [Onderwerp is het aantal zondagen in een jaar; in 1995 waren er 53 in plaats van 52.]
Als ik het goed heb uitgerekend, herhaalt zich deze situatie in het veelbesproken jaar 2000, want dan valt nieuwjaarsdag op een zaterdag en de eerste zondag ***dus*** *op 2 januari.* (MC, 1995)

(9) [Dit is een verslag over de opening van een zwembad. De informatie dat de opening door Hilde Zoer is verricht, is ook al eerder in de tekst genoemd.]
Op het moment dat de eerste burger zich openlijk afvroeg wie de openingshandeling wilde verrichten, ging een groot aantal vingers de lucht in. Het werd ***dus*** *Hilde Zoer.* (Meppeler Courant, 1995)

(10) *Patrick woonde in een zijstraat van de laan van Nieuw Guinea,* ***dus*** *Hendriks* <u>*kwam*</u> *er vlak langs.* (De kunstrijder, 1989)

De acquisitiedata in hoofdstuk 13 bieden additionele ondersteuning voor deze interactie: Nederlandse kinderen vertonen van jongs af aan een voorkeur voor discourse marker *dus* in de zinsinterne positie, terwijl ze het connectief *dus* meestal in de zinsinitiële positie plaatsen.

In hoofdstuk 8 bestudeer ik de reikwijdte van de subjectificatiehypothese van Traugott (1995). Zij veronderstelt dat er bij eventuele diachrone veranderingen altijd een toename is in de mate van subjectiviteit; een uitspraak is subjectiever naarmate deze meer gebaseerd is op ideeën van de spreker en minder gebaseerd is op feiten uit de werkelijkheid. In dit hoofdstuk introduceer ik twee methodes om de mate van subjectiviteit van connectieffragmenten vast te stellen. Hierbij maak ik gebruik van Sweetsers classificatie in termen van domeinen en van een classificatie in termen van *Subjects of Conscienceness* (SOC's). De SOC is de persoon die verantwoordelijk is voor het ontstaan van de causale relatie in de werkelijkheid; deze kan op verschillende linguïstische manieren gerealiseerd worden in een connectiefzin (bv. *hij* versus *ik* versus *de man*). Op basis van mijn analyses van de vier causale connectieven in hoofdstuk 8 kan gesteld worden dat subjectificatie zich wel voordoet bij veranderingen van en naar de functie van connectief (bv. van anafoor naar connectief of van connectief naar discourse marker, maar niet bij verschuivingen binnen het connectiefgebruik (zoals de domeinen binnen het gebruik als causaal connectief). In tegendeel, binnen de connectieffunctie kunnen veranderingen zelfs in de tegenovergestelde richting verlopen. Bovendien is subjectificatie niet beperkt tot het voorkomen van syntactische verschuivingen zoals veranderingen in categoriale status.

In Tabel 4 bied ik een overzicht van de verschillende vorm-functierelaties die zijn voorgesteld en van de resultaten die hierover verkregen zijn vanuit de literatuur en/of de studies in dit boek. De meerderheid van deze resultaten suggereert dat de syntactische en tekstlinguïstische eigenschappen van connectieven inderdaad gerelateerd zijn. Verschillende syntactische structuren die even grammaticaal zijn, kunnen gebruikt worden om verschillende functies op het tekstlinguïstische niveau te ondersteunen. Vaak zijn deze vorm-functierelaties echter niet één-op-één, aangezien meestal slechts één van de syntactische opties beperkt is in zijn mogelijke tekstlinguïstische interpretaties, terwijl de alternatieve opties dat niet zijn.

Tabel 4. Overzicht van de claims en resultaten over vorm/functie-interacties

Syntactische eigenschap	**Interactie met conceptuele eigenschappen?**
Liniarisatie van de connectiefzin	Ja: vooropgeplaatste zinnen hebben een "grounding"-functie (cf. Thompson 1985; Ramsay 1987, Degand 2001)
Woordvolgorde in de matrixzin	Ja: in het geval van conditionele zinnen, is de niet-integratieve woordvolgorde alleen mogelijk bij adverbiale zinnen die apart 'beweerd' kunnen worden (cf. König & Van der Auwera 1988; Van Belle 1997)
Woordvolgorde in de connectiefzin	Ja: achteropgeplaatste V-laatzinnen dwingen een interpretatie van "late closure" af (cf. Verhagen 2001 en de resultaten in hoofdstuk 5)
Woordvolgorde in de connectiefzin	Ja: V-laatzinnen in de "bound subordination constructie" dwingen een content-interpretatie af (cf. Verstraete 2000 en de resultaten in hoofdstuk 6)
Positie van het connectief	Ja: positie onderscheidt tussen connectief- en niet-connectieffuncties van een woord (cf. de resultaten in hoofdstuk 7)
Positie van het connectief	Nee: positie onderscheidt niet tussen de domeinen binnen het gebruik als connectief (cf. de resultaten in hoofdstuk 7 en 13)
Positie van het connectief	Ja: positie onderscheidt toegankelijke van niet-toegankelijke informatie (cf. Ariel 1988 en de resultaten in hoofdstuk 7 en 13)
Categoriale status	Nee: subjectificatie is niet beperkt tot veranderingen in categoriale status (hoofdstuk 8)

Het is goed mogelijk dat andere factoren de één-op-één-relaties verstoren. Zo is het denkbaar dat andere vorm-functierelaties interfereren die gebruik maken van dezelfde beperkte set van syntactische gereedschappen (vgl. de optimaliteitsbenadering van Prince & Smolensky 2004).

De analyses van de diachrone ontwikkeling en de verwerving van connectieven hebben een aantal interessante feiten aan het licht gebracht. Zo laat de diachrone analyse van de vier causalen *want*, *omdat*, *dus* en *daarom* zien dat veel van de hedendaagse eigenschappen van deze connectieven er al in de dertiende eeuw waren. Deze relatief stabiele eigenschappen van de vier causale connectieven staan in Tabel 5.

Tabel 5. Stabiele eigenschappen en belangrijkste veranderingen in het gebruik van vier connectieven

	want	**omdat**	**dus**	**daarom**
Conceptuele functie	causaal connectief	causaal connectief	causaal connectief	causaal connectief
Voorkeursdomein	epistemisch	content	epistemisch	content
Woordvolgorde	V2	V-laat	-	-
Categoriale status	voegwoord	voegwoord	bijwoord	bijwoord

Het opmerkelijk constante profiel van de vier woorden kan verklaard worden vanuit een vorm van specialisatie. Het specifieke profiel van een bepaald connectief garandeert zijn bestaansrecht. Drastische veranderingen zouden ertoe kunnen leiden dat er een volledige

overlap ontstaat met het profiel van een ander connectief, wat weer zou kunnen leiden tot het verdwijnen van één van de twee connectieven.

Tabel 6 toont de belangrijkste veranderingen in het gebruik van de vier connectieven.

Tabel 6. Belangrijkste veranderingen in het gebruik van vier connectieven

	want	**omdat**	**dus**	**daarom**
Conceptuele veranderingen	verlies van temporeel connectiefgebruik	verlies van finalistisch (“opdat”) gebruik	- verlies van anaforische functies - nieuw: functie van discourse marker	verlies van anaforische functies
Syntactische veranderingen	verlies van V-laat woordvolgorde	heranalyse van prepositie *om* + voegwoord *dat* tot voegwoord *omdat*	nieuw: functie van voegwoord	-

Deze tabel laat zien dat specialisatie van connectieven inderdaad een rol speelt. Het verdwijnen van V-laat *want* heeft bijvoorbeeld geresulteerd in een betere taakverdeling tussen V2 *want* en V-laat *omdat*: V2 *want* wordt voornamelijk gebruikt om epistemische relaties uit te drukken, terwijl V-laat *omdat* vooral getypeerd wordt door gebruik in het content-domein. Een tweede geval van specialisatie doet zich voor in de divergentie van *aldus* en *dus*. In het Middelnederlands waren deze twee woorden vrij uitwisselbaar. In het modern Nederlands kan alleen *aldus* noch de anaforische functie vervullen, terwijl *dus* alleen gebruikt kan worden als causaal connectief of als discourse marker. Een derde geval van specialisatie heeft zich voorgedaan bij *opdat* en *omdat*. In het Middelnederlands konden beide connectieven een causale relatie met een gewenst, maar nog niet gerealiseerd consequent markeren; in het modern Nederlands heeft *omdat* deze functie verloren. Als gevolg heeft *opdat* nu een specifiek profiel dat duidelijk verschilt van het profiel van *omdat*. Deze resultaten suggereren dat toekomstige studies naar connectiefveranderingen zich niet kunnen veroorloven om alleen naar geïsoleerde connectieven te kijken, maar dat zij ook eventuele linguïstische alternatieven in hun onderzoek moeten betrekken.

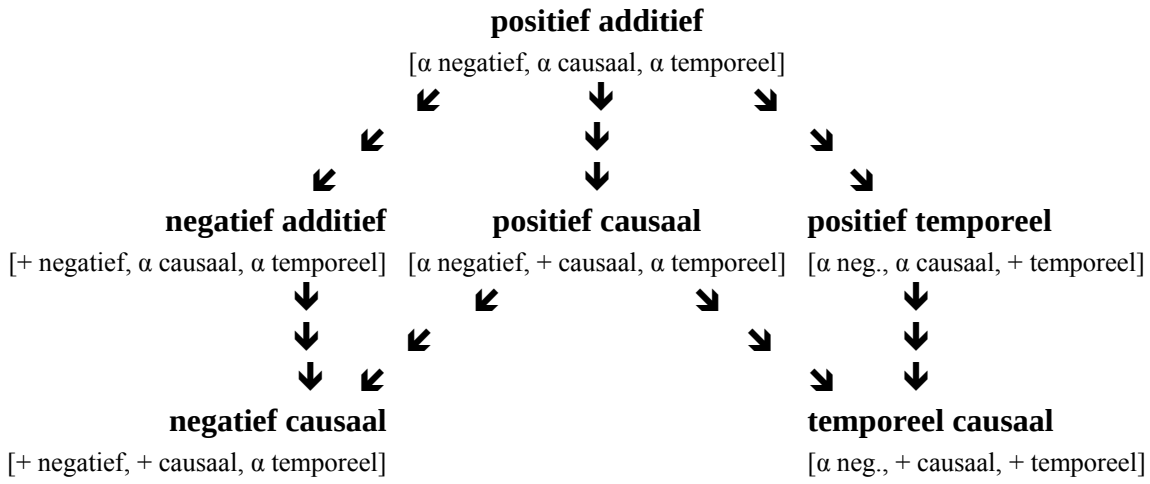
Ook het deel over acquisitie heeft enkele interessante bevindingen opgeleverd. In hoofdstuk 9 geef ik aan dat het belangrijk is om een goede definitie van het begrip ‘acquisitie/verwerving’ te hanteren. Hierbij kom ik tot de conclusie dat een volledig beeld van de connectiefverwerving in longitudinale data alleen verkregen kan woorden a) door te kijken naar het eerste correcte en creatieve gebruik van connectieven (zie hoofdstuk 10), b) door de intra- en interindividuele acquisitiepatronen te bestuderen aan de hand van ontwikkelingscurves (zie hoofdstuk 11), en c) door het connectiefgebruik op verschillende leeftijden op een kwalitatieve manier te bestuderen (zie hoofdstuk 12).

In hoofdstuk 10 tot en met 13 presenteer ik data over de acquisitie van een aantal Nederlandse connectieven. Hierbij suggereer ik dat de acquisitiefeiten verklaard kunnen worden vanuit de notie *cumulatieve complexiteit* (vgl. Brown 1973). Deze complexiteitsbenadering bestaat uit ten minste twee componenten: cumulatieve conceptuele complexiteit en cumulatieve syntactische complexiteit.

In hoofdstuk 10 presenteer ik een multi-dimensionele benadering van cumulatieve conceptuele complexiteit, waarbij ik de verschillende conceptuele noties niet orden op één dimensie (zoals bij Bloom et al. 1980), maar waarbij ieder connectief gedefinieerd wordt op basis van verschillende cognitieve primitieven (overgenomen uit Sanders et al. 1992). Deze

benadering blijkt een goede verklaring te bieden voor de bevindingen over Nederlandse en Engelse connectiefverwerving. Beide talen vertoonden een vast en een variabel deel in de verwervingsvolgorde van connectieven. De vaste routes kunnen verklaard worden vanuit de relatieve complexiteit van waarden binnen een conceptueel primitief (bv. [α causaal] versus [+ causaal]), terwijl de variatie verklaard kan worden vanuit verschillende interacties tussen de conceptuele primitieven die een connectief karakteriseren. Dit ontwikkelingspatroon is schematisch weergegeven in (11).

(11) <u>Acquisitievolgordes gebaseerd op de interactie tussen *basisoperatie*, *polariteit* en *temporele ordening*:</u>



De bevindingen in hoofdstuk 12 indiceren dat ook de syntactische component van cumulatieve complexiteit van belang is. Zo kan de ontwikkeling van zinsinterne nevenschikking naar zinsverbindende nevenschikking verklaard worden op basis van cumulatieve syntactische complexiteit: zinsverbindende nevenschikking is syntactisch gezien meer complex dan zinsinterne nevenschikking omdat het kind met grotere teksteenheden moet omgaan (zinnen in plaats van frases).

Om een verklaring te kunnen bieden voor de verwerving van alle soorten connectieven, moet het multi-dimensionele model nog uitgebreid worden met andere conceptuele en syntactische primitieven. Zo biedt hoofdstuk 13 de eerste resultaten in termen van *domeinen*. Hieruit blijkt dat epistemische relaties het laatst verworven worden en dat het verwervingsbeeld voor content- en speech-actrelaties minder duidelijk is. Een belangrijke conclusie uit dit hoofdstuk is verder dat context een cruciale rol speelt in de productie van de drie domeinen. Het was niet goed mogelijk om stellige conclusies te trekken uit de verwervingsdata, omdat de verzamelaars van deze data niet systematisch gecontroleerd hadden voor verschillende gebruikscontexten.

Het precieze bereik van de benadering van cumulatieve complexiteit moet nog getoetst worden door de invloed van andere factoren nog verder te bestuderen. In hoofdstuk 11 heb ik hiervoor een eerste aanzet gegeven door te kijken naar de rol van ouderlijke input. De kwantitatieve analyses in dit hoofdstuk maken duidelijk dat de rol van ouderlijke input niet volledig uitgesloten kan worden. Wel kan voorzichtig geconcludeerd worden dat ouderlijke input niet als volkomen onafhankelijke factor beschouwd mag worden. Het lijkt erop dat complexiteit de acquisitieroute van kinderen bepaalt en dat ouders principes van ‘audience design’ toepassen: ze passen zich aan de mogelijkheden van hun kinderen aan. Vanuit deze optiek versterken ouders het effect van cumulatieve complexiteit op het acquisitieproces. Het

lijkt dan ook aantrekkelijk om de benadering van cumulatieve complexiteit te combineren met usage-based benaderingen van taalverwerving.

Een volledig beeld van de verwervingsroute kan alleen verkregen worden vanuit data die in verschillende contexten verzameld zijn. Een dergelijk beeld kan verkregen worden door de methodes te hanteren die gepropageerd worden door Tomasello (2003) en zijn collega's.Dit houdt in dat er longitudinale data van interacties tussen kinderen en hun ouders worden bestudeerd. In aanvulling hierop kunnen experimentele studies ingezet worden.

De bevindingen in deze dissertatie suggereren dat verder onderzoek nodig is om meer inzicht te krijgen in de interactie tussen verschillende vorm-functierelaties. Mijn empirisch onderzoek op basis van specifieke operationalisaties laat zien hoe a) zowel kwantitatieve als kwalitatieve analyses, b) zowel diachrone als acquisitiedata, en c) zowel tekstlinguïstische als zinslinguïstische benaderingen kunnen bijdragen aan de ontwikkeling van een theorie over vorm-functierelaties in het gebruik van connectieven en andere tekstlinguïstische elementen.

Dankwoord

Acknowledgements

"Interdependence is a higher value than independence."
(Stephen R. Covey in *The 7 habits of highly effective people*)

Op de kaft van dit boek staat slechts één auteur vermeld. Toch is het schrijven van deze dissertatie absoluut geen solo-actie geweest. Ik ben heel erg blij met en dankbaar voor alle mensen die er op één of andere manier aan hebben bijgedragen. Ik noem hier de belangrijkste.

- Ted Sanders en Fred Weerman. Zonder jullie had dit boek er heel anders uitgezien. In jullie had ik een geweldig stel begeleiders, en ik vind het dan ook heel leuk dat jullie je van dagelijkse begeleiders hebben opgewerkt tot promotoren. Bedankt voor jullie vragen, kritische opmerkingen, tijd, persoonlijke belangstelling, bemoedigende woorden, en – niet te vergeten – jullie humor!
- Huub van den Bergh. Dank je wel dat je me hebt geïntroduceerd in de geheime wereld van GLIM en multi-level analyses. Bedankt voor je geduld bij het uitleggen van statistische details en voor je humorvolle en relativerende opmerkingen. Mocht ik toch nog ergens statistische onzin hebben verkocht, dan ligt dat zeker niet aan jou.
- Ninke Stukker. Bedankt voor al die uren die je in het diachrone onderzoek hebt gestoken, voor onze inhoudelijke discussies en voor alle potten thee die je hebt gezet.
- Marian van Oordt. Dank je wel dat je me enthousiast hebt gemaakt voor het vak *Algemene Taalwetenschap*. Ik vind het erg leuk dat we studenten van de Evangelische Hogeschool nu samen kunnen enthousiasmeren voor "ons" vak.
- Alle mensen die in mijn tijd als "bursaal" een kamer van Trans 10 met mij hebben gedeeld: Nine Elenbaas, Sandra Peters, Maaike Verrips, Taka Hara, Ninke Stukker, Mike Huiskes, Oele Koornwinder en Judith Kamalski. Ik heb genoten van de gezellige praatjes en de serieuze discussies.
- Mijn collega's bij het UiL-OTS en de afdeling Taalbeheersing van het Onderwijsinstituut Nederlands. Ik heb geprofiteerd van jullie commentaar op tussentijdse presentaties en artikelen en ik kijk uit naar onze toekomstige samenwerking!
- Mijn collega's bij de Evangelische Hogeschool. Bij jullie heb ik het vak van docent geleerd en ik vind het een eer om te mogen samenwerken met mensen die zo gedreven zijn en altijd weer mogelijkheden voor verbetering zien. De colleges over praktische vaardigheden zoals vergaderen, presenteren, interviewen en betoog schrijven vormden een aangename afwisseling bij al het theoretische werk op de universiteit.
- De studenten van het *Onderzoekscollege Synchrone en Diachrone Taalontwikkeling*. Het was inspirerend om samen nog meer te ontdekken over taalontwikkeling. Mijn speciale dank gaat uit naar Arvid van Maaren, Marijke van Middendorp en Johanneke Wilson-Birnie, die samen met mij een deel van het onderzoek hebben uitgevoerd.
- Vrienden en familie. Bedankt voor al die keren dat jullie belangstellend vroegen of mijn "scriptie" al af was.
- Mijn "thuisfront". Waar zou ik zijn geweest zonder moeders? Mama Vermeul en Mama Evers, heel erg bedankt! Hoe had het gemoeten zonder extra oppas? Wilma van den Brink, bedankt voor alle broodjes pasta die je voor Nathan en Mirjam hebt gesmeerd!

- Nathan en Mirjam. Ik vond het “vet cool” om voor of na mijn werk samen te voetballen, te schommelen, uit *Jip en Janneke* of *Pluk van de Petteflet* te lezen. Mama’s boek is nu echt af!
- Bob. Je weet wel waarom… Fijn dat jij er wel goed in bent om één ding tegelijk te doen! Enne, het is toch nog gelukt voor 1 november (2005)!

“Moet je nu nog steeds naar school, of weet je alles al?”
Dat was de vraag die mijn schoonvader, Willem Evers, me vaak plagend stelde als ik na mijn universiteitswerk even langskwam. En iedere keer kon ik hem met plezier antwoorden dat er nog veel meer te ontdekken viel en dat ik dus een goede reden had om steeds maar weer naar die universiteit te gaan. Veel mensen zien dit boek als een eindproduct van een aantal jaren werk. Maar in mijn beleving is het eerder een tussentijdse versie. Een boek met onderzoeksterreinen die nog onderbelicht zijn gebleven, met ideeën die nog niet helemaal zijn uitgewerkt, met vragen die nog niet allemaal zijn beantwoord. Kortom: een boek dat niet echt eindigt met een punt of een uitroepteken, maar met een komma of misschien wel een vraagteken. Ik ben dan ook blij dat ik ook de komende tijd nog naar de universiteit mag blijven komen om vragen te stellen. Gelukkig weet ik nog steeds niet alles en hebben we – om met de woorden van meneer Poeder te spreken – nog een eeuwigheid om te onderzoeken hoe alles echt in elkaar zit!

Jacqueline Evers-Vermeul
april 2005

Curriculum Vitae

Jacqueline Evers-Vermeul was born in Purmerend, The Netherlands, on the 9th of December 1973. From 1992 till 1997, she studied both *Algemene Letteren* 'General Arts' and *Algemene Taalwetenschap* 'General Linguistics' at Utrecht University and specialized in *Language and Language Structure*, and in *Communication Studies*. After obtaining a cum laude Master's degree in August 1997, she was employed as a part-time Ph.D. student at the Utrecht Institute of Linguistics OTS. The research carried out at the UiL OTS resulted in this dissertation. From September 1997 on, Jacqueline Evers-Vermeul also worked at the Evangelische Hogeschool in Amersfoort, teaching linguistics and academic skills. Since November 2004, she holds a position as assistant professor of Language use and Discourse studies in the department of Dutch Language and Culture and the UiL OTS at Utrecht University.