The Structure of Discourse
A Corpus-Based Cross-Linguistic Study

Gylling-Jørgensen, Morten

Link to publication in CBS Research Portal

The Structure of Discourse

Copenhagen
Business School
HANDELSHØJSKOLEN

The Structure of Discourse

A Corpus-Based Cross-Linguistic Study

Morten Gylling

PhD Series 38-2013

LIMAC PhD School
Department of International Business Communication

PhD Series 38-2013

The Structure of Discourse

A Corpus-Based Cross-Linguistic Study

Morten Gylling

Supervisor: Iørn Korzen

Ph.D. School LIMAC

Programme in Language and Culture

Copenhagen Business School

Morten Gylling
The Structure of Discourse
A Corpus-Based Cross-Linguistic Study

Table of contents

Foreword

This thesis is the outcome of a number of years of research in contrastive text linguistics, discourse analysis, intercultural rhetoric and translation studies, which started with my master's thesis in 2009 and my concurrent affiliation with the research project the Copenhagen Dependency Treebanks (see Buch-Kromann, Korzen, & Müller, 2009). Parts of the thesis have appeared and been presented elsewhere: at the *Constraints in Discourse IV* in Agay, France (Gylling & Korzen, 2011); at the GSCL2011 pre-conference workshop *Contrastive Linguistics – Translation Studies – Machine Translation – what can we learn from each other?* in Hamburg, Germany (Korzen & Gylling, 2011), and in the subsequent publication in *Translation: Computation, Corpora, Cognition* (Korzen & Gylling, 2012a); at the 19th annual meeting of the Danish Grammar Network (Gylling, 2012a); at the *XII Congresso della Società Internazionale di Linguistica e Filologia Italiana* in Helsinki, Finland (Gylling, 2012b); at the *II International Conference in Communication, Cognition and Media. Political and Economic Discourse* in Braga, Portugal (Gylling, 2012c); at the *COLLOQUE ⱯIA II* in Copenhagen, Denmark (Korzen & Gylling, 2012b); at the *International Conference on Genre- and Register-related Text and Discourse Features in Multilingual Corpora* in Brussels, Belgium (Gylling, 2013); and at the *International Research Conference. Between Romance and Germanic: Language, Text, Cognition and Culture* in Basel, Switzerland (Gylling, forthcoming). All previous papers have been extensively revised and updated for the purpose of this thesis.

Abstract

Effective communication requires texts to be organised into a coherent discourse structure. But languages vary considerably in how they do this, posing a challenge for effective intercultural communication. Instead of relying on our own preferred persuasion style to be the most effective, we need to take into consideration that people with different linguistic and cultural backgrounds do not necessarily employ the same linguistic means in similar communication situations. This is of particular importance in a business context, and a profound understanding of cross-linguistic differences in the organisation of argumentative texts is needed.

In order to address this challenge, this thesis presents a study of structural characteristics in argumentative texts across three different languages. The aim of the study is to examine some of the linguistic means that writers of different languages employ when creating persuasive discourses. The study is based on 150 Danish, English and Italian speeches held by Members of the European Parliament in their native language.

The linguistic means under investigation are conceptualised as belonging to three different structural domains which account for different ways of linking discourse units in a text: a syntactically organised text structure, a rhetorically organised discourse structure and an information packaging organised information structure. The structural domains are defined from a cognitive-functional perspective and juxtaposed into a single analytical framework.

The analyses show that writers across the three languages generally use the same rhetorical relations to build up persuasive discourses. But the analyses also reveal that the Danish, English and Italian writers textualise relations differently. The Danish writers use almost exclusively finite verb forms in coordinate and subordinate structures. The English writers tend to avoid explicating the rhetorical relations between discourse units, and the Italian writers tend to include more units inside the same sentence than the Danish and English writers.

The analyses also suggest that the cross-linguistic differences in textualisation can be correlated with certain persuasive strategies. The Danish writers tend to persuade by analogy, making use of typical features from narratives. The English writers make use of presentational persuasion style, involving themselves in a more personal way than the Danish and Italian writers. And lastly, the Italian writers make use of typical features from quasilogical persuasion style, adopting a formal register and argumentation.

This thesis formulates an analytical framework for a systematic investigation of the structure of discourse across languages, pairing theories and methods from the two parallel disciplines of linguistics and rhetoric in order to gain more insights into effective intercultural communication.

x

Resumé

For at kommunikation kan virke effektivt, skal der være sammenhæng mellem sætningerne i en tekst. Men forskellige sprog strukturer ikke sætningssammenhæng ens, og det kan forringe effektiviteten af interkulturel kommunikation. Frem for at opfatte vores egen foretrukne overbevisningsstrategi som den mest effektive, bør vi forstå at mennesker med forskellig sproglig og kulturel baggrund ikke altid anvender de samme sproglige virkemidler. Dette er ikke mindst vigtigt i en erhvervsmæssig kontekst, og der er behov for en dybere forståelse for hvordan argumentative tekster struktureres på forskellige sprog.

I denne afhandling undersøges hvordan argumentative tekster struktureres på tre forskellige sprog. Formålet er at opnå indsigt i de sproglige virkemidler der anvendes af afsendere med forskellig sproglig og kulturel baggrund. Undersøgelsen er baseret på 150 danske, engelske og italienske taler fra Europa-Parlamentet.

De sproglige virkemidler som undersøges i afhandlingen, betragtes som tilhørende tre forskellige strukturdimensioner, der hver især beskriver forskellige måder at skabe sammenhæng i en tekst på: tekststruktur beskriver syntaktiske relationer, diskursstruktur beskriver retoriske relationer, og informationsstruktur beskriver relationerne mellem informationsenheder. De tre sproglige strukturdimensioner behandles ud fra et kognitivt-funktionelt perspektiv og belyses i én samlet analyseramme.

Analyserne viser at afsendere typisk anvender de samme retoriske relationer til at strukturere argumentative tekster på de tre sprog. Men analyserne viser også at de danske, engelske og italienske afsendere tekstualiserer relationerne forskelligt. De danske afsendere anvender næsten udelukkende finitte verber i side- og underordnede konstruktioner. De engelske afsendere ekspliciterer retoriske relationer mellem sætninger mindre frekvent end de danske og italienske afsendere. Og de italienske afsendere inkluderer flere enheder i samme periode.

Analyserne belyser også at tekstualiseringsmønstrene på de tre sprog kan relateres til forskellige overbevisningsstrategier. De danske afsendere overbeviser ofte gennem en analogisk argumentationsstil, hvor der anvendes typiske træk fra narrativer. De engelske afsendere foretrækker en præsentationel argumentationsstil og fremstår mere personligt involverede end de danske og italienske afsendere. Endelig anvender de italienske afsendere træk fra kvasilogisk argumentationsstil, hvilket gør at register og argumentation fremstår formelt.

Denne afhandling skaber således en analyseramme for en undersøgelse af teksters struktur på tværs af sprog og forener sprogvidenskabelige og retoriske teorier og metoder med det formål at forstå og effektivisere interkulturel kommunikation.

1 Introduction

I first started to gain interest in linguistic structures during my master's thesis and during my time as a student assistant and later research assistant in the Copenhagen Dependency Treebanks project (Buch-Kromann et al., 2009; Buch-Kromann & Korzen, 2010), where I was tasked with the annotation of morphological, syntactic, semantic, anaphoric and discursive features in a parallel corpus of Danish, English, German, Italian and Spanish texts, from 2008 to 2010. The research project was very ambitious in its aim of creating a unified theory for various linguistic levels (from morphology to discourse) in five different languages. The project achieved this to a certain extent, and the methodological and theoretical experience with building up the methodology and theory made me realise the strengths and challenges in using corpora for cross-linguistic studies. While surveying the literature in contrastive linguistics, corpus linguistics, text linguistics and discourse analysis, it became clear to me that only a few scholars had tried to combine the four disciplines (Abelen, Redeker, & Thompson, 1993; da Cunha & Iruskieta, 2010; Ramm & Fabricius-Hansen, 2005; Taboada, Suárez, & Álvarez, 2013), although a number of scholars had recommended that this was done more consistently (Granger, 2003; Hatim, 1997; Knott, 1996, pp. 60–61; Webber & Prasad, 2009, p. 184).

> [T]here are not many studies in languages other than English. Annotated corpora … would be useful in exploring realisation, frequency, and signalling of rhetorical relations. Comparisons across languages and text types – both synchronically and diachronically – would provide insights into language universals and language change. (Taboada & Mann, 2006a, p. 449)

As many robust theories and useful methods have been developed specifically for the study of each of these four disciplines, I recognised that it might be fruitful to respond to this plea by designing a corpus-based study of linguistic structures across three different languages.

Selecting parliamentary discourse as the basis of my data not only had to do with the availability of the corpus employed but also with the comparability of the texts. Philipp Koehn (University of Edinburgh) has compiled the Europarl Corpus (Koehn, 2005), which is a large parallel corpus of all official languages in the European Union. The corpus contains the proceedings of the parliamentary debates of the Union in original, non-translated versions (L1) together with their corresponding official translations (L2) in the EU languages. This means that the texts, in addition to their primary parallel use (L1-L2), can also be used for comparative

purposes (L1-L1), as the corpus contains 'texts of the same genres in the same domains in a range of different languages in the same period' (McEnery & Xiao, 2007, p. 20).

The next step in designing my study related to the choice of theories and methods. The aim was to examine linguistic structures from a product and process point of view. As many higher level linguistic structures such as text and discourse are based on concepts from lower levels such as morphology and syntax, I chose to follow some of the ideas of the schools of Functional Linguistics and Systemic Functional Linguistics. Halliday & Hasan (1976), Halliday (1985) and Martin (1992) propose that dichotomies such as syntactic coordination versus subordination and finite versus non-finite realisation of clauses belong to the basic structural components of syntax and text structure. Next, I decided to describe discourse structure from a linguistic perspective. For this purpose, I adopt Mann & Thompson's (1988) Rhetorical Structure Theory and their approach to describing discourse structure as the connection of text constituents through interpretable rhetorical relations between these constituents. The text constituents under investigation here shall henceforth be referred to as discourse units, essentially considered to be propositions textualised as sentences, clauses and, in some cases, phrases without an explicit subject or verbal element. In addition to the investigation of syntactic relations between discourse units, I focus on the semantic and pragmatic source of coherence, and the rhetorical hierarchy of discourse units. Finally, I also examine the linkage of discourse units by drawing on the insights of Chafe (1976), Fabricius-Hansen (1996), and Vallduví & Engdahl (1996)'s notions of information structure, and investigate whether the rhetorical relations are signalled or not by discourse cues.

Many cross-linguistic studies have a tendency to complete the investigation by pinpointing differences and similarities between two languages without discussing the source of these. In this thesis, I shall try to go one step further by considering patterns in the three linguistic structures as instances of a particular contextually embedded persuasion style. I do this by interpreting variations in linguistic structures across the three languages under investigation from an intercultural rhetoric perspective (Connor, 2002, 2004, 2011). In intercultural rhetoric, structural variation is correlated with contextual factors such as social and cultural values. As instances of a particular contextually embedded writing style, the discourses under investigation can be described in terms of the structural characteristics they include. Those characteristics are, in turn, described in terms of the effect they may have on the argumentation.

The six main objects of analysis are conceptualised as a combined notion of linguistic structures, understood as rule-bound combinations of linguistic components applied by language

users. Previously, Brown & Yule (1983, p. 24) and Widdowson (1979) have divided the three concepts of text, discourse and information into two distinct approaches: a static and a dynamic one. The static approach is concerned with text as a product and does not take into consideration how the text is received by the reader. The dynamic approach looks at discourse and information as a process and examines how the writer attempts to communicate a message to the reader in his online reading process. Instead of considering text on one side and discourse and information on the other as two distinct approaches for studying discourse units within and above the sentence, I combine the two approaches in order to examine in which ways the use of discourse units and rhetorical relations differs in different languages. The six main objects of analysis are:

- *Dependency* (syntactic coordination versus subordination of discourse units)
- *Realisation* (finite versus non-finite realisation of discourse units)
- *Hierarchy* (rhetorical coordination versus subordination of discourse units)
- *Source of coherence* (semantic versus pragmatic description of rhetorical relations)
- *Linkage* (intersentential versus intrasentential juxtaposition of discourse units)
- *Signal* (implicit versus explicit marking of rhetorical relations)

Discourse units that occur inside the same sentence can be either syntactically coordinated or subordinated. This is referred to as *dependency*. Furthermore, discourse units can be realised by finite or non-finite verb forms, or as verbless constructions. I term this *realisation*. Just as discourse units can be syntactically coordinated or subordinated, discourse units can also be rhetorically coordinated or subordinated, depending on their co-textual salience. This is labelled *hierarchy*. Discourse units can be related rhetorically to other discourse units in terms of semantic (information) or pragmatic (intentions) relations. This is called *source of coherence*. Discourse units can be related within the same sentence or across sentence boundaries. I term this intrasentential or intersentential *linkage*. Rhetorical relations between discourse units can be marked explicitly by a number of linguistic items, referred to as discourse cues, or they can remain implicit. I name this phenomenon *signal*.

The aspects chosen come together as characteristics of text structure, discourse structure and information structure, as shown in Figure 1.1. The figure shows how the objects of analysis are divided into three distinct linguistic structures. In the first, two objects are considered. Dependency and realisation of discourse units yield different aspects of text structure. With respect to dependency, the relative frequency of coordinate versus subordinate discourse units is

studied. As far as realisation is concerned, the relative frequency of finite and non-finite verb forms together with verbless constructions is examined. The second linguistic structure is discourse structure, which is realised through rhetorical relations that enter in a rhetorical hierarchy, coordinate or subordinate, which then can be ascribed a source of coherence, either semantic or pragmatic. The final two objects of analysis are related to both text and discourse structure, as they focus on discourse units and rhetorical relations simultaneously, although they have been placed in a separate box labelled information structure. The objects are intrasentential versus intersentential linkage of discourse units and the signalling of the rhetorical relations between these.
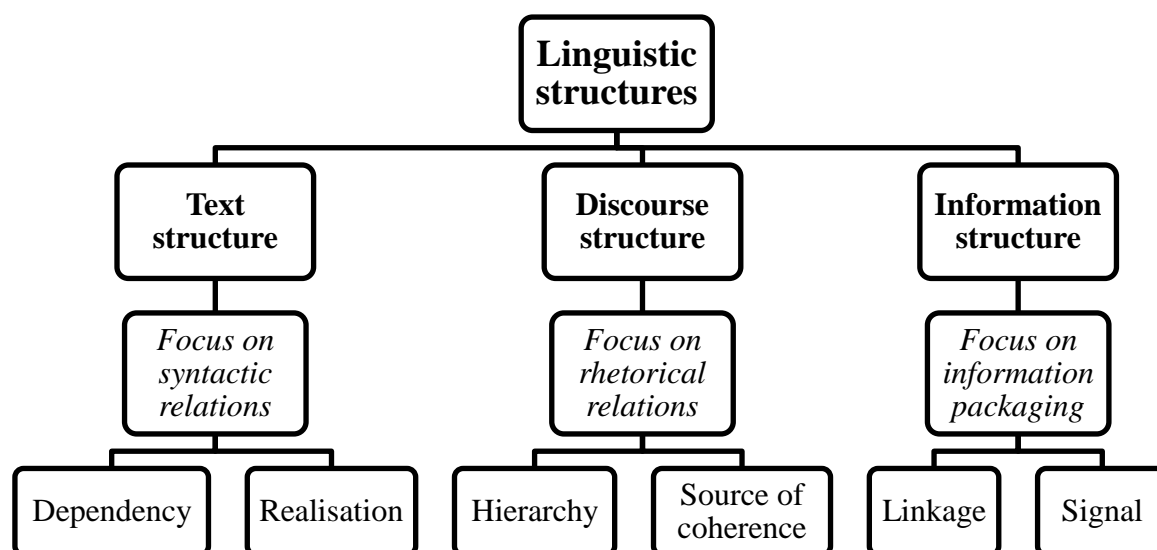
Figure 1.1: Linguistic structures, foci and objects of analysis

Apart from accounting for the theoretical framework of the present study, Figure 1.1 also outlines the structure of the thesis: each linguistic structure has been assigned its own chapter which focuses on various aspects of the given structures by investigating the objects of analysis shown in the bottom-level boxes of the figure. The top level, linguistic structures, is introduced in the present chapter and in Chapter 2, and discussed in more detail in Chapters 7 and 8.

Admittedly, the sharp division of linguistic structures shown in Figure 1.1 is more operational than the definitions and foci of the individual structures actually permit. In this way, it could be argued that information structure together with its two objects of analysis should be considered a part of text structure and discourse structure rather than a distinct structure. It could also be argued that the two objects of analysis dependency and hierarchy are so closely related

to each other that it does not make sense to separate these from each other under different structural characteristics. This is true, but similar distinctions have been made by other scholars in the literature (see the relevant chapters for references), and within this field of linguistics, it is very difficult to define categories that do not have overlaps. As it is, the distinction turned out to be effective in the annotation process and in the subsequent analyses.

The theories and methods employed in the thesis stem from the disciplines of text linguistics, discourse analysis and intercultural rhetoric, all taken from the field of Functional Linguistics. Both the theories and the methods have mainly been developed for the English language. Nevertheless, most have also been tested on other languages. It was necessary to perform some minor revisions of the theories and methods in order to carry out rigorous comparisons between the texts in the three languages, but my results show that English constitutes a useful point of departure from a linguistic perspective, although some apparently similar grammatical constructions differ in terms of semantics.

I have already hinted at some of the research questions. In the rest of this introductory chapter, I shall outline those questions and discuss them in more detail. Below, the basic research questions that I set out to answer in this study have been formulated.

1. How are various linguistic structures within and above the sentence level related to each other?
2. How do linguistic structures manifest themselves in argumentative discourse?
3. How do linguistic structures vary in argumentative discourse across Danish, English and Italian?
4. To which extent can linguistic structures be said to be influenced by contextual factors?

The approach to the answer to the first question has already been outlined: through the analysis of the six above-mentioned objects, I expand present views on linguistic structures. The second and third question will be answered throughout the thesis. In each chapter, I point out the differences across the three languages under investigation, focusing in particular on the textualisation of discourse units and the rhetorical relations between these. At the same time, these differences shall also be studied with respect to the typological nature of the texts and of the three languages. Finally, the fourth question brings us to an explanation of the variations in linguistic structures found in Danish, English and Italian parliamentary discourse.

The main body of the work starts in Chapter 2, with a survey of the theories of linguistic structures and the outline of an analytical framework of these. I then provide a thorough description of the corpus, the data collection and the annotation process. This description is to be found in Chapter 3.

Chapter 4 deals with the description of text structure phenomena, followed by a discussion of dependencies and realisations of discourse units for all three languages, by taking into consideration their language-specific features. Next, Chapter 5 examines the elements of discourse structure in the corpus through a study of the hierarchy of discourse units and sources of coherences of the rhetorical relations, again contrasting the three languages. The survey of linguistic structures is completed in Chapter 6, which describes how discourse units and rhetorical relations are linked and signalled in the texts of the corpus. Chapter 7 takes a different route, which leads to a discussion of the variations observed between Danish, English and Italian from an intercultural rhetoric perspective. The study concludes with a discussion of results and implications of this thesis (Chapter 8).

A few notes on form are given below. I have limited my discussions of previous research to the research closely related to the areas I explore. Excellent summaries on the areas of contrastive linguistics, corpus linguistics, text linguistics, discourse analysis, Functional Linguistics, and intercultural rhetoric are provided in, among others, Brown & Yule (1983), Chesterman (1998), Connor (2011), De Beaugrande & Dressler (1981), Dirven (2004), Halliday (1985), Hoey (2001), McEnery, Tono & Xiao (2006), Renkema (2004, 2009a) and Thornbury (2005). Chapter 2 provides a review of the analytical framework of the study through the description of linguistic structures. In all cases, I examine the main aspects of the objects of analysis prior to discussing the analysis and the results obtained.

An omnipresent constraint on form consists in how to refer to the authors of the texts in question. Gender, age or educational or personal background of the speakers in the corpus will not be considered; I shall concentrate exclusively on their roles as speakers in the given context. Therefore, the names of the speakers only appear in Appendix A. Extra-linguistic information such as applause from the audience or interruptions by the president of the meeting shall not be included in the analyses.

Finally, a note on terminology. I shall refer throughout to the writers and readers of a text where this role is intended more generically, covering terms like producers, speakers, addressors, and recipients, hearers, addressees, too. Following the general usage to circumvent the biases of English, writers will referred to as 'she', and readers as 'he'. When describing the

single texts of the corpus, I will use the appropriate pronoun to refer to the speaker's actual gender.

I began this introduction by explaining the choice of parliamentary discourse. I shall finish with a few words on the choice of title. I see the study of the structure of discourse as an analysis of language in use. The linguistic forms are dependent on the purposes and functions they are designed to serve in human affairs, and human languages are instruments shaped to facilitate communicative and social functions in accordance with the cognitive, physiological and social skills of human beings. As argued in Danish Functional Linguistics (Harder, 2005, p. 11), languages are not autonomous structures that exist in *vacua*; on the contrary they exist in a symbiotic relationship with the extra-linguistic world. Different languages mould the extra-linguistic substance differently into unique structures – they 'cut the pie differently', as it were. But all human languages are essentially subject to same constraints and efforts, and hence comparable. In the context of this study, the competence to construct a coherent, cohesive and persuasive discourse is taken to constitute an essential human cognitive ability which is coded universally in the languages of the world. As such, it constitutes the functional-conceptual *tertium comparationis* that allows a comparison of three different languages from a linguistic point of view. However, languages are also shaped by various contextual factors such as linguistic and rhetorical traditions. Consequently, linguistic variation is expected between different languages.

2 A framework for the analysis of linguistic structures

This chapter provides the framework for the rest of the thesis. I consider linguistic structures within and beyond the sentence level, and devise a type of analysis based on what should be included in a cross-linguistic analysis of linguistic structures. Such an analysis first requires an operational definition of what linguistic structures are, and a description of how to perform an analysis based on these. This chapter supplies the current view on upper level linguistic structures, and outlines the type of analysis to be carried out.

Section 2.1 provides some theoretical background, a comparison of the various linguistic structures that together define language in action, and a review of different perspectives of the study of linguistic structures. Section 2.2 outlines how linguistic structures have been studied in other disciplines such as second language teaching, translation studies and contrastive linguistics. I then consider what components need to be included in a cross-linguistic analysis of linguistic structures in Section 2.3. Section 2.4 pre-empts Chapter 3 and refers to previous studies of linguistic structures in parliamentary discourse. We start with a tour of the concept of linguistic structures.

2.1 Linguistic structures

In the present thesis, linguistic structures are understood as the pairing of meaning and form at various levels, accounting for the rules pertaining to language use, consciously or subconsciously combined by speakers of a given language. In this way, linguistic structures are found in various fields of linguistics, from phonetics to semiotics. In this thesis, the focus is on linguistic structures within and above sentence level. In particular, I concentrate on text structure, discourse structure and information structure. There is no consensus on the definition of linguistic structures, or on the boundaries of these, even though the study of linguistic structures above sentence level is treated across various subdisciplines of linguistics: text linguistics, text analysis, discourse analysis, and discourse linguistics of texts (Connor, 1996, p. 11). As shown in Figure 2.1, which is a revised model of the different cognitive dimensions from Functional Linguistics proposed by Korzen & Lundquist (2003, p. 11), the various linguistic structures interact with each other. In the context of this thesis, the interaction between syntax and text, discourse and information structure is of particular interest, see also Carlson & Marcu (2001, p. 2): 'the boundary between discourse and syntax can often be blurry'. The idea behind Figure 2.1 is that communication is language moulded as texts that have specific purposes in a given context. This is why the model contains three main dimensions, each

divided into two substructures indicated by the dotted lines: the smallest level (I) represents the microstructure of words and phrases; the middle level (II) represents the macrostructure of sentences and whole texts; and the large level (III) represents context (see van Dijk, 1980, 2008). Human communication, then, is the result of a synergy between all dimensions. In this study, the main focus is on the structures of Level II B of text, discourse and information structure. However, since this study is cross-linguistic in nature and aims at correlating linguistic variation with contextual factors, I shall also include Level I and Level III. Different languages have different syntactic patterns, which partially are result of differences in lexical and morphological features, and partially result of various contextual factors.
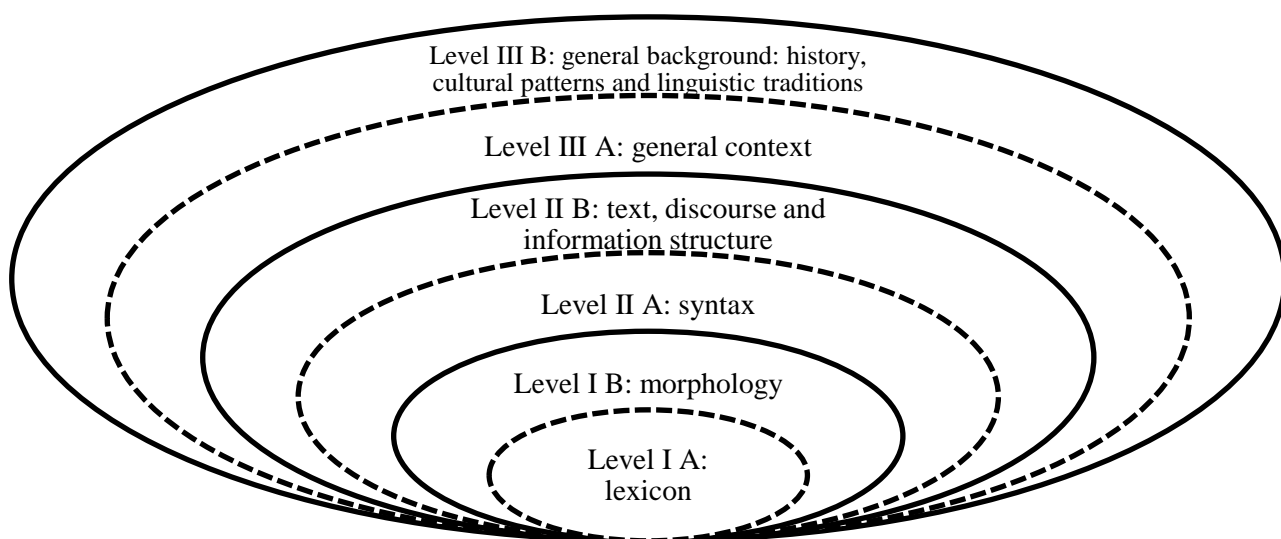
Figure 2.1: Linguistic structures and cognitive dimensions

Differences in Level I and Level II features mean, for example, that some languages tend to express the same semantic content within the same sentence or even word, whereas other languages tend to express the same content in several sentences or words (Gale & Church, 1993, p. 76; Korzen & Lundquist, 2003, pp. 10–11; Pierini, 2004, p. 186; Skytte & Korzen, 2000, p. 16). Or as argued by Longacre & Woods (1976, p. 2):

> [A]ll work on lower levels is lacking in perspective and meets inevitable frustration when the higher levels – especially discourse and paragraph – have not been analysed … In the view of these considerations, discourse analysis emerges not as an option or as a luxury for the serious student of language, but as a necessity.

Just as Figure 2.1, most of the work reviewed here originates from cognitive-functional linguistics, which is also the linguistic framework of this study. A combination of the functional and cognitive approach may seem odd, but it does not constitute an actual problem for the overall idea of language as human communication in context, cf. one of the founders of Cognitive Linguistics, Langacker (1991, p. viii), who refers to 'a cognitive-functional conception of language', where the nature of language is based on the same axioms (*pace* Jansen, 2003, p. 14). In addition to the cognitive-functional framework, I apply some of the ideas from intercultural rhetoric on how 'texts and interpretations of texts are shaped by the world and shape the world' (Eisenhart & Johnstone, 2008, p. 11), and on how 'language and writing are cultural phenomena, different cultures have different rhetorical tendencies' (Connor, 2002, p. 494). Even though intercultural rhetoric has been developed within the disciplines of second language acquisition, writing and learning, or more precisely English as a Foreign Language (EFL), it draws from all of the previously mentioned disciplines relevant to this thesis, as shown in Figure 2.2 taken from Connor (1996, p. 9). I use insights from intercultural rhetoric here to explain variations across languages with contextual factors. The boxes in the middle column of the figure show the assumptions applied by intercultural rhetoric.

This section outlines some of the different viewpoints of linguistic structures and language in general by Functional Linguistics, Cognitive Linguistics and rhetoric. We have already seen how various terms are used interchangeably in this area of linguistics. Accordingly, the section also reviews the terminology of two of the most important terms in this context, namely text and discourse. In addition, I also define the concept of information as used in this study, and sketch the differences and similarities between two other related phenomena, cohesion and coherence.
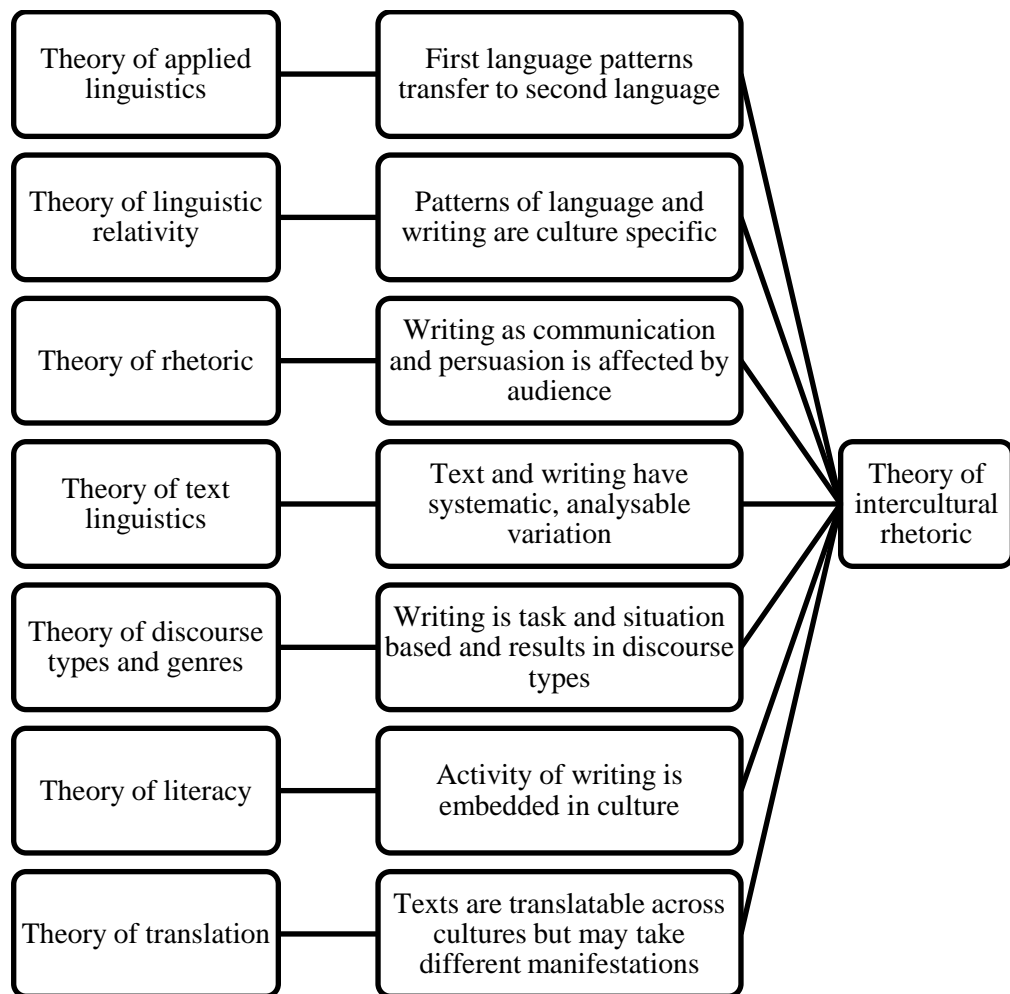
Figure 2.2: The theoretical framework of intercultural rhetoric

2.1.1 A cognitive-functional approach

Within the structural linguistic paradigm, the description of linguistic structures stops at sentence level, as no systematic description can be given of the structures above. This was to a large extent also the point of departure within the formal and generative paradigms, although some formal theories of discourse have extended the description of language to also include structures above the sentence (e.g. Asher & Lascarides, 2003; Asher, 1993; Grosz & Sidner, 1986; Polanyi, 1988). In Functional Linguistics, on the other hand, the description of structures has been extended to include text and context as objects of analysis, since language consists of components that fulfil different functions in human communication (Halliday, 1985). Finally, in Cognitive Linguistics, no actual distinction is made between linguistic structures, as all structures are symbolic (Langacker, 1991). Cognitive linguists do not see any qualitative difference between lower levels (morphology and syntax) and higher levels (text, discourse and

information structure); they see only a quantitative difference, as higher levels typically are more complex structures than lower levels in that they often contain more information.

One of the fundamental ideas in Functional Linguistics is that language is not an independent system, but a system that has developed to fulfil basic functions in human communication (Halliday, 1985). Language is used in texts that again are used in contexts. This means that language, or text, is shaped by context, and context shapes language. Functional linguists agree that meanings realised in a text never take place in a vacuum, but in a situation and in a context, although this realisation is not straightforward, as pointed out by Eggins & Martin (1997, p. 236):

> [A]n interactant setting out to achieve a particular cultural goal is most likely to initiate a text of a particular genre, and that text is most likely to unfold in a particular way – but the potential for alternatives is inherent in the dialogic relationship between language and context.

This idea of a relationship between context and language was in direct opposition to the idea of the structural and formal paradigms of language as an abstract system of rules – in structuralism referred to as '*langue*', and in formal or generative linguistics as linguistic 'competence', both being part of the well-described dichotomies '*langue-parole*' (de Saussure, 1959, p. 13ff) and 'competence-performance' (Chomsky, 1965, p. 4). Functional Linguistics manages to combine these dichotomies through the three layers of 'register' embedded in the so-called situational features: the 'field' of discourse (what is going on), the 'tenor' of discourse (what are the relations among the participants), and the 'mode' of discourse (what is the channel and genre). These three elements have direct realisations through the 'metafunctions' of language, referred to as the 'ideational', the 'interpersonal' and the 'textual' components. Accordingly, the ideational metafunction is realised through field, the interpersonal one through tenor, and the textual one through mode. A similar model of the relationship between text and context is found in Hymes' SPEAKING model (Hymes, 1972).

In the context of this thesis, Halliday's (1985) three metafunctions can be used to describe how a text hangs together: the ideational metafunction prescribes that the text is designed around one common topic, the interpersonal defines the purpose of the text, and the textual determines the relations between the various parts of the text. As such, the ideational and the textual metafunctions are of particular interest to the present study, because it is the textual

metafunction that relates the ideational metafunction, also referred to as 'the propositional content' (Lundquist, 2003, p. 231). Being mutually related, all three metafunctions do, nevertheless, play an important role in the description of a text: if one of the metafunctions is not represented in a text, the text is not a text. However, it should be noted that not being represented is not the same as not being present. As noted by Taboada (2009), a majority of the parts of text, or what I refer to as discourse units, do not contain any explicit marking of the textual or rhetorical relationship between the units, but this does not mean that the texts in which these discourse units occur are not texts. High degrees of implicitness in texts are only feasible because the participants, through their background knowledge, are able to infer which specific relations hold between the different parts. This phenomenon is usually referred to as frames, scripts and schemata, which are all different ways of representing and explaining background knowledge. The three notions have been widely applied in Cognitive Linguistics, where they are categorised as either static or dynamic structures: frames (Minsky, 1975) are static structures that can be semantic categorisations of terms (a cat is a subordinate term for an animal); scripts (Schank, 1972) are dynamic structures often referred to in terms of actions (a classroom script involves a teacher, students, chairs, books, learning, and so on); and schemata are organised instances of background knowledge that help us predict aspects of a discourse and its interpretation (Tannen, 1980). Schemata are thus used to describe the influence of different cultural backgrounds and different interests in the interpretation of discourses. Inferences fulfil an essential function in establishing structures in texts: they enrich the content of the discourses by adding information, and at the same time demonstrate what kind of background knowledge the writer requires of the reader (Irmer, 2011). Common to all theories of background knowledge is that they postulate some form of internal organisation that helps us interpret and predict events in general, and linguistic events or discourses in particular (Brown & Yule, 1983; Johnson-Laird, 1980).

In Cognitive Linguistics, a text is seen as the result of a linguistic codification of mental representation models (Skytte & Korzen, 2000, p. 18), and linguistic structures are motivated by general cognitive processes (Langacker, 1991). A mental model is the representation that a person has of a given input. If the input is supplied in the form of a text, the first words and sentences will automatically activate a mental model based on the input of the recipient, and the text will be interpreted in an interactive process of the linguistic input and the mental model (Johnson-Laird, 1983; Skytte & Korzen, 2000, p. 20ff).

The notion of mental models is important for this thesis, because it explains the links between language and knowledge on the one hand, and the process of text interpretation and discourse processing on the other. The model is also closely related to the idea of linguistic structures in Functional Linguistics, where internal organisation to a large extent is created through the ideational metafunction, which in fact represents how we imagine the world around us and textualise it by means of language. In this way, Functional Linguistics and Cognitive Linguistics can be reconciled without further complications, as both paradigms claim that language is not an independent faculty (Cruse & Croft, 2004, p. 2). In a cross-linguistic context, this notion of non-independence and the hypothesis that language is influenced by context are important, because they may serve as a cognitive-functional *tertium comparationis* (Chesterman, 1998; Jansen, 2003, p. 17; Moreno, 2008).

In sociolinguistics, language is assumed to co-constitute social life, because language users can index socio-cultural information (Ochs, 1996, p. 409). This means that specific linguistic forms (cf. the notion of textualisation above) in particular contexts have the potential to reflect and constitute social meaning (Duranti, 1997, pp. 17–20; Johnstone, 2008, pp. 133–134). In the same way, Hymes (1972, p. 56) talks about speech events that are 'activities, that are directly governed by rules or norms for the use of speech', which can be said to constitute the loci where 'communities are formed and held together' (Duranti, 1997, p. 289). This conceptualisation of how language interacts with social life and culture is also a central tenet in Register and Genre Theory (Eggins & Martin, 1997, p. 230), which seeks 'to theorise how discourse, or texts, are alike and unlike each other, and why'. In order to do this, they establish two common themes:

> Firstly, they focus on the detailed *analysis* of variation in linguistic features of discourse: that is, there is explicit, ideally quantifiable, specification of lexical, grammatical and semantic patterns in text. Secondly, [the approaches of the theory] seek to *explain* linguistic variation by reference to variation in context: that is, explicit links are made between features of the discourse and critical variables of the social and cultural context in which the discourse is enacted. *Register* and *genre* are the technical concepts employed to explain the meaning and function of variation between texts. (Eggins & Martin, 1997, p. 234)

Eggins & Martin (1997, p. 234) specify register as 'a theoretical explanation of the common-sense observation that we use language differently in different situations'. Register defines the

probabilistic relationship between context and language as either strong, where a high number of contextual factors have an impact on the text (Hymes, 1972), or as weak, in which texts are realisations of a limited number of contextual dimensions, e.g. the three metafunctions mentioned above. Register variation is not only observable at lower levels of syntax, as noted by Eggins & Martin (1997), but also at higher levels of text, discourse and information structure.

2.1.2 A rhetorical approach

Owing to their dual focus on production and processing of discourse, the three linguistic structures studied in this thesis – text structure, discourse structure and information structure – are closely related to three of the five canons in classical rhetoric, namely *inventio*, *elocutio* and *dispositio* (the other two being *memoria* e *actio*) (Lo Cascio, 1991, p. 236). I do not claim that there is a one-to-one relationship between the three linguistic structures and the three canons; on the contrary, I assume that the three canons are integrated parts of all three structures simultaneously. As a result, linguistic structures can also be approached from a rhetorical point of view, rhetoric being defined as the study of persuasive and convincing communication.

In a different but comparable view to the cognitive-functional one, Renkema (2004, p. 145ff) relates the concept of register to that of style. Thus, the notion of register as understood from a stylistic viewpoint is that variation in style can be explained by taking into account the situational factors or constraints in the SPEAKING model (Hymes, 1972) of 'setting', 'participants' and 'norms'. In this way, Korzen and colleagues (Korzen & Gylling, 2012b; Korzen & Lundquist, 2003; Korzen, 2003) talk about a correlation between formality and linguistic structures: in a formal or formalised setting such as the one of the texts employed in this study, the participants are placed in a hierarchical role structure, and a mental distance between them is created. This implies a more formal register for the purpose of the depersonalising communication. This formal register expresses an intellectualisation or 'logification' of the content; in other words, an interpretation of the discourse units and of the rhetorical relations between these is required. The depersonalising 'filter' between the participants reduces the spontaneity and the personal dimension and sensitivity. In this way, we may talk about an opposition between the social and the personal dimensions, between social status and personal involvement. A high degree of socio-cultural hierarchy will – all other things being equal – cause a high degree of logification, interpretation and intellectualisation; on the other hand, a low degree of formality is characterised by a higher degree of personal spontaneity and personal involvement. Typically, these phenomena have been investigated under the

sociolinguistic subfield of diasystems (diachronic, diatopic, diastratic, diaphasic and diamesic variations) (Bazzanella, 2008; Coseriu, 1956; Koch & Oesterreicher, 2001; Völker, 2009; Weinreich, 1954). Logification and interpretation are reflected in various ways in language, but a feature that is particularly sensitive to the different levels of formality is the morphological codification of the verbal content in discourse units, or the explication – for instance, due to the verbal conjugations – of distinctions between various pragma-narrative levels, that is, between the foreground and the background (Tomlin, 1985). The foreground-background distinction is more or less equivalent to the nucleus-satellite distinction in Rhetorical Structure Theory (Mann & Thompson, 1988), which I shall return to in Chapter 5. The morphological explication of verb forms in discourse units reflects, or grammaticalises, a cognitive hierarchy of the content in question, and may be seen as a particular effect of the intellectualisation and of the logification of a text. In this way, seeing linguistic structures from a rhetorical perspective is closely related to the viewpoints taken by the cognitive-functional approaches.

In other words, it is a matter of adapting, consciously or unconsciously, one's linguistic style to the given situation by varying factors such as ornamentation, sentence structure and word choice (Gabrielsen & Juul Christiansen, 2010, p. 43). A number of scholars (e.g. Biber, 1991; Johansson, 2007; Stubbs, 1996) have demonstrated how texts of different genres, types and languages vary across stylistic choices in terms of lexis and grammar, stressing the so-called 'Humboldt principle' (Abraham, Givón, & Thompson, 1995) of a one-to-one relationship between form and meaning. Applied to stylistics, this principle prescribes that each formulation has its own stylistic meaning, and that there is no such thing as free variation, even though some formulations may appear equivalent. Renkema (2004, p. 148) recognises three possible views on style, which, in this context, corresponds to the notion of register:

- Style as a possible form for a specific content
- Style as a choice of specific patterns
- Style as a deviation from expectations

In this thesis, I will adopt a combination of the two latter views, seeing, firstly, style and register as choice patterns available to the writer in phrasing what she would like to say. The choices may not be the same in all languages, and perhaps more important, the preferences for certain patterns may not be the same across languages. Secondly, style and register are seen as being influenced by contextual factors and by language-specific routine patterns developed to meet

certain expectations from the readers. This means that particular contexts or situations entail a specific style in order to be conceived as appropriate or fitting. Consider the following example, which in many contexts would probably be perceived as stylistically awkward (cf. Neff, Dafouz, Díez, Prieto, & Chaudron, 2004; Nir & Berman, 2010; Slobin, 1996a), even though it is grammatically correct and perfectly coherent.

1) Larry does not eat meat, and he is a vegetarian, and he lives together with Annie, and she eats a lot of meat.

Instead of the many coordinated clauses employed in example 1), readers would expect subordinate clauses expressing causal and elaborative relationships between the clauses. As argued in Chapter 3, the genre of the texts studied in this thesis, namely parliamentary discourse, is expected to be characterised by a somewhat formal and jargon-like register quite different from that of example 1). Thus, both writers and readers have expectations concerning the lexical complexity and informational density.

Even though Renkema's two views seem to differ in terms of approaching style from an objective (style as specific patterns) and a normative perspective (deviation from expectations), I consider the study of this thesis purely objective in describing and explaining differences in formulation patterns. In contrast, a normative approach would entail a study of whether the speeches by the members of the European Parliament had been able to meet the expectations of the audience. This is beyond the scope of this thesis, but it is interesting to note how much (normative) effort the European Union as an institution puts into converging the language of its politicians, employees and institutions. In a style guide from the European Commission (*How to write clearly/Scrivere chiaro/Skriv klart*, 2011), ten recommendations on how to write clearly are listed almost identically across the 23 official languages of the EU. The hints are largely similar to those provided for Danish by Pontoppidan (2013), and for English by Thornbury (2005; see below), and closely related to the Gricean Maxims (Grice, 1975), but are, of course, specifically related to writing text in the EU institutions. An interesting difference between the different languages is that the Danish and English versions recommend a mean sentence length of 20 words, while the equivalent Italian guide (p. 6) does not recommend specific sentence length.

> The recommendation, which appears in some style guides, not to use sentences containing more than 20-25 words on average is likely to be counterproductive to

our documents. The documents risk having too fragmented sentences, with the subsequent possibility of ambiguous sentences or sentences without any sense. In fact, the Italian language allows you to coordinate and/or subordinate clauses in a flexible and at the same time clear manner.[1]

In another non-language specific style guide from the European Union (*Writing for translation*, 2010, p. 7), the recommendation for translators reads as follows:

> Long and unwieldy sentences create many problems for the reader, so avoid squeezing too much information into each sentence. This does not mean that you should write very short sentences throughout the text. It is, in fact, not the length itself that creates reading and translation problems, but rather a surplus of contracted sentences, subordinate clauses or other intrusive phrases which hamper the readability of the text.

It would appear that approaching style from a normative perspective is much more a matter of personal taste, ideology or policy than approaching style from an objective perspective. Neither the formalised descriptive apparatus of Cognitive Linguistics, nor the systemic analytical framework of Functional Linguistics is employed in any strict sense in this thesis. Yet, the conception of language as a non-independent faculty plays a central role in the understanding of linguistic structures presented in the chapters to come.

2.1.3 Text and discourse revised

We have already seen how the two terms 'text' and 'discourse' are used interchangeably in this area of linguistics. In this study, the view is taken that the concepts should not, as suggested by some scholars, be conceived as synonyms or a distinction between medium and channel, since the description of linguistic structures entails the ability to distinguish between various levels mutually intertwined. The majority of the literature under review seems to either conflate the two terms, giving them the same meaning, or to contrast the terms by dealing with only one of them. However, this 'either-or' strategy appeared in need of a revision in this cross-linguistic study, which I argue can be done by drawing an analytical distinction between text as the

---

[1] My translation of *"La raccomandazione, che figura in alcune guide di stile, di non usare frasi contenenti in media più di 20-25 parole rischia di rivelarsi controproducente per i nostri documenti in quanto atta a produrre periodi troppo frammentati, con la conseguente possibilità di frasi equivoche o prive di senso compiuto. La lingua italiana effettivamente consente di coordinare e/o subordinare diverse frasi in modo agile ed al tempo stesso chiaro."*

product and form, and discourse as the process and rhetorical organisation of verbal communication.

When placing the concept of text in the general scheme of language, a useful point of departure is found in the origins of the word. The English term has its origin in the Latin word *textum* which means 'weave' or 'fabric' (Irmer, 2011, p. 44), an association which corresponds neatly with the current widespread use of the word, as found in the literature of Functional Linguistics: 'The word TEXT is used in linguistics to refer to any passage, spoken or written, of whatever length, that does form a unified whole' (Halliday & Hasan, 1976, p. 1). Hasan & Halliday's notion of a 'unified whole' refers to the fact that the units of a text are mutually related and that texts have a structure. But since texts can assume an almost infinite variety of structures and forms, from single words to thousands of words (Christiansen, 2011, p. 31), the meaning conveyed by the text is also dependent on the context. Among other characteristics, Thornbury (2005, p. 19) lists seven criteria for constructing a text as a unified whole. A text must:

- Be self-contained

- Be well-formed

- Hang together (i.e. cohesive)

- Make sense (i.e. coherent)

- Have a clear communicative purpose

- Be a recognisable text type

- Be appropriate to the context of use

Although Thornbury's almost normative criteria for constructing a text would appear to require a reasonable degree of knowledge of technical concepts such as cohesion, coherence, communicative purpose and text types as a prerequisite for creating meaningful texts, it is interesting to discover, as noted by Knott (1996, p. 3), that inventing non-texts in *any* context is actually quite difficult and rare, and that readers, as a result, will go to great lengths to find an appropriate interpretation of a text. In this way, a text may be considered a means of successful communication, cf. the cooperative principle of the Gricean Maxims (Grice, 1975).

Discourse finds its origin from Latin *discursus*, which means 'the process of understanding, reasoning and thought', giving discourse a more dynamic nature than text. However, the study

of discourse as a linguistic property is often closely related to the study of text, in particular when talking about texts and their purpose in the context of use. For instance, Brown & Yule (1983, p. 1) see the analysis of discourse as the analysis of 'language in use', just as Renkema (2004, p. 1) describes discourse studies as the investigation of the relationship between form (e.g. a statement) and function (e.g. an invitation) in verbal communication. A common ground of most discourse studies consists in investigating linguistic phenomena beyond the boundaries of the sentence, psychological structures and processes along with social interaction (Chafe, 2003, p. 441). With these definitions in mind, the two concepts of text and discourse have, not surprisingly, often been considered to be synonyms.

The term discourse has also come to be used to describe activities at the intersection of distinct disciplines such as those of social science and linguistics. At this intersection, the work of Foucault (1969), in which discourse is related to the concept of power, has fostered the interdisciplinary Critical Discourse Analysis approach (CDA) led by scholars such as Fairclough, Wodak and van Dijk (Fairclough, 1989; Flowerdew, 2008; Wodak & van Dijk, 2000). In CDA studies, the typical data consists of political or economic discourse which is critically examined from a linguistic viewpoint in terms of how discourses exhibit degrees of power distance and means of suppression. Even though this thesis deals with political texts, the CDA approach will not be pursued, as the primary analytical focus is on *form* rather than on *content*.

As outlined above, the view taken in this thesis is best characterised as a combination of the text-as-product (static) and the discourse-as-process (dynamic) views. This distinction between linguistic structures resonates with previous definitions formulated by Brown & Yule (1983) and Widdowson (1979), which also have been employed in Rhetorical Structure Theory (RST), where 'unstated but inferred propositions [i.e. rhetorical relations] … arise from the text structure in the process of interpreting texts' (Mann & Thompson, 1987, p. 3). The main difference between the two concepts is, indeed, based on a distinction between studying text as a product without taking into consideration how it was produced or how it could be received, and studying discourse as a process where words, clauses and sentences are considered to be evidence of the writer's attempt to communicate a given message to the reader. In the context of RST, discourse analysts are particularly interested in investigating how a reader might comprehend the writer's intended message, especially with regard to how the different parts of the texts are related to each other. This approach is summarised by Brown & Yule (1983, p. 24) in the following way:

This is clearly an approach which takes the communicative function of language as its primary area of investigation and consequently seeks to describe linguistic form, not as a static object, but as a dynamic means of expressing intended meaning.

Although Brown & Yule originally used the text versus discourse distinction to promote the discourse approach rather than that of text, a combination of the two approaches is, nevertheless, possible and has accordingly been adopted by scholars in Functional Linguistics, e.g. Martin (1992) and Matthiessen & Thompson (1988), who combine the study of syntactic relations (text) with that of rhetorical relations (discourse). In this respect, the notions of conjunctive relations (Halliday & Hasan, 1976; Martin, 1992) in Systemic Functional Linguistics, which in many respects are similar to rhetorical relations, are based on grammatical description and classified into coordinate and subordinate relationships within and across sentence boundaries. In computational linguistics, Knott (1996) also used a number of syntactically motivated 'cue phrases' in order to build a set of coherence (i.e. rhetorical) relations.

The deceptively simple nature of the text versus discourse distinction deserves more discussion than it has received in the literature so far, which the following review of previous approaches and definitions clearly demonstrates (see Korzen & Gylling, 2012a). There is little agreement in the literature on how the two concepts should be defined. The conceptualisation of text being the same as discourse is a particularly common standpoint, as found in the *International Encyclopedia of Linguistics* (Chafe, 2003, pp. 439–440):

The term discourse is used in somewhat different ways by different scholars, but underlying the differences is a common concern for language beyond the boundaries of isolated sentences. The term text is used in similar ways. Both terms may refer to a unit of language larger than the sentence: one may speak of a *discourse* or a *text*.

Irmer (2011, p. 43) is a recent example of a scholar who uses the terms interchangeably: 'Generally, a *text* or a *discourse* is a sequence of natural language utterances'. The same viewpoint is found earlier in Stubbs (1996, p. 4), who adds that 'sometimes this terminological variation [between text and discourse] signals important conceptual distinctions, but often it does not, and terminological debates are usually of little interest'. Also, Halliday & Hasan (1976, pp. 2–4) use both terms in their definition of a text:

> A text has texture and this is what distinguishes it from something that is not a text. … The texture is provided by the cohesive RELATION [which is set up] where the INTERPRETATION of some element in the discourse is dependent on that of another.

Similarly, both terms – discourse and (co-)text – are found in Rijkhoff (2008, p. 90), who states that 'discourse in the sense of co-text is a linguistic entity'.

Conversely, there are scholars who adhere to the view that text refers to written language and discourse to spoken language. For instance, Stubbs (1983, p. 9) notes that 'One often talks of *written text* versus *spoken discourse'*, and similarly Riazi (2003) states that:

> The first [approach] is *discourse analysis*, which mainly focuses on the structure of naturally occurring spoken language, as found in such *discourses* as conversations, commentaries, and speeches. The second approach is *text analysis*, which focuses on the structure of written language, as found in such *texts* as essays and articles, notices, book chapters, and so on.

However, this distinction is rejected by other scholars such as Tannen (1982, p. ix), who argues that:

> Discourse … refers to both text and talk, and these not as two separate genres to be compared and contrasted, but rather as overlapping aspects of a single entity. As the object of study, spoken discourse is "text", much as words spoken in a speech are commonly referred to as the text of the speech. In this sense, *discourse* and *text* are synonymous.

Christiansen (2011, p. 34) also states that '[i]n non-linguistic and non-semiotic circles, text is sometimes used for examples of written language and discourse for the spoken. Nowadays linguists accept that such a distinction based only on medium and channel is simplistic'.

A third group of scholars see discourse structure as the rhetorical organisation of a text (Mann & Thompson, 1988), an organisation definable as a series of rhetorical relations between text segments, created in the process of human communication (Brown & Yule, 1983; Scarpa, 2001, pp. 28–32; Widdowson, 1979). Widdowson (Widdowson, 2004), who overtly criticises one of the first scholars to introduce discourse studies, Harris (1952), for conflating the terms text and discourse, states more precisely that:

> Discourse in this view is the pragmatic process of meaning negotiation. Text is its product. ... The discourse may be prepared, pre-scripted in different degrees. ... But whatever the degree of prescription, the text, the actual language that realises the interaction, is immediate to it, and is directly processed on line.

Seemingly inspired by Halliday & Hasan (1976, p. 300) and their claim that 'discourse … come[s] to life as text', Christiansen (2011, p. 34) similarly states that 'text [is] the form, discourse the content', and along the same lines, Cornish (2009, pp. 99–100) concludes his definition of the two terms by stating that:

> Text, then, refers to the connected sequences of signs and signals, under their conventional meanings, produced by the speaker … Discourse, on the other hand, refers to the hierarchically structured, mentally represented product of the sequence of utterance, propositional, illocutionary and indexical acts that the participants are jointly carrying out as the communication unfolds … Text, in normal circumstances of communication, on the other hand is essentially linear, due to the constraints imposed by the production of speech in real time.

Similarly, Ruiz Ruiz (2009) remarks that:

> [T]he two concepts [discourse and text] should not be confused or equated. Indeed, every piece of discourse has a textual form or can acquire it; the same text may include different discourses or the same discourse may adopt different textual forms.

As mentioned above, in this thesis I shall follow this latter group of scholars and their definitions of discourse as the process and rhetorical organisation of verbal communication and text as the (oral or written) product and form. Both discourse and text can be analysed with regard to their internal relations and structures, but methodology and terminology vary. An important factor in defining the two concepts also consists in the scope and aims of the study in question. In this thesis, where the study has a cross-linguistic perspective and a corpus-based approach, the notion of text structure shall be applied to account for dependency and realisation of the single text parts (discourse units), while discourse structure shall refer to the hierarchy of rhetorical relations between the discourse units and the sources of coherence that can be used to describe these.

2.1.4 Defining information

Although the term information has many meanings, as noted by Jansen (2003, p. 29ff), the use of the term in connection with linguistic structures can be considered quite uniform compared to those of text and discourse. Information structure accounts for the transition from the first element in a clause towards the rest of the clause, usually referred to as thematic realisation and thematic progression (Perfetti & Goldman, 1974). The general assumption in information structure is that every discourse unit has backward-looking links, while making the discourse move forward. This is how language is able to utter something new. Discourse units thus link to previous material and contribute with something new. This holds at the level of the text, but also at the syntactic level. Vallduví (1992, p. 35) mentions the overall agreement − despite disagreements on details – that 'in the sentence there is some sort of informational split between a more informative part and a less informative part'. The split happens at the point where the part, or unit, that is used as an anchor to the information, in order to guarantee that the information will enter the reader's knowledge store, gives way to the informative part. Lambrecht (1994, p. 1) begins his book on information structure by pinpointing the terminological disagreement on the details, too:

> There has been and still is disagreement and confusion in linguistic theory about the nature of the component of language referred to in this book as INFORMATION STRUCTURE and about the status of this component in the overall system of grammar.

Names for the parts of this component of discourse vary, and the exact definitions are also elusive. Vallduví & Engdahl (1996, p. 465) revise the different labels to what they call 'focus' and 'ground'. For focus: new, NewInfo, rheme, dominant constituent. For ground: background, presupposition, open-proposition, given, theme, topic. More detailed descriptions of these terms can be found in Firbas (1974), Halliday (1967), Hockett (1958), Krifka (1993), and Vachek (1966).

The idea that some parts of a discourse are more informative than others is used in this thesis to represent how some information is packaged within and across sentences and clauses. Information structure shall be used to refer to the notion of information packaging as proposed by Fabricius-Hansen (1996, 1999) and Vallduví & Engdahl (1996), and that of informational density as used by Fabricius-Hansen (1996) and Jansen (2003):

[W]e would probably say that the informational density is higher in A than in B if at least one of the following conditions holds, other things being equal: i. the average amount of discourse information per sentence is higher in A than in B; ... (Fabricius-Hansen, 1996, p. 529)

More precisely, it will be assumed that the informational density is higher in A than in B if both contain the same amount of information but A is shorter than B, and/or A contains more discourse units than B. The concept of discourse units is defined in more detail in the following chapter, but corresponds essentially to clauses. In this thesis, information structure includes two objects of analysis (introduced in Chapter 1), namely linkage of discourse units, which can be either intrasentential or intersentential, and the signalling of rhetorical relations, which can be either explicit or implicit.

2.1.5 Cohesion and coherence

Two other phenomena of particular interest when defining linguistic structures are cohesion and coherence. In the literature, cohesion has by some scholars (Halliday & Hasan, 1976; Stede, 2011) been described as a central phenomenon of textuality and thus related to text structure as defined above, because detecting cohesion (or cohesive ties) does not require deep understanding of a text. By contrast, noting coherence involves uncovering how the discourse units of a text are mutually related to each other. This view has been questioned by other scholars who argue that cohesive ties also require a certain degree of interpretation, and that these may, in fact, be used by writers to facilitate reading and comprehension by readers (Källgren, 1979). Since the literature on cohesion and coherence is notorious for its terminological confusion (see Christiansen, 2011, p. 32), I shall briefly explain how the notions of text, discourse and information outlined above relate to the concept of cohesion and its closely related companion, coherence.

Following Halliday & Hasan (1976), I take cohesion and coherence to constitute two different, albeit mutually related and overlapping phenomena which are present in all texts with internal continuity and organisation. As such, cohesion can be defined in terms of the connections which have their manifestations in the text itself (Renkema, 2004, p. 103). Thus defined, the five types of cohesion traditionally recognised in the linguistic literature on text and discourse, i.e. substitution, ellipsis, reference, conjunction and lexical cohesion (as introduced by Halliday & Hasan, 1976) are seen as different ways of constructing a meaningful text. In this sense, cohesive ties constitute the 'blocks' that bind together the single parts of a text. In

contrast with coherence, cohesion refers to the entities mentioned explicitly in the text, and the most important (and perhaps studied) device is anaphora (Irmer, 2011, p. 45), defined as an expression (e.g. a noun phrase or a verb phrase) referring back to a previously mentioned entity or concept also known as the antecedent. While I shall refrain from endeavouring to define anaphora *per se*, it will be useful to make reference to anaphora since anaphors are not only restricted to lexical and syntactic devices such as pronouns, noun phrases or verb phrases, but also occur in a number of discourse cues such as coordinators, subordinators, conjunct adverbs, prepositional phrases and phrases which take sentential complements (Quirk, Greenbaum, Leech, & Svartvik, 1985). In example in 2) taken from the corpus employed in this thesis, the prepositional phrase in the last sentence, *at this point*, containing a propositional anaphor refers back to the antecedent, *1 January 2003*, in the first sentence. Other cohesive signals are found in connectives, word repetition, related words and pronouns; in other words, all devices that make a text hang together; see Thornbury's criteria of textuality above.

2) Of course, at its core is equal access to Community waters which was derogated until 1 January 2003. Only at this point will the CFP fully come into force. <ep-02-05-29.txt:37>[2]

Coherence, which by many scholars is considered one of the most salient issues in the study of discourse (Renkema, 2004; Taboada, 2009), refers to linking different parts of the texts together at the level of semantic and pragmatic interpretation, thus capturing 'the inner logic' of the text (Stede, 2011, pp. 79–80). Due to the often implicit nature of coherence, definitions of coherence vary in the literature. The same holds true for the terminology. One of the central phenomena in the study of coherence is coherence relations. Coherence relations, which have also been termed 'discourse relations' (e.g. Asher & Lascarides, 2003) or 'rhetorical relations' (Mann & Thompson, 1988), performing essentially the same function, refer to the specific semantic or pragmatic relationship that holds between various units of text. Henceforth, I shall only use the term rhetorical relations. Varying from theory to theory, typical semantic relations include cause, result, time and elaboration, whereas pragmatic relations typically refer to speech acts such as motivation and enablement. The rhetorical relations employed in this thesis are presented in Chapter 5, and the definitions of these can be found in Appendix B.

---

[2] The numbers following each Europarl ('ep-') text should be read as follows: YY-MM-DD; 'txt' indicates SPEAKER ID. In this thesis, the examples originate from non-translated Europarl texts. Examples in Danish and Italian shall be followed by an English translation – in some cases, the official EU version.

Coherence, thus, spans the domains of cohesion and vice versa, often occurring intertwined. This relation is captured in the Figure 2.3, which Irmer (2011, p. 52) attributes to Markus Egg. [U_1] and [U_2] are two discourse units between which a coherence (or rhetorical) relation can be interpreted. [R_n] represents the way discourse units are grouped recursively into text spans. In the lower part of the figure, the cohesive ties between events [e] and entities [x] and [y] of the discourse units are represented by direct/co-referential and indirect/associative anaphors.
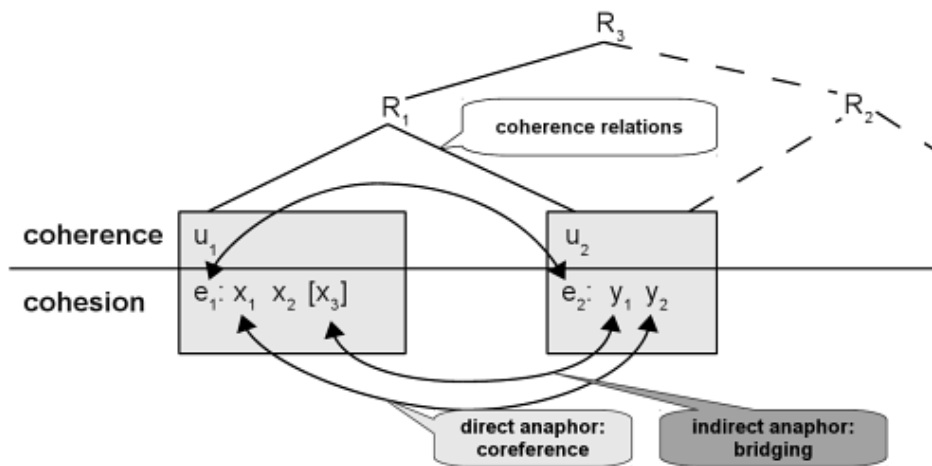
Figure 2.3: Coherence and cohesion

This clarification of the relation (or difference) between cohesion, as an overt textual phenomenon signalled in the surface of a text, and coherence, as a covert phenomenon, is important for the present thesis, because the interpretation of the specific rhetorical relations between two discourse units is often not explicitly marked (Bateman & Rondhuis, 1997, p. 3), but may in fact be interpreted through explicit cohesive ties such as anaphora and discourse cues that occur in the given discourse units (Irmer, 2011, p. 358). As noted by Stede (2011, p. 20), signals of cohesion in text typically serve as indicators of topic continuity, whereas the absence of cohesive ties typically indicates a topic shift. In this way, establishing coherence depends on cohesion and vice versa.

In the traditional branch of text linguistics pursued by De Beaugrande & Dressler (1981), cohesion and coherence also play a central role, being two of the seven criteria that distinguish a text from a non-text. The five other criteria are: intentionality (the intention of the writer), acceptability (the attitude of the reader), informativity (new information connected to known), situationality (place, time, social situation, etc.), and intertextuality (reference to other texts). Many of these criteria overlap with those of Halliday's (1985) metafunctions and Hymes' SPEAKING model (1972).

2.1.6 Linguistic structures: an overview

The definition of linguistic structures that I propose here is one where linguistic structures are primarily organisational-determining characteristics of texts. A given text is perceived as being a text because of its structural characteristics, that is, because the text is organised in a certain way. For that recognition to happen consensus must be established that a series of contextual, cultural and social factors shape the textualisation of linguistic structures. The linguistic structures appropriate to a text are determined by the function of the text in a given situation. The broad term 'function' encompasses two different aspects: the communicative purpose of the text and its social function. This is not a novel definition, nor a ground-breaking reformulation of the concept, but an operational definition, which will prove useful in approaching the texts in the corpus as instances of writing influenced by contextual factors.

The units of analysis are discourse units, which effectively correspond to clauses, and the rhetorical relations between these. In text structure, the focus is on how the discourse units are syntactically dependent on each other, and on how the discourse units are realised syntactically. In discourse structure, the focus is on the hierarchy of the rhetorical relations and on the sources of coherence. And finally, in information structure, attention is paid to the packaging of information, understood as the linkage of discourse units and as the signalling of rhetorical relations. The three linguistic structures, their foci and their individual objects of analysis studied in this thesis are summarised in Table 2.1.

	Text structure	Discourse structure	Information structure
Description	form and product	content and process	information packaging
Focus	syntactic relations	rhetorical relations	discourse units and rhetorical relations
Objects of analysis	i) syntactic dependency ii) realisation	i) rhetorical hierarchy ii) source of coherence	i) discourse unit linkage ii) relation signalling
Taxonomy	i) coordination/ subordination ii) finite/non-finite/verbless	i) coordination/ subordination ii) semantic/pragmatic	i) intrasentential/ intersentential ii) explicit/implicit

Table 2.1: Overview of linguistic structures

41

As mentioned in Section 2.1.2, the use and frequencies of the elements in the different taxonomies are likely to be affected by the contextual factors of the given communication situation. In particular, I shall focus on linguistic and rhetorical traditions in the three languages that I investigate. In intercultural rhetoric research, the concept of context is treated within the large domain of culture. Culture is divided into large and small cultures, following Holliday (1994, 1999). Large cultures have ethnic, national, or international features as essential components and tend to be normative and prescriptive. Small cultures, on the other hand, are non-essentialist and based on dynamic processes that relate to organised behaviours within social groupings. Small cultures avoid culturist ethnic, national, and international stereotyping (Holliday, 1999, p. 240), and are rooted in activities. A specific discourse is one of the products of small cultures (Holliday, 1999, p. 251f). The texts employed in my study can therefore be seen as the discourse of a discourse community that shares many small culture features (all texts have been created by persons with the same profession, and from the same institution), but also as discourses of a discourse community that exhibits large culture differences such as nationalities. This is the way in which the context of the texts shall be interpreted. Contextual influences on linguistic structures are considered in Chapter 7.

2.2 The relevance of linguistic structures to other disciplines

Linguistic structures within and above the level of the sentence have become an area of stimulating empirical research, but, as mentioned in Chapter 1, many of the studies are of a monolingual nature, examining English. In intercultural rhetoric research, which is mainly aimed at second language learning, Kaplan's (1966) study of paragraph development in essays by students learning English as a foreign language is generally considered the foundation of the research discipline (Connor, 2002, p. 494), which at that time was referred to as 'contrastive rhetoric' (Connor, Nagelhout, & Rozycki, 2008). The main conclusion of Kaplan's paper was that different cultures use different strategies to organise paragraphs and hereby content – a conclusion that inspired many to start including cultural aspects in foreign language teaching. However, it also provoked others to argue that the study was too general and failed to address the many aspects of culture (see Connor (1996, pp. 1–28) for an overview and discussion). One of the reasons that Kaplan's work was criticised is found in his division of cultures into four classes according to their paragraph development, where English, or Anglo-European, texts were said to be developed linearly, whereas Romance languages included material that, from a linear point of view, is irrelevant. Although this division is highly controversial, not based on

sufficient empirical evidence and very ethnocentric from an Anglo-European point of view, it is widely cited, also by non-English scholars. In the Romance literature, Pierini (2004, pp. 186–187), for instance, claims that Kaplan's distinction corresponds neatly with the different styles of thought found in English and Italian argumentative texts.

> The two [argumentative texts] exhibit different stylistic features, determined by different rhetorical norms operating in Italian and English. The Italian segment is made up of one very long, highly complex sentence, where phrases are rich in modifiers. The English segment is made up of five short, simple sentences, consisting of one or two clauses, where the structure of phrases is rather simple. Researchers have observed that generally, Italian prefers hypotaxis (subordination), while English prefers juxtaposition and parataxis (coordination). They have also observed that English has a 'direct', linear way of structuring information, while Italian, like other Romance languages, displays freedom to digress or to introduce extraneous matter. … Researchers have observed that the more specialised the genre, the more the individual style recedes to make way for genre-specific conventions, which can range from the use of special terms to the preference for certain syntactic structures and phraseological units…

Pierini and Kaplan's claims are supported by empirical evidence from a number of other scholars. Among those, a number of Danish scholars (Jansen, 2003; Skytte, Korzen, Polito, & Strudsholm, 1999; Skytte & Korzen, 2000) have shown how the structural differences between Italian and Danish are similar to those found between Italian and English (cf. also Lo Cascio 1991). In the same way, Lundquist (2005) shows differences between a more complex French sentence and text structure and a more simple Danish sentence and text structure: French sentences show a higher degree of subordination and non-finite realisations, whereas Danish sentences more frequently are coordinated with finite verb forms (cf. also Lehmann 1988). In joint works by Germanic and Romance scholars (Baron, 2003a; Korzen & Herslund, 1998; Korzen & Lavinio, 2009; Korzen & Lundquist, 2003), these differences are attributed to typological and contextual differences between Germanic/Scandinavian and Romance languages in general. Within the Germanic languages, a number of scholars (Fabricius-Hansen, 1996, 1998, 1999; Fetzer & Speyer, 2012; Hansen-Schirra, 2007; Ramm & Fabricius-Hansen, 2005; Teich, 2003) have shown that English, German and Norwegian also differ in terms of how

information is structured internally from a translational point of view. As a general observation, German tend to employ a more complex sentence structure than Norwegian, and to some extent also English. In this way, knowledge of how different languages structure information at various linguistic levels plays an important role for second language learning, linguistic typology and translation studies.

2.3   Components of a cross-linguistic analysis of linguistic structures

The discussion so far has concentrated on the concept of linguistic structures. Next, we shall consider the application of the approach to the analysis of linguistic structures. Before we embark on this study, we first need to consider the kind of analysis and the tools that shall be used. This section discusses a 12-step analysis from the domain of intercultural rhetoric research, using comparable corpora proposed by Moreno (Connor & Moreno, 2005; Moreno, 2008). The basis for conducting such an analysis is to establish the *tertium comparationis* at various levels, the two most important being the choice of primary data and the creation of comparable textual concepts, such as cohesion or coherence, which shall lead to comparable analyses of the linguistic and textual data. The 12 steps, reproduced from Connor (2011, p. 49), are:

1. Formulating clear hypotheses about the relationship between writing cultures and how textual meanings are expressed.
2. Defining the population of accomplished, or expert, L1 texts that can be considered comparable and specifying the basis of the similarity constraints.
3. Selecting a representative sample of that population in each of the writing cultures compared.
4. Identifying comparable textual units.
5. Validating those units of analysis as recognisable functional or pragma-discursive units by language users in each culture either through literature review or further research.
6. Quantifying the occurrence of these textual universals in each corpus.
7. Devising objective criteria to describe the textual realisations of the universals proposed in the languages.
8. Applying any devised analytical criteria to the description of the corpora independently.
9. Juxtaposing the taxonomies.
10. Contrasting the quantitative results for each comparable qualitative category.

11. Interpreting the significance of quantitative similarities and differences by statistical analysis.

12. Drawing conclusions about the relation between writing cultures and how textual meanings are expressed on the basis of the comparative results.

In this thesis, these steps shall be followed more or less slavishly. As a first step, the hypotheses were formulated above, while steps 2 to 7 shall be presented in Chapter 3. Steps 8-10 are studied in Chapters 4-6. I interpret the results and draw contrastive conclusions of steps 11 and 12 in Chapter 7, and conclude the thesis in Chapter 8 with a summary and discussion.

2.4   Previous studies of linguistic structures in parliamentary discourse

The corpus employed in this study contains instances of a specific text type. More specifically, the corpus consists of instantiations of a particular genre exhibiting certain, often conventionalised, linguistic and textual characteristics. Some of these require further illustration and explanation, as it is assumed that structural phenomena such as rhetorical relations and textualisations exhibit different frequencies in different genres and text types (Biber, Conrad, & Reppen, 1998; Biber, 1991; Korzen & Gylling, 2012b; Swales, 1990; Taboada & Mann, 2006a, p. 449; Van der Vliet & Redeker, forthcoming). As such, the corpus also contains a particular variety of the three languages under investigation, pertaining to a specific context of use which we conceive as deviating in some way from what might be considered a general norm or standard. The risk, however, lies in allowing ourselves to believe that the object of our interest in some way represents a discrete linguistic phenomenon separate from the language as a whole. For if this were the case, we would of necessity be dealing, not with an inclusion-like relationship between related phenomena, but rather with quite different, mutually exclusive, languages, making it difficult to determine the nature of the relationship between specific languages such as Danish, English and Italian, and specific varieties of such languages. In what follows, the notions of general language and institutional language are considered and, more specifically, of parliamentary discourse as a subset of institutional language. In doing so, the works of Bayley (2004) and van Dijk (2000) are drawn upon.

While discursive practices have shaped some institutions, other institutions have shaped discursive practices themselves. The study of how language functions within specific institutional contexts has attracted a considerable amount of interest over the last few decades (Bayley, 2004, p. 7), both from what Okulska & Cap (2010, p. 13) refer to as 'political linguistics' and 'analysis of political discourse', notably conducted within the paradigms of

Systemic Functional Linguistics (Halliday, 1985; Martin & Rose, 2003) and Critical Discourse Analysis (Fairclough, 1989; Wodak & Chilton, 2005). However, only a few studies on the language of parliament have been carried out; exceptions are some of the previously mentioned authors such as Bayley (2004), Chilton (2003), Fløttum (2013) and Wodak & van Dijk (2000). These studies typically examine debates held in national European parliaments, only a few of which adopt a cross-cultural perspective. Bayley (2004, pp. 9–10), who is one of the exceptions, argues that even though parliamentary discourse is freely available on the Internet, parliamentary discourse suffers from lack of academic interest. One of the reasons for this could be that the attention paid by the mass media and the general public is limited. In this regard, the European Parliament is no exception. Among previous key studies on the linguistic properties of the institutions of the European Union, attention is devoted to aspects such as the European language constellation (De Swaan, 1999), institutional multilingualism and its possible reforms (Mamadouh, 1999), linguistic capital and symbolic domination in the EU (Loos, 2000), the representation of EU in the media (Just, 2009), and European language identity related to the EU enlargement (Bellier, 2002). Other studies, such as those of Abélès (1992) and Mamadouh (1995), have examined (language) practices at a specific EU institution, such as the European Parliament, or analysed EU organisational discursive practices (Baron, 2009; Born & Schütte, 1995; Ilie, 2010; Muntigl, 2000; Wodak, 2000).

Linguistic practices in the EU have also been examined from a contrastive point of view. In an introspective paper, former Italian EU Translation Director, Cosmai (2000), points out some interesting aspects of the EU language from a translational perspective. Citing other scholars, politicians and news reporters, he refers to the EU language as a sublanguage that is so difficult to understand for non-experts that it is actually a language on its own.

> To define such a language, the mass media of the 15 member states has shown great imagination. The British have referred to it as *Eurojargon*, *Eurofog* or *Eurospeak*, by analogy with the Newspeak described by George Orwell, the French *eurobabillage*, *brouillard linguistique européen* or argot du *Berlaymont*, the symbolic name of the building housing the European Commission, the German *Eurowelsch* or *Eurokauderwelsch*, while in Italy it is disparagingly referred to as

"eurocratese." All these terms appear strongly negative, and seem to indicate a vague and impenetrable language ... (Cosmai, 2000, p. 2)[3]

The complexity of the EU language is also discussed by Mori (2003), who claims that Eurolect (*euroletto*), as she calls it, acquires its morpho-syntactic features from English and its textual features from a more standardised transnational model. Lundquist (2005), who has also observed this, argues that these standardised textual features, *in casu* pragmatics, may jeopardise democracy in the EU, since different nationalities, and language users, do not interpret EU texts in the same way. But while most of the previous studies have confined themselves the examination of legal texts such as treaties, regulations and rules, that is, texts carefully elaborated and formulated and texts that need to follow certain conventions, the texts under investigation in this thesis have a more spontaneous and personal nature, entailing different characteristics.

Despite the language heterogeneity of the present study, it is assumed here that the corpus texts are in fact homogeneous with respect to communicative function. In addition, all the speeches were held by members of the European Parliament in the same span of time. For this reason, the texts can be referred to as speeches, hereby categorising them as a functionally discrete text genre based on the criterion of communicative purpose suggested by Swales (1990). This categorisation is also to be found in the works of Girnth (1996) and Reisigl (Reisigl & Wodak, 2009, pp. 90–91; Reisigl, 2007, pp. 34–35), who differentiate between various 'fields of political actions', 'political (sub)genres' and 'discourse topics'. The fields of action of parliamentary discourse are multiple, ranging from law-making procedures, formation of public attitudes, opinion and acceptance of political control. According to Bayley (2004, p. 1), parliamentary discourse is the most formal and institutionalised variety of political subgenres. The characteristics of parliamentary discourse are also described by van Dijk (2000, p. 47) as consisting of a number of prototypical non-exclusive features, as also stated by Bayley (2004, p. 13):

---

[3] My translation of *'Per definire tale linguaggio, la stampa dei 15 Stati membri ha dato prova di grande immaginazione. Gli inglesi lo hanno denominato* Eurojargon, Eurofog *o* Eurospeak*, in analogia con il* Newspeak *descritto da George Orwell, i francesi eurobabillage,* brouillard linguistique européen *o* argot du Berlaymont*, dal nome dell'edificio simbolo della Commissione europea, i tedeschi* Eurowelsch *o* Eurokauderwelsch*, mentre in Italia esso viene chiamato spregiativamente "eurocratese". Tutti questi termini appaiono di segno fortemente negativo, e sembrano indicare un linguaggio nebuloso e impenetrabile ...'.*

[I]ts peculiar turn-taking procedures and the authorised terms of address that are associated with it, do seem to be typical of this discourse model. … Thus it might be argued that particular combinations of certain features involving levels of linguistic and discursive analysis, such as some phonological features, interaction strategies, intervention length, terms of address, metadiscursive and argumentative lexis, direct and indirect quotation, explicit expressions of belief and opinion, epistemic modality tending towards low certainty rather than probability, and complex structures of subordination favouring conditionals and concessive, give parliamentary language its distinctive and recognisable flavour.

From a structural point of view, many of these features are of interest, and in this cross-linguistic context, any difference between the languages calls for further investigation. In fact, Bayley (2004, p. 14) argues that parliaments, and consequently their linguistic practices, are sensitive to the context of culture and history in the widest sense, causing a ritualised and rule-bound language. He gives two examples from the national parliaments of the United Kingdom (*House of Commons*) and Italy (*Camera dei deputati*), where rules determine linguistic choices: whereas in the House of Commons, members are not allowed to read from a prepared text, reading from a written text in the Italian *Camera dei deputati* is common practice. The Danish parliament (*Folketinget*) in this regard resembles Italian rather than British practice. I assume that these differences may affect the speakers' way of constructing their speeches in the European Parliament. In the same way, van Dijk (2000, p. 69) states that ritual and procedural rules also determine the formal staging structures, or schemata. In the debates of the European Parliament, the procedural rules entail that the debate is composed of a sequence of monologues which are intertextually and intratextually related as parliamentarians respond to what has been said previously. Thus, it is multi-voiced and − of its nature − dialogic. The speeches can be regarded as parts of a larger negotiation, where compromises, modifications and new proposals are instantly negotiated. However, the texts of the corpus employed in this study do not contain speeches that are directly related to each other as responses, following the criteria described in Chapter 3.

Two possible weaknesses may be perceived in this general outline of the characteristics of parliamentary discourse. First of all, a large proportion of parliamentary work is dedicated to matters of routine and to uncontroversial questions. It is not uncommon that speeches are made to practically empty seats, especially in the European Parliament. Viewing the speeches in this

corpus as controversial and polemic debate may be a misinterpretation, as only a small number of debates are actually 'real'. Secondly, rather than dealing with parliamentary discourse in its complete sense, the analyses are based on official transcripts. These may be inaccurate, and even if they are accurate, they lack a fundamental dimension of parliamentary discourse: its spokenness (Slembrouck, 2009) and most extra-linguistic material. In Chapter 3, I describe how these two weaknesses have been taken into consideration.

2.5   Summing up

My goal in the analysis of these texts is to study the general textual patterns in each of the three languages Danish, English and Italian independently. My goal is also to examine how choices in text structure, discourse structure and information structure constitute the features of argumentative parliamentary discourse. I explore the deviations from the general patterns, that is, where and when those appear. And finally, I compare all three languages in order to reveal possible linguistic strategies and persuasive styles captured by the linguistic structures. I have chosen an approach inspired by insights of cognitive-functional linguistics and intercultural rhetoric because such an approach can help us 'understand texts from textual, contextual, and social points of view' (Connor, 2011, p. 44).

The first step is to conduct an analysis of the three different linguistic structures in the parliamentary discourses in each of the three languages:

- text structure
- discourse structure
- information structure

Once I have studied these three structures, their foci, and the objects of analysis (see Table 2.1) for each language, I shall establish similarities and differences in the textualisation of the three structures in each language. Chapters 4 to 7 cover the characteristics and comparisons of each of the elements of analysis. Before describing the analysis, Chapter 3 provides a detailed description of the corpus, how it was collected and how it has been annotated.

Cognitive-functional approaches propose a universal theory of the relationship between texts and contexts. Intercultural rhetoric approaches propose a plausible link between linguistic style and contextual factors. This study provides a description of linguistic structures that maps contextual variables to linguistic strategies and persuasive styles in argumentative texts. Once we understand these strategies and styles, we will be able to recognise how contextual factors

such as linguistic and rhetorical traditions affect the internal organisation of texts, and hopefully be able to communicate with people from other cultures more efficiently.

3 Data description

The texts used in this study were extracted from the Europarl Corpus, which is an open source text collection compiled by Philipp Koehn at the University of Edinburgh. The present description of the corpus is based on the Europarl reference paper (Koehn, 2005) and on information from the Europarl Corpus website (http://statmt.org/europarl/). Officially, Europarl is a parallel corpus consisting of both L1 and L2 texts, but it could also be seen as a collection of comparable corpora including 'the same proportions of the texts of the same genres in the same domains in a range of different languages in the same period' (McEnery & Xiao, 2007, p. 20). Moreover, the corpus could also be considered a special corpus containing sublanguage material or a monitor corpus, understood as a constantly moving collection of text. In this thesis, I mainly use the corpus comparably.

The corpus, which as of June 2013 has been issued in seven versions, the latest in May 2012, consists of parallel texts (i.e. original texts aligned with their translations) in all EU languages, amounting to some 60 million words per language. The texts are transcripts of speeches held by the members of the European Parliament and other politicians in the years 1996-2011. The speeches are also published in the proceedings of the European parliamentary debates. Owing to the large number of languages and words included, the corpus plays an important role in training and evaluating statistical machine translations and is one of the most widely used corpora in computational linguistics and translation studies (Cartoni & Meyer, 2012; van Halteren, 2008). However, the corpus can also be used comparably for cross-linguistic studies, as some texts contain metadata about which language has been used, who held the speech, and the political affiliation of the speaker. In this way, it is possible to compile a corpus containing texts in different languages that in many respects are similar and hence comparable.

There are three reasons for using the Europarl Corpus in the present study. Firstly, whereas most scholars using Europarl seem interested in the size and languages available in the corpus, my aim is to approach the corpus from a broader perspective taking contextual and linguistic features into consideration. The second reason for using Europarl consists in the argumentative nature of the texts. In line with previous studies of discourse structure (e.g. Stede, 2004; Van der Vliet & Redeker, forthcoming), the choice of using Europarl was also based on the assumption that argumentative texts typically offer a more interesting discourse structure (understood as a wider range of different rhetorical relations) than narrative or expository text types. In particular, parliamentary discourse is often characterised by discourse moves (see Biber, Connor, & Upton, 2007) such as claim (or counter argument) followed by support for claim (or reason) (Bayley,

2004, p. 24), entailing the use of both pragmatic and semantic rhetorical relations. In a corpus-based study of the distribution of rhetorical relations, Sanders (1997, p. 138) showed that in persuasive (i.e. argumentative) text types, pragmatic rhetorical relations will occur relatively more frequently than in other non-persuasive text types. Thirdly, given the cross-linguistic approach of the thesis, the high text comparability between the languages in Europarl was essential. The comparability concerns textual features such as text genre (parliamentary debates), text type (argumentative), register (relatively formal) and contextual features such as shared topics (though they vary within a political continuum), shared location (the European Parliament) and shared profession among speakers (Members of the European Parliament, MEPs). A fourth, perhaps scientifically less valid reason for choosing Europarl as my empirical basis is that the corpus is freely available, apparently not restricted by any copyrights, that it contains complete texts, and that it is considered to be 'an invaluable resource for cross-linguistic studies' (Cartoni & Meyer, 2012, p. 2) of authentic, natural language.

3.1 The texts: genre and typology

As an empirically based study, this thesis is similar to a large and increasing body of linguistic studies whose data draw on a corpus of authentic, natural language texts often chosen to represent, in some more or less specific manner a given language or sublanguage. Unlike most of those previous studies, however, the present investigation neither seeks to develop new theories or annotation methods, nor to present new approaches to data mining or query processing. Instead the study aims at studying structural phenomena in argumentative texts across three languages. Before describing the corpus used in this study, this section shall discuss the characteristics of the sublanguage or variety of parliamentary discourse.

One important linguistic property shared by the corpus texts is that the languages employed are the standard varieties: Standard Danish (see Galberg Jacobsen & Skyum-Nielsen, 2003, p. 64ff), British English/Received Pronunciation[4] (Collins & Mees, 2003; Peters, 2004) and Standard Italian (see Loporcaro, 2013; Maiden & Parry, 1997; Marcato, 2007); thus regional varieties or social varieties have not been considered. This constitutes an immediate constraint on the present work in that anything that can be concluded about the linguistic structures will apply only to texts such as the ones that occur in the contexts of these three standard language varieties.

---

[4] Henceforth, this standard shall be referred to as English.

As mentioned above, the debates originate from the plenary sittings of the European Parliament and are by virtue of their nature argumentative, although they do contain smaller sequences of other speech acts (cf. Skytte & Korzen, 2000, p. 30ff). Parliamentarians are very explicit about the construction of arguments – similar to van Dijk's (1980)'s notion of macrostructure – which, according to Bayley (2004, p. 24), results in a very high frequency of meta-argumentative lexis containing words as *argument, issue, point*, etc. The plenary sittings are chaired by the President of the European Parliament, who typically opens the sitting, sometimes with a tribute or a speech on a current topic. During the sittings, the President calls upon speakers and ensures that the proceedings are properly conducted according to the detailed framework described in the Parliament's rules of procedure. One of the rules that appear important in this context is Rule 146 on language. It states (1) that 'all documents of Parliament shall be drawn up in the official languages'; and (2) that 'all Members shall have the right to speak in Parliament in the official language of their choice. Speeches delivered in one of the official languages shall be simultaneously interpreted into the other official languages and into any other language the Bureau may consider necessary'. This rule is without doubt more ideological than practical. As a representative of all European citizens, the Parliament describes the assembly's multilingualism as one of its most important aspects. This is important in order to guarantee the transparency and accessibility of the Parliament's work for EU citizens and it is unreasonable to require that Members are completely fluent in more than one of the languages of the Union. In addition, in its capacity as legislator, the European Parliament is obliged to guarantee that the linguistic quality of all the laws it adopts is beyond reproach in all 23 official languages. According to both the website of the Parliament, the metadata provided by the Europarl Corpus, and previous observations (De Swaan, 1999, p. 15), a high percentage of the speakers deliver the speeches in their native language. This means that it is difficult to identify potential strategies of simplification employed by the speakers to minimise misinterpretations made by the interpreters. In this regard, Wright (2007) divides the native English speakers into two groups on the basis of their awareness of linguistic clarity[5] (cf. also "Clear writing throughout Europe," 2011): one very aware and another less aware. Concerning the first group, Wright (2007, p. 153) states:

---

[5] Wright (2007) does not provide any details of Danish and Italian native speakers' linguistics awareness but groups the two languages together with other 'Nordic, Dutch and German MEPs' and 'Mediterranean and Southern MEPs' respectively.

They felt that their contribution should be to 'confine (themselves) to a standard format in order to be understood', 'to boil down their style', to observe the parliamentary rule of KISS (keep it short and sweet). They saw no place for rhetorical flourishes and extravagant style since this could hamper comprehension in a multilingual setting. In the plenary sessions, parliamentary procedure discourages this practice anyway, as speakers are allocated a precise and short time to make their point and are cut off if they run over their allotted slot. … Their linguistic behaviour in the public sessions did tend to confirm that they attempted to articulate carefully and to use a clear, plain style.

In opposition to this, the performance of the second group is characterised as (ibid.):

Their oral performance suffered from one or more of the following problems: too many metaphors, archaic idiom, colloquialisms, rambling syntactic structures, a failure to articulate clearly and a tendency to speak very fast.

This last quote may seem highly counterproductive to the analyses in this thesis since the validity of the linguistic quality and representativeness of the texts is questioned. In the selection of texts, however, the above-mentioned problems have been taken into account and incoherent and incomplete texts have been excluded – which in terms of numbers turned out to be less dramatic than expected. In any case, all texts are products of individual writers with different linguistic capacities, and the stringent revision of the proceedings, which will be discussed below, may have affected 'rambling syntactic structures' in a more positive direction.

Another important constraint on the reproduction of the speeches is Rule 181 on the verbatim report of the plenary sittings. It states that (1) 'a verbatim report of the proceedings of each sitting shall be drawn up in all the official languages'; (2) 'speakers shall return corrections to typescripts of their speeches to the Secretariat within one week'; (3) 'the verbatim report shall be published as an annex to the Official Journal of the European Union'; and (4) 'members may ask for extracts from the verbatim report to be translated at short notice'. Even though the reports, that is, the texts of the present corpus, are defined as verbatim, i.e. corresponding word for word to what has been said, and even though the original Europarl corpus contains some metadata about the speakers' names, no wholly satisfactory means can be found of knowing to what extent the verbatim reports correspond fully to the original speeches. The problem arises because each speech goes through a number of steps in becoming a verbatim report: the first is

optional and regards the preparation phase; the second is the actual delivery during the sitting; the third is the transcription conducted by the Secretariat of the Parliament; the fourth consists of the revision; and the fifth of the translation into the other languages. Since this process and the constraints mentioned above apply to all three languages analysed in this study, no further reservations will be made regarding this.

Another constraint on corpus observations is related to the nature of speeches of all kinds: it can often be difficult to ascertain their specific origin, i.e. who composed, elaborated or dictated the text since many politicians have professional speech writers (to help) write their speeches. This calls into question whether the individual speaker can also be considered the originator of the text. This constraint will also not be of high importance to the findings of this study as the study object is language use in general, and not the language use of individuals. In summary, this example from the English part of the corpus neatly captures the various constraints of the Europarl texts, as argued above.

3) Mr President, I am aware of the bad parliamentary habit of writing one's speeches in advance and then making them without listening to the rest of the debate. Indeed I have even been guilty of that myself on occasion. <ep-99-05-04.txt:159>

3.2   Data criteria

For this analysis, I decided to extract an equal number of Danish, English and Italian texts. I wanted to include only non-translated texts consisting of speeches held by parliamentarians in their native language (L1 texts). Out of the approximately 28 million words in the second version of the Europarl releases (which was the version available at the time of the data collection), approximately 1,800 Danish texts, 13,500 English texts, and 5,000 Italian texts from the period 1996-2003 were identified by language tags as L1 texts. These texts were compiled in an unbalanced subcorpus, as shown in Figure 3.1.

I wanted to have a balance between the number of speeches and the number of words. 50 texts in each of the three languages were selected as randomly as possible, which yielded a total of approximately 15,000 words per language. It was decided that 150 texts would constitute an acceptable basis for my analyses, given the detailed annotations I intended to perform. The reason for this relatively modest size of the corpus is that high level features such as discourse structure are far more labour-intensive and difficult to annotate than lower levels such as syntactic structures (Biber et al., 1998; Stede, 2011; Van der Vliet & Redeker, forthcoming (b)). Other annotation projects have used varying corpus sizes: Taboada (2004a) annotated 60

conversations with a total of 16,000 words; the RST Treebank (Carlson, Marcu, & Okurowski, 2003) 385 texts totalling some 176,000 words from the Wall Street Journal Corpus; the Penn Discourse Treebank (Prasad et al., 2008) over one million words also from the Wall Street Journal Corpus; the Potsdam Commentary Corpus (Stede, 2004) 170 texts from newspaper commentaries; Wolf & Gibson (2005) 135 texts from the Wall Street Journal Corpus; and in the Copenhagen Dependency Treebanks approximately 60,000 words were annotated in Danish, English, German, Italian and Spanish (Korzen & Buch-Kromann, 2011). Note, however, that these studies have employed different annotation techniques, had different scopes, drew on different theoretical frameworks and, most importantly, employed a varying number of trained annotators.
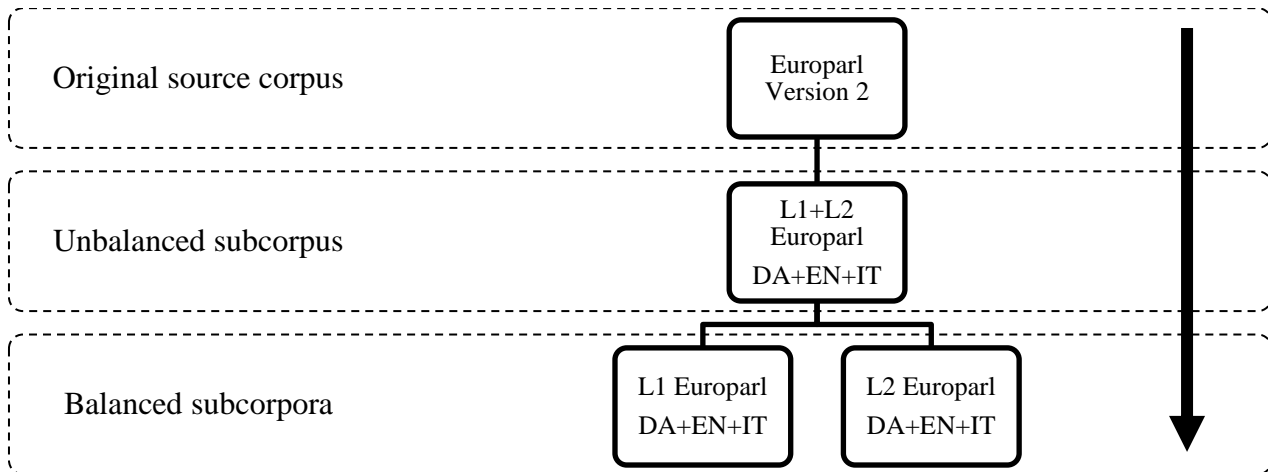
Figure 3.1: Extraction process

The 150 texts were chosen from the Europarl subcorpus following a number of textual requirements: that there should be no more than one text by the same speaker, no texts from the same date and no texts from the same debate (referred to as CHAPTER in the metadata), or at least, that there should be as much variety as possible. This was to some extent feasible in English and Italian, since there was a larger pool from which to choose. In Danish, I was forced to use more texts by the same speaker; see Appendix A for an overview.

A textual requirement was also laid down concerning speech length, so that the texts were also comparable in terms of size. The texts in the balanced subcorpus contain between 150 and 700 words. The reason for this choice of size is found in van Halteren (2008, pp. 937–938), who argues that the short texts in Europarl (<380 words) tend to be more argumentative than the long texts (>2500 words), which he characterises as presentations of written reports (i.e. descriptive

or expository text types). It may be difficult to ascertain how long these speeches were in terms of delivery, but a qualified estimate based on the following extract of an English speech from the corpus containing 268 words, and based on own read aloud experiments, would be between two and eight minutes. The first line introduced by the tag SPEAKER ID contains relevant metadata such as a specific speaker id (54), the language in which the speech was held ('EN' for English), the name of the speaker (Bushill-Matthews), and sometimes political affiliation (not indicated here). The numbers in the second line indicate year ('01' for 2001), month ('01' for January), date (15th) and repetition of SPEAKER ID (54).

<SPEAKER ID=54 LANGUAGE="EN" NAME="Bushill-Matthews">

<ep-01-01-15.txt:54>

Mr President, I wish to begin by saying that although I have been allocated four minutes, I should like to do my bit for the simplification and streamlining of bureaucracy by speaking for less than half that time.

Figure 3.2: Excerpt from a Europarl text

Gender and political affiliation are not variables in the study, and no generalisations are made with respect to these. The main reason for excluding these variables was, as mentioned above, that it is hard to determine whether all speeches were actually written by the MEP indicated in the metadata or by a professional speech writer. Nonetheless, one might argue that it could be relevant to take the speakers' gender, social background or political standpoints into consideration and investigate whether speakers belonging to either the left-wing or right-wing groups in the Parliament use sentences of a certain length or more non-finite verb forms than their opponents. But since the way of analysing discourse in the present thesis is not a Critical Discourse Analysis approach studying, for instance, the distribution of power in the Parliament, those parameters have not been taken into consideration. A quick look at the overall statistics of the texts (cf. Appendix A) also reveals that it is arguable whether such a pattern is actually present: in the Danish texts, where there are several speeches by the same speakers, the difference between the shortest and longest sentences in speeches held by the same MEP is above 100 % (e.g. the speeches by Blak and Krarup). In addition, I found no evidence that gender and political standpoint (i.e. left or right) affect text structure (e.g. differences in syntactic structures), discourse structure (e.g. the use of different rhetorical relations) or

information structure (e.g. differences in sentence lengths). Table 3.1 summarises the basic facts and data on the balanced L1 subcorpus of Europarl.

	Danish L1 texts	English L1 texts	Italian L1 texts	*Total*
Number of texts	50	50	50	*150*
Number of words	14,737	14,666	14,781	*44,184*
Number of different speakers	21	39	40	*100*

Table 3.1: Basic numbers of the balanced L1 Europarl subcorpus

As we can see from Table 3.1, the pools of Danish, English and Italian speakers are slightly different. Whereas the vast majority of the English and Italian speeches have been held by different speakers, the Danish speeches are more often held by the same speakers. The main reason for the lower number of different speakers in the Danish texts compared to the English and Italian texts has to be found in the discrepancy of the number of seats allocated in the European Parliament, and in the fact that only a few Danish speakers are frequently participating in the parliamentary debates through speeches. At the time of writing, the countries represented by the languages in this study have been assigned the following number of seats in the Parliament: Denmark 13 (1.8 % of the total number of seats in the Parliament), Ireland 12 (1.6 %), Italy 72 and United Kingdom 72 (9.8 %). This could pose a problem for generalisations based on the Danish data because the sample is not as heterogeneous as the English and Italian samples are.

In addition to the subcorpus of Danish, English and Italian L1 texts, a parallel corpus of corresponding L2 in the three languages was created. The idea behind creating this parallel subcorpus was that although most translations were very similar to their source text in terms of linguistic structures, some translators had changed the syntactic structures in the translations, rendering explicit the rhetorical relations that they had inferred between two or more discourse units. This is, for example, the case in the Danish L2 text excerpt in example 4) translated from English L1, where the underlined English relative clause has been transformed into a subordinate finite clause with discourse cue *fordi* (because), shown in bold-faced type.

4) Turning to the definition of child pornography in the Karamanou report, my group has problems with the definition <u>which includes creating the impression that the person depicted is a child</u>. <ep-01-06-11.txt:59>

*Så til definitionen af børnepornografi i Karamanou-betænkningen. Her har min gruppe problemer med definitionen, **<u>fordi</u>** <u>den indbefatter tilfælde, hvor man skaber det indtryk, at den afbildede er et barn</u>.*

Of course, the translator's reproduction of the L1 construction needs to be approached critically. But in most cases, I found the syntactic changes very useful, sometimes confirming my own analyses of rhetorical relations. The L2 subcorpus has not been annotated in any way and has only been used as potential support for my L1 annotations and analyses. The L2 subcorpus was not used in the same way as the L1 subcorpus because parallel texts are best suited for improving machine translation since they permit L1-L2 text alignment and evaluation, a matter which has been pointed out by several scholars (Baroni & Bernardini, 2005; McEnery et al., 2006). On the other hand, comparable texts (i.e. texts in different languages or varieties that deal with the same overall topic) are well-suited as the empirical basis for descriptive, and possibly cross-linguistic, comparisons. Translated texts are inappropriate because the filter of the translator and the translation strategies get in the way, and L2 texts may end up with a text structure very similar to that of the L1 (i.e. non-translated). Baroni & Bernardini (2005, p. 260) refer to this phenomenon as 'translationese', a term adopted from Gellerstam (1986):

> It is common, when reading translations, to feel that they are written in their own peculiar style. Translation scholars even speak of the language of translation as a separate 'dialect' within a language, which they call third code … or translationese ... Translationese has been originally described ... as the set of "fingerprints" that one language leaves on another when a text is translated between the two.

In the same vein, McEnery et al. (2006, p. 49) state that:

> source and translated texts … alone serve as a poor basis for cross-linguistic contrasts, because translations (i.e. L2 texts) cannot avoid the effect of translationese ... [C]omparable corpora are a useful resource for contrastive studies and translation studies, when used in combination with parallel corpora.

A further precaution to take when using parallel corpora is ensuring that one knows whether the translation has been produced directly or indirectly, that is, through another language. This is particularly important in EU texts due to the EU's system of relay or pivot languages; remember that my subcorpora contain only Europarl texts from the period 1996-2003:

> From personal discussion with a translator at the European Parliament, we know that after 2003, a pivot language was used (English), which implies that all statements were first translated into English and then into the 22 other target languages. Before 2003, however, it seems that the translations were made directly from all languages into others. (Cartoni & Meyer, 2012, p. 3)

Table 3.2 summarises the total number of words in the three parallel L2 subcorpora of Europarl.

into from	Danish L1	English L1	Italian L1
Danish L2	-	13,718	15,569
English L2	16,732	-	15,909
Italian L2	15,799	14,456	-

Table 3.2: Total numbers of words in the parallel L2 Europarl subcorpora

3.3 Terminology and translations

Some clarifications are in order here with regard to the use of terms such as 'writer', 'reader', 'words', 'speech' and 'text'. Following the terminology of one of the theories employed in this thesis, Rhetorical Structure Theory, I use 'writer' and 'reader' where this role is intended more generically covering terms like producers, writers, addressors, and recipients, readers, addressees.

'Word' is used in the same way in the three languages, although it may be argued that an Italian word does not necessarily correspond to a Danish or English word. For example, Italian is a pro-drop language, and a finite verb form such as *arrivano* contains the indication of the third person plural, *they arrive*. In the same way, Danish includes the definite article as a suffix attached to the noun: *huset*, which corresponds to *the house*. These differences mean that many reservations should be made when conducting linguistic measurements based on words. However, measurements such as mean words per sentence might serve as a first indication of significant cross-linguistic differences. Sentence length has also been used to measure

informational density in other studies (Fabricius-Hansen, 1998; Neff et al., 2004; Taft, Kacanas, Huen, & Chan, 2011; Teich, 2003).

'Speech' and 'text' are used synonymously throughout the thesis. Admittedly, the former term is associated with spoken language, whereas the latter is related to written language. But as the texts used in this study can be characterised using the theatrical terminology of Nencioni (1976) as written-to-be-spoken (*scritto-parlato* in Italian), I refer to them they as either speeches or texts.

The Danish and Italian examples have been translated into English. Generally, I use my own translations but I have indicated when the official EU translations have been employed. The translations are divided into discourse units, represented in the examples with square brackets: [EDU]. The unit markers are also included in the source texts in order to facilitate the mapping between the two languages. I have tried to preserve as many language-specific characteristics as possible, so that the correspondences are more transparent.

3.4 Annotation

All L1 texts in the three languages have been annotated with regards to text structure, discourse structure and information structure. Annotation was done in two different modes: one related to text structure; another related to discourse and information structure. In the first mode, which is usually referred to as a 'bottom-up approach' (e.g. Marcu, Romera, & Amorrortu, 1999), all texts are segmented into minor discourse units: Section 3.5.1 provides an introduction to these. As this first phase is of great importance to the present study, the principles of text segmentation shall be discussed in more detail in the following sections. Once identified, the discourse units were annotated in terms of their syntactic function (dependency) and in terms of finiteness (realisation). These two steps were relatively uncomplicated and the annotations were made in a spreadsheet.

In the second mode, the rhetorical relations between the discourse units were interpreted. This was done using the rhetorical relations as defined by Rhetorical Structure Theory, which shall be introduced in Chapter 5, see also Figure 5.1. This annotation mode was carried out using the RSTTool (O'Donnell, 2004) which allows the annotator to segment the text into discourse units and then build up the discourse structure with rhetorical relations. During this annotation mode, the hierarchy (nucleus-satellite relationship) and sources of coherence (semantic or pragmatic) were annotated. In addition, the type of linkage (intrasentential or intersentential) and signals (implicit or explicit) – both part of the information structure of a text

– were included in the annotation spreadsheet. This mode was far more labour-intensive than the former, see Section 3.4.1 and Chapter 5.

All annotations were done manually by me to ensure that the annotations followed the same set of principles and guidelines. I started with the English texts, which allowed me to compare my annotations with RST annotations by other scholars and with the numerous examples of text segmentation provided by the RST Treebank annotation guidelines (Carlson & Marcu, 2001). Later, the Danish and Italian texts were annotated in the same manner. The three languages turned out to pose different annotation challenges: the English texts had a low frequency of explicitly marked discourse cues through which the rhetorical relations between discourse units can more easily be interpreted; and the Italian texts had longer sentences and a higher number of embedded discourse units than Danish and English. All annotations have been revised three times, and five per cent of the texts were additionally annotated by another annotator and subsequently discussed. On average, the complete annotation of a text took six hours. In Figure 3.3, an annotation excerpt from an English text (<ep-02-09-23.txt:62>) shows how the different objects of analysis have been conjoined in the same spreadsheet. Annotations of rhetorical relations and signals written in capital letters indicate intersentential linkage and lower case letters intrasentential linkage. The first column (#) indicates the sequential number of the EDU in the given text; the second (*Discourse unit*) shows the linguistic material of the EDU; the third (*Textualisation*) refers to the syntactic relation between the two EDUs; the fourth column indicates how the EDU is syntactically related to one of the adjacent EDUs; the fifth column (*Rhetorical relation*) shows how the EDUs are related rhetorically to each other; the sixth column indicates the morphological realisation of the main verb of the EDU; and lastly, the seventh column (*Signal*) specifies the explicit marking of discourse cues, if signalled.

Figure 3.4 shows the RST annotation of the last four discourse units (#15-18) of the text from Figure 3.3. The annotation style is an RST tree built up recursively. The top node unit expresses the nucleus, which is the unit that contains the most important information in the text span. The lower nodes cover the satellites. Between the two lowest nodes, we see a multinuclear relation (Conjunction), meaning that the discourse units are of equal rhetorical importance.

#	Discourse unit (EDU)	Textualisation	EDU syntactically related to EDU#	Rhetorical relation	Morphological realisation	Signal
14	So for example, we can agree with paragraph 12, that immediate and general detention should be avoided.	independent sentence	12-13	VOLITIONAL RESULT	finite main verb	SO
15	We also fully support conclusion 5 of the report,	matrix clause	12	LIST	finite main verb	ALSO
16	which stresses the need for an holistic approach towards asylum and immigration	relative clause	15	volitional cause	finite main verb	-
17	looking at this as a horizontal policy objective	present participial construction	16	means	gerund	-
18	and considering all policy areas, particularly those with an external dimension, such as trade, development, environment and agriculture.	coordinate present participial construction	17	conjunction	gerund	and

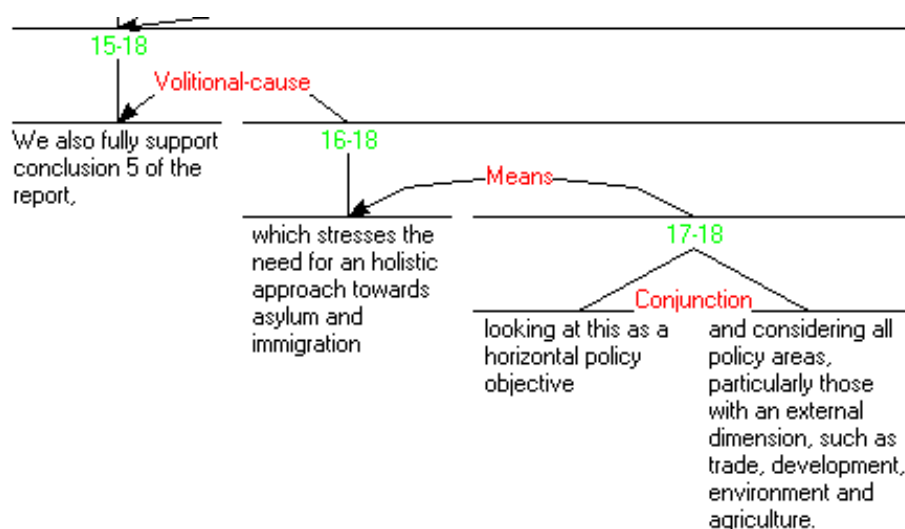**Figure 3.3: Annotation excerpt**

Figure 3.4: Annotation excerpt of a discourse structure

3.4.1 Inter-annotator agreement

The annotation of linguistic structures above the sentence is by no means an easy task. During my time as an annotator in the Copenhagen Dependency Treebanks from 2008 to 2010, I experienced how annotating low level features such as syntactic structures required much less effort than annotating high level features such as discourse structure; some of the struggles are documented in Buch-Kromann (2010). The same problems were experienced when the RST Treebank was created. Carlson et al. (2003, p. 103) note that:

> Developing corpora with these kinds of rich annotation is a labour-intensive effort. Building the RST Corpus involved more than a dozen people on a full or part time basis over a one-year time frame (Jan-Dec 2000). Annotation of a single document could take anywhere from 30 minutes to several hours, depending on the length and topic. Re-tagging a large number of documents after major enhancements to the annotation guidelines was also time consuming.

One of the main reasons for this is that syntactic roles such as subject and object are much better defined, recognised and accounted for in the literature than rhetorical relations. The extra effort needed in annotating higher level features consists not only in interpreting which relations hold between the various units of a texts but also in obtaining a certain degree of tagging consistency, often referred to as inter-annotator agreement[6] in cases where more than one annotator is involved (e.g. Cook & Bildhauer, 2011; Marcu et al., 1999).

The purpose of computing inter-annotator agreement scores is to demonstrate that the annotation guidelines can be understood and applied by people other than those who developed the coding schemes (in my case, the schemes include the EDU segmentation principles, RST trees and the inventory of RST relations). In addition, inter-annotator agreement is computed in order to ensure reproducibility of annotations. In one of the most frequently cited review articles on inter-annotator agreement, the goals of agreement studies are summarised as follows:

> Researchers who wish to use hand-coded data—that is, data in which **items** are labelled with **categories**, whether to support an empirical claim or to develop and test a computational model—need to show that such data are **reliable**. The fundamental assumption behind the methodologies discussed in this article is that

---

[6] Other places in the literature, inter-annotator agreement is also referred to as inter-coder agreement (e.g. Artstein & Poesio, 2008).

data are reliable if coders can be shown to **agree** on the categories assigned to units to an extent determined by the purposes of the study […]. If different coders produce consistently similar results, then we can infer that they have internalised a similar understanding of the annotation guidelines, and we can expect them to perform consistently under this understanding. Reliability is thus a prerequisite for demonstrating the **validity** of the coding scheme—that is, to show that the coding scheme captures the "truth" of the phenomenon being studied, in case this matters: If the annotators are not consistent then either some of them are wrong or else the annotation scheme is inappropriate for the data. (Just as in real life, the fact that witnesses to an event disagree with each other makes it difficult for third parties to know what actually happened.) However, it is important to keep in mind that achieving good agreement cannot ensure validity: Two observers of the same event may well share the same prejudice while still being objectively wrong. (Artstein & Poesio, 2008, pp. 556–557)

The literature on computing agreement scores proposes a number of ways to calculate these according to the nature of the study and to the methods and theories employed. In the analysis of RST trees as applied in this thesis, Marcu et al. (1999, p. 52) propose mapping the hierarchical RST structures into sets of units that are labelled with categorical judgments using Cohen's (1960) Kappa coefficient. This is because decisions at one level of the discourse tree affect decisions at other levels, which means that the levels are not independent of each other (Van der Vliet & Redeker, forthcoming (b)). The parameters and categorical judgments considered are:

- EDU segmentation (categories: yes or no)
- Spans (categories: yes and no)
- Nuclearity (categories: (nucleus, satellites or none)
- Relation labelling (categories: the 32 different RST relations; see Appendix B)

For the present study, approximately five per cent of the 150 texts have been selected for computing agreement scores. The fellow annotator chosen was an experienced annotator from the Copenhagen Dependency Treebanks Project (CDT), who was familiar with the RST annotation style, the RST relations and the theory in general. However, as the other annotator had done most of his annotations with the CDT relation inventory, which is quite different from the RST inventory (see Buch-Kromann, Gylling, Jelsbech Knudsen, Korzen, & Müller, 2010), I

arranged a meeting prior to the annotation task, during which the most important aspects of RST annotation and relations were presented and discussed. Moreover, I selected a number of troublesome cases from the corpus, which I anticipated the fellow annotator would also find challenging. After the meeting, the annotator was given seven texts to annotate following my guidelines. During the annotation process, a number of clarifying questions were asked about the EDU segmentation principles and the relation inventory, which I answered as generically as possible without being introduced to the specific cases.

Table 3.3 shows the overall agreement scores using the methods proposed by Marcu et al. (1999). It can be seen from the table that there are acceptable levels of agreement on all four parameters, from the highest agreement of K=0.95 on EDU segmentation to the lowest agreement of K=0.63 on relation labelling. Landis & Koch (1977) regard Kappa values between 0.6 and 0.8 as 'substantial' results, and Kappa values between 0.8 and 1.0 (the maximum) as 'perfect'. See also Artstein & Poesio (2008) and Marcu et al. (1999) for further information on how Kappa values and inter-annotator agreement scores are computed in general.

Agreement on	Kappa values
EDU segmentation	0.95
Spans	0.85
Nuclearity	0.80
Relation labelling	0.63

Table 3.3: Inter-annotator agreement

Although Carletta (1996) suggests that Kappa values between 0.67 and 0.8 allow only 'tentative conclusions', the agreement obtained on relation labelling of 0.63 must be seen in light of the results obtained by others. Considering the scale and the limits of being a one-man project, cf. also the discussion in Section 3.2, I regard the obtained agreement numbers as satisfactory. In addition, the results are also very much in line with previous studies (Buch-Kromann et al., 2010; Marcu et al., 1999; Van der Vliet, Berzlánovich, Bouma, Egg, & Redeker, 2011; Wolf & Gibson, 2005), in which the Kappa values on relation labelling are significantly lower than the other values such as EDU segmentation. Artstein & Poesio (2008, p. 580) also note that:

> The analysis of discourse structure—and especially the identification of discourse segments [EDU segmentation]—is the type of annotation that, more than any other, led C[omputational] L[inguistic] researchers to look for ways of measuring reliability and agreement, as it made them aware of the extent of disagreement on

even quite simple judgments … Subsequent research identified a number of issues with discourse structure annotation, above all the fact that segmentation, though problematic, is still much easier than marking more complex aspects of discourse structure, such as identifying the most important segments or the "rhetorical" relations between segments of different granularity. As a result, many efforts to annotate discourse structure concentrate only on segmentation.

In general, acceptable agreement scores are important as coding or annotation of data plays a crucial role in the analysis of a study. In the present cross-linguistic study, however, what is perhaps more important is that the annotations have been carried out in exactly the same way across the three languages under investigation. This allows us to conduct analyses of the annotations in Danish, English and Italian without having to worry about different annotation preferences or diverse readings of guidelines. Moreover, as I myself have annotated all 150 texts in all three languages, it has been possible for me to continuously adapt the segmentation principles to language-specific constructions and to compare textualisations of rhetorical relations in one language with textualisations in the other two languages by comparing both L1 with L1 texts and L1 with L2 texts.

Lastly, it must be mentioned that the increased focus on inter-annotator agreement scores has been criticised by a number of scholars. Buch-Kromann (2010, p. 9) warns scholars building treebanks, that is, annotated corpora, about excluding detailed descriptions of various linguistic phenomena only to achieve high agreement scores:

Measuring Treebank quality is probably one of the hardest and most important outstanding problems in the field, and any research that can address these problems even tentatively should be encouraged by the field. […] [M]ore importantly, if used as a proxy for annotation quality by treebank designers and reviewers, an exaggerated focus on agreement may lead to distortions in the way treebanks are designed.

In the same way, Reidsma & Carletta (Reidsma & Carletta, 2008) argue that even Kappa measures above 0.8 are no guarantee that the results actually are reliable. Instead of concentrating on numbers purely, scholars should look for any patterns in the disagreement among annotators and assess what impact they will have; in an RST context, the scholars behind the MTO Corpus (Van der Vliet & Redeker, forthcoming (b)) propose a reconciliation of

different annotations as a possible solution. Figure 3.4 gives examples of each parameter considered in Table 3.3 where there was disagreement between me and the other annotator. In the first example of EDU segmentation, the disagreement relates to whether the interrogative sentence should be segmented into one or three EDUs. In the second example, we can see how differently the first three sentences of a text can be annotated in terms of spans, nuclearity and relation labelling, simply because the relation between two EDUs (#2+#4/6) has been interpreted differently (Concession versus Contrast).

	EDU segmentation
My anno-tation	[Hvilke regler gælder,] [og hvilke rettigheder har vi,] [hvis noget går galt?] *What rules apply, and what rights do we have if something goes wrong? <ep-99-05-06.txt:42>*
Fellow annota-tor	[Hvilke regler gælder, og hvilke rettigheder har vi, hvis noget går galt?]
	Spans, nuclearity and relation labelling
My anno-tation	 *[Mr President, International Consumers' Day was celebrated in March, with the theme of electronic commerce and consumer protection.] [Commerce on the Internet does not in itself raise new problems of consumer policy,] [but [since we are talking about a new medium,] there is a need to establish security and confidence.] [Commerce on digital networks should be at least as secure and safe as commerce in the physical world.] <ep-99-05-06.txt:42>*

Fellow annota-tor	 Background 1-5 Handel via digitale net bør være mindst lige så tryg og sikker som handel i den fysiske verden. Elaboration Hr. formand, den internationale forbrugerdag i marts blev markeret under temaet elektronisk handel og forbrugerbeskyttelse. 2-5 Contrast Handlen på Internettet rejser ikke i sig selv nye forbrugerpolitiske problemstillinger, 3-5 Same-unit men 4-5 Nonvolitional-cause da der er tale om et nyt medie, er der behov for, at der etableres tryghed og tillid.

Table 3.4: Disagreement examples

The question to be raised here is really whether a large number of texts can be annotated by two persons in exactly the same manner. Admittedly, some interpretations are more correct than others, but since annotators are asked the extremely difficult task of having to infer what type of rhetorical relations the writer intended to express between two units, I believe that if one were to achieve a Kappa agreement score on relation labelling of 1.00, something would be scientifically inaccurate simply because texts are ambiguous. Or as argued by Buch-Kromann (2010, p. 10) based on the experience from the annotations in CDT:

> The experience of the CDT annotators, and many others in the field, is that semantic distinctions are really hard to make, and that disagreements are often caused by truly ambiguous texts where the two differing analyses either lead to essentially the same meaning, or the context does not contain sufficient information about the speaker's true intentions. But that does not necessarily imply that the distinction does not encode important information, it is just noisy information.

That being said, the agreement numbers of my annotations are acceptable, and no further reservations will be made in this regard.

3.5 Text segmentation

To understand how a text is composed presupposes an analytical approach in which the smallest or minimal units of a text are individuated. Once individuated, these units can be examined in terms of distribution and internal relationships (Jansen, 2003, p. 67). The units of interest in a study of the present kind are discourse units, which can best be characterised in terms of syntax, but also in terms of semantics. Discourse units constitute the informational basis of a text in

such a way that the sum of the discourse units can be said to form a coherent text. Obviously, these characteristics do not tell us much about the formal criteria for individuating one discourse unit from another, and a survey of the literature on the segmentation of discourse units quickly reveals that discourse units are difficult to define without having to make concessions and exceptions, and even harder to individuate automatically for computers. However, some of the most recent and promising attempts to define these units have actually been performed within computational linguistics.

According to Irmer (2011, p. 128), every theory of discourse structure has to account for the basic structuring units. The basic units of a text, also referred to as the minimal unit of analysis (Taboada & Zabala, 2008, p. 65), are usually defined in accordance with the object under investigation. In studies of spoken language, the sentence is not necessarily regarded as the optimal unit of analysis (Schiffrin, 1987, p. 32), though in other studies it is considered appropriate (Polanyi, 1988). In an overall attempt to categorise the different types of units, Mosegaard Hansen (1998, pp. 113–128) identifies three different types.

- form-based units (sentence, clause, turn, tone unit and utterance)
- content-based unit (proposition)
- action-based units (speech act and communicative act)

The first group of form-based units has often been associated with theories and projects investigating written discourse (e.g. Givón, 1983; Grimes, 1975; Longacre, 1983). Within this group, the sentence has been defined as a main or matrix clause with all its modifiers, including subordinate clauses, though not independent coordinate clauses which are treated as two units, cf. the examples in Table 3.5 below. Also when dealing with spoken data, form-based units have been used to determine turn of sentential, clausal, phrasal and lexical constructions (e.g. Sacks, Schegloff, & Jefferson, 1974). The advantage of using this type of discourse unit is that segmentation can be carried out relatively consistently since it is based on well-established grammatical and syntactic categories. The disadvantage, however, is that the segmentation may not be fine-grained enough to catch all units of interest in the particular study.

The second group of content-based units includes only propositions, a notion which according to van Dijk & Kintsch (1983, p. 14) can be found at the surface of a text (i.e. in clauses) stressing the close relationship between semantic propositions and clauses in this way:

In principle, we assume that there is a strategic one-to-one relationship between propositions and clauses: One clause expresses one proposition. This means, however, that our propositions must be *complex*, according to the usual models from logic or philosophy. Word meanings will usually correspond to what is called an atomic proposition.

Similarly, Renkema (2004, p. 88) denotes a proposition as the minimal unit of meaning containing a verb (the predicate) and one or more arguments that relate to it. By using propositions as the minimal discourse units, the analyst ensures that all possible units of interest are included. However, the segmentation process is complicated in that it demands a large inventory of rhetorical relations. Furthermore, as noted by Renkema (2004, p. 89), hardly any criteria can be given to test the accuracy of the analysis, a circumstance which also makes it very difficult to automate the segmentation.

The third group of action-based units, which I shall refrain from elaborating on in detail, also draws on insights from philosophy and the contributions to this field, in particular to speech act theory (Austin, 1962; Searle, 1969). A minimal unit within this group comprises units that function to express a complete thought or idea on the part of the speaker, ranging from the sentence level and below.

Mosegaard Hansen's three groups partially overlaps with previous approaches such as that of Enkvist (1985, p. 14), who suggests four different 'models': a sentence-based text model which attempts to individuate the criteria for clauses to form a text; a predication-based text model according to which a text should be seen as one particular arrangement, or textualisation, of a set of underlying elements; a cognition-based *or* cognitive text model which tries to explain where the predications and units come from, and an interactional text model which aims at explaining the motives for which the speaker has chosen to textualise the predications in the given manner.

To sum up, Mosegaard Hansen's and Enkvist's proposals display terminological differences but also some degree of consensus regarding the general idea of finding and defining the smallest unit of analysis. According to Carlson & Marcu (2001, p. 3), scholars across different theoretical stances agree that the units must be non-overlapping spans of texts. As observed by Taboada & Zabala (2008, p. 68), it is obvious 'that each group of researchers, each field that studies discourse, or even each new research project, devises a definition to suit their purposes', but whether this actually constitutes a problem for the study of discourse is doubtful. The minimal unit of discourse must, in fact, be segmented bearing in mind the purpose of the

analysis and how the applied theory has defined these units previously. Admittedly, this makes it more difficult to compare one approach to another, but the alternative would be to omit relevant units from the analysis or to include irrelevant material. In any case, the most efficient means to obtain a reliable segmentation is a detailed set of guidelines.

The units considered for my study are included in the first group of form-based units and to some extent also those of the interactional text model, as I will examine differences in the textualisations of discourse units across the three languages. Since I have chosen RST as my theoretical framework, the annotations in the present study will be based on sentence and clause-like units as described in the annotation guidelines by Carlson & Marcu (2001). Henceforth, the minimal discourse units will be referred to as Elementary Discourse Units.

3.5.1   Elementary Discourse Units

Elementary Discourse Units (EDUs) is a term proposed by Carlson & Marcu (2001) to cover the (alleged) 'minimal building blocks' of the discourse structure of a text. The purpose of dividing a text into EDUs is to be able to identify the linguistic structures of a text by individuating the various relations between these EDUs. From a textual perspective, the EDUs can be linked to other EDUs in terms of syntactic relations; from a discursive perspective, EDUs can be related to each other by rhetorical relations; and from an informational perspective, EDUs can be related to each other by packaging relations. One of the main differences between the treatment of EDUs from a discursive and a textual perspective is the fact that EDUs can relate to other EDUs rhetorically both within and across the sentence boundary, whereas, from a textual point of view, EDUs typically relate to the immediately adjacent EDU inside the same sentence.

As mentioned in the previous section, EDUs are in the terminology of Carlson & Marcu (2001) sentence and clause-like, which means that they can be assigned to form-based units. Stede (2011, p. 89) tries to define EDUs more precisely as a:

> span of text, usually a clause, but in general ranging from minimally a (nominalisation) NP to maximally a sentence. It denotes a single event or type of event, serving as a complete, distinct unit of information that the subsequent discourse may connect to. An EDU may be structurally embedded in another.

The idea behind this definition is finding a suitable balance between granularity and robust EDU identification (Carlson & Marcu, 2001, p. 3) using lexical and syntactic clues to help determine the boundaries. Another way of identifying EDUs is a more intuitive approach to whether a

rhetorical relation could hold between one unit and another segment (Tofiloski, Brooke, & Taboada, 2009, p. 78; Van der Vliet et al., 2011, pp. 160–161). Within this approach, the annotator may, if in doubt, test whether a particular sentence contains several EDUs by trying to link the candidate EDUs with one of the rhetorical relations present in the theoretical framework employed. Polanyi et al. (2004, p. 110) combine the two above-mentioned approaches in trying to define EDUs, characterising them as 'the syntactic constructions that encode a minimum unit of *meaning* and/or discourse *function* interpretable relative to a set of contexts'. However, this rather intuitive bottom-up approach requires well-trained annotators, comprehensive definitions of the rhetorical relations and exhaustive annotation guidelines in order to succeed.

As noted in Section 3.4.1, the task of producing large, consistently discourse annotated corpora is extremely difficult, partly 'because the boundary between syntax and discourse can be very blurry' (Carlson & Marcu, 2001, p. 2). In Table 3.5, this fuzzy boundary is evident in a number of different constructions, which range from distinct sentences to single clauses, all conveying essentially the same meaning, but packaged in different ways. Loock (2010) refers to this phenomenon as 'competing allostructures'. Notably, in the first example of a single sentence, the discourse structure is very hard to capture, as it only consists of a subject, a verb and an object. Thus, single-sentence constructions that express some kind of causality within their own boundaries are one of the compromises that annotators inevitably have to make when segmenting EDUs (Carlson & Marcu, 2001, p. 3). In all the other examples, however, the discourse structure exists through the rhetorical relations between the various types of sentences and clauses, and these have therefore been treated as two EDUs – here shown between square brackets.

Within the RST framework, several definitions and approaches to text segmentation exist: in the classical variety of RST (Abelen et al., 1993; Mann, Matthiessen, & Thompson, 1992; Mann & Thompson, 1987, 1988; Tofiloski et al., 2009), EDUs are defined as clauses, with the exception of subject and object clauses, complement clauses and restrictive relative clauses. In one of the more modern versions of RST (Carlson et al., 2003), the EDU inventory has been extended to comprise complements of attribution verbs and phrases that begin with strong discourse cues (e.g. *because, in spite of, according to*), as well as restrictive relative clauses, nominal postmodifiers and clauses that break up other EDUs.

Syntactic realisations	Examples
One single sentence	[E-commerce needs consumer confidence.]
Two distinct sentences	[E-commerce is going nowhere.] [We need consumer confidence.]
Two coordinate main clauses	[We do not have consumer confidence] [and (therefore) e-commerce is going nowhere.]
Matrix + finite adverbial clause	[E-commerce is going nowhere,] [unless we have consumer confidence.]
Matrix + relative clause	[E-commerce, [which has not got consumer confidence,] is going nowhere.]
Non-finite adverbial clause + matrix clause	[Without consumer confidence,] [e-commerce is going nowhere.] <ep-00-05-03.txt:120>

Table 3.5: Competing allostructures

A number of more recent proposals (Grabski & Stede, 2006; Lüngen, Puskàs, Bärenfänger, Hilbert, & Lobin, 2006; Van der Vliet et al., 2011) suggest variations on the classical and modern approaches, arguing for the inclusion or exclusion of certain syntactic constructions in order to obtain either a high inter-annotator agreement or a less constrained definition of EDUs. Common to most approaches is that they are based on syntactic constructions present in English grammar, which for some languages has shown to be sufficient to cover constructions in other languages like German (Grabski & Stede, 2006) and Dutch (Van der Vliet et al., 2011). This could, however, be more troublesome when it comes to languages that are considered to be typologically more distant from English such as Arabic, Japanese or Chinese. Since we are comparing Danish, English and Italian in this thesis – the first two Germanic languages and the last a Romance language – language typological differences need to be taken into consideration when establishing the segmentation principles. For instance, as noted by Korzen and colleagues (Korzen & Gylling, 2012a, p. 39; Skytte & Korzen, 2000, p. 41), the Italian gerund construction does not exist in Danish, and it is not completely identical to the English –*ing* form (Pierini, 2004, p. 124). For these reasons, a total adoption of already existing text segmentation principles within the RST framework would perhaps not capture all possibly relevant aspects of the linguistic structures of a text; thus, an additional modified set of a text segmentation principles is needed.

3.5.2 Text segmentation principles

As mentioned above, there are a number of syntactic constructions that have to be dealt with in a more detailed manner in order to exclude and include relevant material, as the 'task of defining the criteria for EDUs in a detailed manner is inevitable…' (Stede, 2011, p. 88). In the following, I shall discuss the reasons for excluding and including these constructions in the definition of EDUs, accompanied by examples from my corpus in the three languages under investigation.

I shall use mainly English examples where there are no notable cross-linguistic differences between Danish, English and Italian, but for constructions not present in the English part of the corpus, either Danish or Italian examples shall be selected. EDUs are again marked in square brackets; the source of the example (<ep-#>) is shown at the end of the example; and relevant lexical material is underlined. Examples in Danish and Italian are translated into English; these translations are reproduced in italics immediately below the source text.

The section also includes a number of troublesome syntactic phenomena inspired by those of Carlson & Marcu (2001), Mann & Thompson (1988) and Stede (2011). The segmentation used in the present study is fairly fine-grained; it is guided by the already existing rhetorical relations (the extended RST inventory of rhetorical relations as found on the RST website, see also Chapter 5 and Appendix B), and it is based on a combination of syntactic and semantic segmentation criteria as well as punctuation.

3.5.2.1 Sentences

Sentences are understood as a group of words expressing a complete event or situation. They contain a subject and predicate and must have a finite verb (Verspoor & Sauter, 2000, p. 170). Simple sentences containing a single clause are considered to be EDUs, as seen in the following example 5) of a sentence consisting of a subject (*I*), a verb (*welcome*) and a direct object (*this debate*).

5) [I welcome this debate.] <ep-00-11-13.txt:41>

Inside sentences, clauses that function as complements of a verb are not considered to be EDUs. This holds for finite and non-finite complement clauses like the ones below, where an infinitive (*essere poveri*/*to be poor*) functions as the subject of the clause in the Italian example 6), and an English *–ing* participial phrase (*allowing…*) functions as the direct object in example 7).

6) ...[essere poveri è sempre difficile,]... <ep-02-04-11.txt:43>
    ...[to be poor is always difficult,]...

7) ...[which effectively means <u>allowing the Spanish fleet into the rest of British waters,</u>]...
   <ep-02-05-29.txt:37>

Similarly, constructions of the type *I will try to wait for you* are also treated as a single EDU, because *to wait* is considered to be part of the valency of the main verb, here *try*.

In contrast to Carlson & Marcu (2001), complements of attribution verbs (e.g. speech acts and other cognitive acts) are not treated as EDUs. Marcu & Carlson include complements of attribution verbs as they state that a rhetorical relation can be interpreted between a matrix clause containing an attribution verb and a complement clause. They label this relation 'Attribution'. Treating complements of attribution verbs as EDUs, however, is not unproblematic (Dinesh et al., 2005; Hardt, 2013), and it remains one of the main differences between the various versions of RST relations found in the literature. One of the reasons for not including complements of attribution verbs in the present thesis is that attribution does not constitute a 'real' rhetorical relation since there is no regular semantic or pragmatic relationship between the attribution complement and the matrix clause; this viewpoint is, among others, also found in Stede (2011, p. 90). On the other hand, a reason for including attribution as a rhetorical relation is that complements of attribution verbs play an important role, in terms of contents, in text summarisation as applied to various tasks related to information retrieval and artificial intelligence (cf. Carlson et al., 2003; Mann & Thompson, 1987; Webber, Egg, & Kordoni, 2011). In examples 8) and 9), which contain verbs of attribution (*hope* and *say*), we can see that no segmentation has been applied to separate the complements from the matrix clauses.

8) [<u>We hope that</u> the Member States will continue to monitor that.] <ep-00-10-02.txt:84>

9) [Mr President, <u>I should like to say that</u> I will be broadly supporting the Dell'Alba report]... <ep-98-06-15.txt:42>

3.5.2.2 Coordinated clauses
Clauses can be either coordinated with other clauses or subordinated to other clauses. The characteristics of the former are the same as those of the sentence plus that a main (not matrix) clause can stand by itself or be part of a sentence, see example 10). However, in this study, I also include elliptical coordinated clauses, where two or more clauses share the same verb or share a noun phrase as their common subject. This goes against Carlson & Marcu (2001). An

example of this is found in 11) where the two clauses share the same verb phrase (*has announced the deployment of*), which has been omitted in the last of the two clauses. The reason for including these is that the two clauses can be seen as two distinct predicates. Example 12) illustrates how a subject (*My group*) can also be omitted in the second clause. An exception to this principle is found in 13) where a serial verb construction of two coordinated verbs (*go out and sell*) does not yield two distinct predicates and therefore has not been segmented into two EDUs: in instances of this kind, the first verb usually expresses only the manner in which the action of the second verb is being carried out.

10) [In my view we are now prepared to resort to extreme measures] [and I accept Commissioner Byrne's view that in the present situation they are the minimum necessary.] <ep-01-02-01.txt:12>

11) [Britain has already announced the deployment of some 30,000 troops] [and the United States 130,000 troops]... <ep-03-01-29.txt:28>

12) [My Group will therefore support the common position] [and looks forward to the enactment of the legislation]... <ep-00-01-17.txt:40>

13) [We have to go out and sell the case for enlargement to public opinion in all the Member States of the European Union.] <ep-00-09-04.txt:46>

Coordination is also found between most types of subordinate clauses, which are therefore segmented as separate EDUs. In example 14), the second underlined subordinate finite adverbial clause (*og såfremt man vil/and if we wish*) is coordinated with the first subordinate finite adverbial clause (*såfremt koncentrationen/if concentrations*), resulting in two separate EDUs for the reasons stated above.

14) [Vedtagelsen af denne betænkning vil få alvorlige konsekvenser for landbrugserhvervet i EU,] [såfremt koncentrationen af f.eks. fosfor og kvælstof ikke må overskride baggrundsværdierne for disse stoffer i vandmiljøet,] [og såfremt man vil fastsætte en grænseværdi på 0 for pesticider i vandmiljøet.] <ep-00-02-15.txt:141>

[Adopting this report will have serious consequences for agriculture in the EU] [if concentrations of, for example, phosphorus and nitrogen cannot exceed the background

values for these substances in the aquatic environment] *[and if we wish to establish a limit of zero for pesticides in the aquatic environment.]*

3.5.2.3 Subordinate finite clauses

There is a general consensus in the literature that subordinate clauses can be divided into three main classes across various languages: nominal, adverbial and relative clauses (Gast & Diessel, 2012, p. 1; Mortensen, 2013, pp. 64–65). In this thesis, I will exclusively focus on the last two classes. Subordinate clauses are defined according to their syntactic relationship to the matrix clause, which can be described as more or less dependent or integrated. In the following, I consider subordinate clauses from a verbal perspective, focusing on subordinate finite or non-finite clauses. The finite subordinate clauses are characterised by some of the same features as a sentence, but they typically start with a subordinator; they cannot stand on their own, and they function as a clause or phrase constituent (Verspoor & Sauter, 2000, p. 170). The main types of subordinate finite constructions are finite adverbial clauses, as shown in 15) and 16), and relative clauses (both restrictive and appositive), shown in 17)-19).

15) [When the European finance ministers meet next week,] [we need action on jobs.] <ep-96-07-03.txt:75>

16) [We still have a lot of work to do in this area] [as recent events have proved.] <ep-00-01-17.txt:40>

17) [Madam President, first of all I should like to thank Mr Koch for his report] [which has, at its heart, the issue of transport safety.] <ep-00-01-17.txt:40>

18) [We are faced with criminals] [who will use every possible means to avoid detection.] <ep-00-11-13.txt:41>

19) [Mr President, it is chilling to see how the preparations to invade Iraq have gathered pace,] [despite the extra time] [given to the weapons inspectors] [which, of course, we all welcome.] <ep-03-01-29.txt:28>

One of the reasons for including restrictive relative clauses, which goes against the classical version of RST and a number of other approaches (Tofiloski et al., 2009; Van der Vliet et al., 2011), is, as argued by Gylling (2012a), that these are able to express rhetorical relations other

than Elaboration. The inclusion of restrictive relative clauses is also supported by Polanyi et al. (2004). In the first example in 20), the restrictive relative clause – here embedded – could be interpreted as expressing a conditional relationship to the matrix clause (paraphrased *If they are innovative, they speak English*). In example 21), the relative clause – here a so-called echo apposition repeating the noun *commissione/committee* – expresses a purpose relationship to the matrix clause (paraphrased *The regional government has established a committee **in order to** ascertain to which extent these texts are influenced by Marxism*).

20) [Alle, [der er innovative,] taler engelsk] ... <ep-02-11-07.txt:11>

   [Everyone [who is innovative] speaks English] ...

21) [Questo ha infatti costituito una commissione per il controllo dei libri di testo di storia nelle scuole, commissione] [che deve controllare fino a che punto [...] questi testi sono inficiati di marxismo.] <ep-00-11-13.txt:31>

   [The regional government has established a committee to monitor the history textbooks used in schools, a committee] [which is to ascertain to which extent [...] these texts are influenced by Marxism.]

Another reason for including restrictive relative clauses is that, in practice, distinguishing between restrictive and appositive relative clauses can be a difficult task (see Bache & Kvistgaard Jakobsen, 1980).

Appositive relative clauses are to a larger extent capable of expressing different rhetorical relations, and should, therefore, always be included in the EDU segmentation of a text, see Loock (2010), who, from an information packaging perspective, shows a number of 'competing allostructures' (e.g. sentential parentheticals, juxtaposed/coordinated independent clauses, adverbials or noun modifiers) that essentially convey the same information as that of the appositive relative clause.

However, following Carlson & Marcu (2001), syntactic focusing devices such as clefts, pseudo-clefts and extrapositions are not treated as EDUs, as these constructions purely provide emphasis on a particular element in the sentence without expressing any rhetorical relation. In 22), we can see how a cleft construction is not segmented; in 23), how a pseudo-cleft construction is also not segmented; and in 24), how an extraposition (Carlson & Marcu, 2001, p. 15f) is not segmented as a distinct EDU either.

22) [Det var 1996-regnskabet, der fældede den tidligere Kommission,]... <ep-00-01-18.txt:255>

[It was the 1996 accounts which led to the downfall of the previous Commission,]...

23) [What fellow Members may not know is that the main source of the disease was contaminated paté] [sold from three Belgian plants.] <ep-99-07-22.txt:25>

24) [It is important to sort out the structural problems in our economies] [but this will take time.] <ep-96-07-03.txt:75>

3.5.2.4 *Subordinate non-finite clauses*

Subordinate non-finite clauses contain a predicate (the subject is normally implicit), do not have a finite verb, may or may not start with a subordinating adverb or a preposition, and cannot stand on their own, but function as a clause or phrase constituent (Verspoor & Sauter, 2000, p. 170). The main types of subordinate non-finite constructions are non-finite adverbial clauses, adjectival modifiers (not treated as EDUs by Carlson & Marcu (2001)), present and past participial clauses and gerunds/English –*ing* clauses. In 25), it can be seen how two infinitive clauses (*in order to isolate… and to prevent…*) have been segmented as EDUs.

25) [We will be asking for a split vote] [in order to **isolate** these words,] [to **prevent** such an inference.] <ep-02-07-01.txt:103>

Infinitival clauses as in 26) are also segmented as EDUs, provided that the clause is not a complement to the verb as in 27), see Carlson & Marcu's (2001, pp. 6–7) discussion on 'infinitival clauses' versus 'infinitival complements'.

26) [Now they are coming back to translate that into our Rules of Procedure] [to **ensure** that our work is transparent.] <ep-01-11-12.txt:56>

27) [I therefore ask the Commissioner to **take** that on board.] <ep-97-09-16.txt:102>

Example 28) shows that English –*ing* clauses (*Despite often being seen…*) along with the Italian gerund in 29) (*votando favorelmente*) are treated as individual EDUs even in cases where the participial clause is not marked with a subordinator and regardless of whether the clause is placed before or after the matrix clause.

28) [Despite often **being** seen as a local issue] [there are huge costs to European society and to business of stress-related illnesses] [to which noise is a big contributory factor.] <ep-97-06-09.txt:99>

29) [Il nostro gruppo sostiene questo impegno] [**votando** favorevolmente,]... <ep-98-05-02.txt:62>

*[Our group supports this commitment,] [**voting in favour,**]...*

In 30), we observe how an adjectival postmodifier (*now present…*), also referred to as reduced relative clauses with an adjective (Carlson & Marcu, 2001, p. 21), has been individuated as an EDU.

30) [Indeed, according to estimates, carbon dioxide accounts for 50 % of the greenhouse gases] [now **present** in the atmosphere.] <ep-96-04-15.txt:28>

In 31) and 32), two examples of an English –*ing* construction (*facing* and *containing*) have been treated as EDUs. However, –*ing* forms that in dictionaries are classified as prepositions such as *including, regarding, concerning* etc. have not been treated as separate EDUs in any of the three languages under consideration.

31) ...[and this very week convenes a first meeting of local women's groups on questions] [**facing** minority ethnic women.] <ep-98-01-28.txt:80>

32) [Tax incentives must be used to encourage the early introduction of vehicles] [**containing** advanced anti-pollution equipment.] <ep-98-09-15.txt:49>

Examples 33) and 34) display similar usages of past participial modifiers (*caused* and *committed*) to those shown above. Again, these are segmented as distinct EDUs.

33) [My constituents in London also suffer the problems of sleep disturbance] [**caused** by early morning flights at Heathrow Airport]... <ep-97-06-09.txt:99>

34) [The Western European Union is a military alliance] [**committed** to the nuclear deterrent.] <ep-98-05-13.txt:201>

A phrase is a group of related words that typically does not express a whole event or situation, it does not have both a subject and a predicate, and it cannot stand on its own, but functions as a clause or phrase constituent (Verspoor & Sauter, 2000, p. 170). In a phrase, the verbal element is normally omitted or strongly deverbalised in the form of a nominalisation. Phrasal constructions have been handled in different ways in the literature. Tofiloski et al. (2009) do not regard these as EDUs, whereas others do (e.g. Carlson & Marcu, 2001; Grabski & Stede, 2006; Van der Vliet et al., 2011). In this study, I will follow the latter group by considering prepositional phrases as separate EDUs on the grounds of an overtly-signalled rhetorical relation through discourse cues. For instance, in 35), the Italian phrase starting with the subordinator *anche se (even if)* is segmented as an EDU. In examples 36)-38), similar cases are exemplified with the discourse cues *because of* in 36), *despite* in 37), and *without* in 38). Note that these belong to different syntactic classes: *anche se (even if)* and *because of* are subordinators; *despite* and *without* prepositions.

35) [...e debbo dire anche che il ruolo del Parlamento è aumentato rispetto a Amsterdam,] [anche se con minore progresso rispetto a quanto...] <ep-99-12-14.txt:40>

[...and I also have to say that Parliament's role has been increased, with respect to Amsterdam,] [even if with less progress compared to what...]

36) [We did, however, vote against the report in committee] [because of a number of outstanding concerns,]... <ep-02-07-01.txt:103>

37) [Mr President, firstly I should like to pay tribute to Mrs Díez de Rivera Icaza for her passionate pursuit of the noise issue,] [despite her illness.] <ep-97-06-09.txt:99>

38) [Without consumer confidence,] [e-commerce is going nowhere.] <ep-00-05-03.txt:120>

Prepositional phrases with clausal objects that are related to the entire matrix clause are also considered to be individual EDUs, as they typically express Circumstance or Means relations; see example 39), which also contains an *–ing* form (*saying*).

39) [I would respond to the honourable Member] [by saying that we will certainly conduct further discussions with Member States.] <ep-99-01-12.txt:178>

In line with Van der Vliet et al. (2011), non-clausal appositives or nominal appositions are not recognised as separate EDUs, e.g. the construction underlined in 40). This also holds true for non-clausal reformulations, exemplifications, comparatives and postmodifiers.

40) [He is a very distinguished Danish poet, <u>one of your own countrymen,</u> Mr President, <u>Thorkild Bjørnvig,</u>]... <ep-96-10-21.txt:42>

3.5.2.6   *Punctuation*

A relevant means of determining EDU boundaries is punctuation. However, punctuation and the usage of this may also be misleading in a cross-linguistic context such as the present. For this reason, I shall explain the conventions for determining EDU segments involving various types of punctuation, following the text segmentation principles laid down by Van der Vliet et al. (2011).

First of all, full stops (41), question marks (42) and exclamation marks (43) indicate separate EDUs, unless full stops are used in abbreviations, acronyms, dates and so forth.

41) [Mr President, there is no limit on the curiosity of scientists.] <ep-97-03-11.txt:36>

42) [How best can we achieve those aims?] <ep-96-12-09.txt:33>

43) [Attenti quindi al voto di domani!] <ep-98-10-05.txt:17>
   [So all eyes will be on tomorrow's vote!]

Secondly, colons and semicolons only separate EDUs when the subsequent material is a clause such as in 44) and 45), though not if it is a non-clausal expression as in 46).

44) [<u>Io credo che sul serio dovremmo, noi tutti, considerare gli anziani un patrimonio dell'umanità:</u>] [la loro saggezza, la loro conoscenza, la storia [di cui sono viventi interpreti], <u>in realtà non sono fino in fondo valorizzate e utilizzate.</u>]
   [<u>I feel that we should all genuinely see elderly people as our human heritage:</u>] [their wisdom, their knowledge and the history [of which they are the living exponents] <u>are not fully valued or exploited.</u>]

45) [We will do so] [<u>and we will support you;</u>] [<u>but we need a commitment from the Commission that this is a real attempt to salvage something from this deal.</u>] <ep-97-09-16.txt:102>

46) ...[against that, we see: relative stability.] <ep-02-05-29.txt:37>

Thirdly, I do not consider non-clausal expressions in parentheses, brackets and quotation marks to be EDUs; however, the frequency of these constructions was close to zero in the texts investigated in this project.

And lastly, independent fragments without a verbal element ending with a full stop are considered to be EDUs.

47) [Competition rules pure and simple.] <ep-96-09-04.txt:83>

3.5.2.7 *Embedded EDUs*

An EDU may be broken up by an embedded construction such as a relative clause or a finite adverbial clause, see examples 48) and 49). Following Carlson & Marcu (2001), the pseudo-rhetorical relation of Same-unit is used in the annotation scheme to indicate the relation between the embedded unit and the main unit; see Table 3.4 and Chapter 5 for a detailed description of how this relation is employed practically.

48) [It is clearly our duty to ensure that the standards [by which they are governed or regulated] attain the best levels of corporate governance.] <ep-02-03-11.txt:79>

49) [Manufacturers say that [if this amendment goes through], they will no longer be able to manufacture these goods.] <ep-01-10-01.txt:83>

3.6 Summing up

In this chapter, I have described the main characteristics of the corpus and the annotation principles employed. One of the goals was to argue that the texts contained in the corpus should be seen as a sample of a general population of which the texts under scrutiny are representative. Another goal consisted in making the data collection and annotation process as transparent as possible to the reader. With respect to data collection, the issue is always how much the analyst can generalise from the results obtained. Given the restrictive nature of the corpus, no bold claims can be made about the characteristics of the languages in general, but it remains my belief that this is an exhaustive study of the construction of parliamentary discourse, at least as found in the European Parliament. Generalisations with respect to other genres and text types are speculative at this point, although a number of previous studies confirm some of the observations; see Section 2.2. I shall return to this later in the thesis.

Concerning the second element, the annotation process, the literature on the annotation of linguistic structures above the sentence boundary clearly shows how much uncertainty there is. The present study is by no means an exception to this, but by presenting the scores on inter-annotator agreement together with other statistics, the reader should have some idea of which precautions to take when reading the following three chapters.

The last section of this chapter was used to describe the segmentation process of the Elementary Discourse Units (EDUs). The overall strategy was to stick as close as possible to previous studies, which proved to be somewhat difficult as none of the previous studies included Danish and Italian. As a consequence, a number of adjustments were carried out in order not to favour one language over another. Common to all of the choices described in this chapter is that certain compromises had to be made. Although I recognise that other possible solutions exist and that adopting a more coarse-grained approach to, for example, text segmentation could have made some processes easier, I am confident that the decisions made constitute an ideal operational point of departure for this cross-linguistic study, which has no comparable predecessors in the literature I am familiar with. In the next three chapters, the main body of the analyses of this study shall be presented.

4   The text structure of parliamentary discourse

The present chapter deals with the study of linguistic structures in parliamentary discourses through the analysis of text structure. The previous chapters defined the concept of linguistic structures and introduced the notion of Elementary Discourse Unit (EDU) as the minimal unit of analysis. This chapter investigates how these EDUs are related to each other in terms of syntactic relations using the coordination versus subordination distinction. I refer to this phenomenon as the dependency of EDUs. Furthermore, this chapter studies how EDUs are realised in terms of verbal finiteness, that is, either through finite or non-finite verb forms, or through verbless constructions. I refer to this phenomenon as the realisation of EDUs.

Lehmann (1988), when describing different aspects of syntactic relations, points out that it is relevant in cross-linguistic studies to separate various syntactic parameters of clause linkage from semantic-pragmatic parameters because different grammatical types – what I call EDUs – can express a wide range of different rhetorical relations. That is, there is no one-to-one relationship between the syntactic relations of EDUs and the rhetorical relations between them.

> While [the semantic nature of the relation between two clauses] has always played a prominent role in the classification of subordinate clauses, it does not appear to be constitutive of cross-linguistically valid types of clause linkage. That is, there is no cross-linguistic notion of, say, the concessive clause which would possess any constant structural correlates. It rather appears that the grammatical types … cut across the semantically different clause linkage relations. (Lehmann, 1988, p. 183)

My first step in analysing the linguistic structures of texts across languages therefore focuses on the syntactic characteristics of EDUs rather than on the rhetorical relations between them. In Chapter 5, I compare the textualisations of EDUs with the type of rhetorical relations they establish. The main cross-linguistic results regarding text structure are described in Section 4.6. Before presenting the results of the analysis, I provide, in Section 4.1, a brief account of how the dependencies and realisations of EDUs have been applied to the analysis of text structure, followed by a complete description of a so-called deverbalisation scale which shall be used for analysing text structure, in Section 4.2.

Section 4.3 reviews some of the previous cross-linguistic approaches to text structure in Danish, English and Italian in different text types. The section concludes with a research hypothesis on similarities and differences in the text structure for each language. This is followed by two sections on results, testing whether the hypothesis can be confirmed or not.

4.1   Dependency and realisation of EDUs

In this section, the two main aspects of text structure under investigation shall be defined, namely dependency and realisation. At the same time, a review is provided of the most important literature on the two aspects. Bearing in mind the six dimensions of the cognitive-functional communication model presented in Chapter 2 regarding linguistic structures and cognitive dimensions, this chapter shall demonstrate how the text structure dimension is not only deeply intertwined with that of syntax but also closely related to the lower level of morphology and to the higher level of context (see Figure 2.1). Both dependency and realisation are traditional grammatical parameters which have been described in the literature on syntax.

The first parameter shall be referred to as dependency to account for EDUs that are syntactically coordinated with other EDUs and for EDUs that are syntactically subordinated to other EDUs. Coordination and subordination are well-established grammatical terms, and:

> [I]n what is probably their most widespread application, 'subordination' and 'coordination' – along with their adjectival cognates 'subordinate', 'coordinate', etc. – are *syntactic notions* denoting relations between parts of a complex syntactic unit. That is, they concern the structure of sentences or clauses and their parts. (Fabricius-Hansen & Ramm, 2008, p. 2).

An example of two coordinated EDUs is found in 50), and a subordinate EDU is found in the second EDU in 51), both underlined. In both examples, the coordinator (*but*) and subordinator (*as*) have been highlighted in bold-faced type.

50) [Mr President, I am firmly of the belief that people have a right to smoke,] [**but** my personal taste is that they should do so only in the privacy of their own homes.] <ep-00-12-11.txt:42>

51) [Mr President, this is not reform,] [**as** the Commission itself points out.] <ep-02-05-29.txt:37>

In the first example, the two EDUs are coordinated by the coordinator *but*, which usually expresses some kind of contrast in terms of rhetorical relation. The two EDUs are also sentence-like in that they are grammatically independent of each other and can easily stand alone, that is, the comma after *smoke* can be replaced by a full stop, resulting in two independent sentences. In

the second example, the second EDU is subordinated to the first EDU, a matrix clause *this is not reform*, due to the presence of a subordinator *as*, and cannot stand on its own. As a rule of thumb, it is the connectives – coordinators or subordinators – that determine the dependency between EDUs placed inside the same sentence. But in cases where there is no explicit connective, it is the morphological realisation of the main verb in the subordinate EDUs that determines the dependency. This is the case in example 52) where a matrix clause is followed by a subordinate construction with two English *–ing* forms (*causing* and *fuelling*). It is also the case in example 53) where two coordinated main clauses are separated by a semicolon expressing a juxtaposition of the two EDUs (cf. Ferrari & Zampese, 2000, p. 236). In EU texts, the recommended usage of colons and semicolons between EDUs advises that colon should be used to separate subordinate EDUs and semicolons to separate coordinate EDUs (see *English Style Guide. A handbook for authors and translators in the European Commission*, 2012, p. 15). However, the results of the next chapter suggest that this advice is not always followed.

52) [It would make the world a much more dangerous place,] [**causing** more resentment] [and **fuelling** more terrorism.] <ep-03-01-29.txt:28>

53) [Relative stability is based on track record;] [equal access is based on fleet tonnages.] <ep-02-05-29.txt:37>

The concept of dependency has mainly been examined in the field of syntax, but in discourse linguistics, the notion of coordination has also attracted much attention. Some scholars believe that syntactic coordination equals rhetorical coordination, that is, dependency across sentences (e.g. Asher & Vieu, 2005; Tomlin, 1985; Txurruka, 2000), whereas others claim the opposite, that is, syntactic coordination does not necessarily correspond to rhetorical coordination (Blühdorn, 2008; Haspelmath, 2004, p. 3f; Ramm & Fabricius-Hansen, 2005; Skytte & Korzen, 2000). This also holds true for syntactic versus rhetorical subordination.

> Undoubtedly, the idea to link syntactic subordination to discourse nuclearity [i.e. rhetorical subordination] has intuitive appeal, and moreover it is not difficult to accumulate evidence in support of such a link: Quite often, a syntactically subordinate clause is in fact 'less central to the writer's purposes' than the matrix clause is. But on the other hand, one can also gather evidence for the opposite position – the writer's purposes running against syntactic subordination, which

altogether suggests that matters are more complicated … (Stede, 2008a, pp. 47–48)

The disagreement on syntactic versus rhetorical coordination and subordination probably arises as a result of the stereotypical and vague coordinator *and* together with its counterparts in other languages. In an attempt to separate the two kinds of linguistic dependencies from each other, Lehmann (1988, p. 182) refers to syntactic parataxis and hypotaxis versus rhetorical or discourse coordination and subordination. In this thesis, however, I apply only the coordination versus subordination distinction, always specifying the syntactic or rhetorical nature of this. In general terms, I shall refer to syntactic dependency and rhetorical hierarchy.

Accordingly, syntactic coordination is considered a syntactic relation between two EDUs of the same syntactic construction, which may consist of all the EDU categories described in Chapter 3. Syntactic subordination, on the other hand, is used to refer to EDUs that are syntactically dependent on another EDU but not vice versa (Ferrari & Zampese, 2000, p. 141). Again, I refer to the types of subordinate EDUs presented in Chapter 3. However, such a coarse distinction is by no means unproblematic. In fact, one of the other categorisation issues in this area of grammar is that of treating coordination and subordination as two independent groups or as two poles on the same scale. Cristofaro (2003, p. 15ff) is one of the more recent scholars who suggests that no sharp binary distinction between coordination and subordination should be drawn. In fact, coordination and subordination are the two extremes of a gradient where many coordinated constructions show various degrees of subordination (cf. also Cosme, 2008, p. 109f; Johannessen, 1998, p. 237ff; Kortmann, 1996, p. 56ff; Quirk, Greenbaum, Leech, & Svartvik, 1985, p. 927f).

Lehmann (1988, p. 182) describes coordination as involving the linking of EDUs of the same syntactic construction. The same viewpoint is taken by Quirk et al. (1985, p. 918), who specify that in coordination 'the units are constituents at the same level of constituent structure, whereas in subordination they form a hierarchy', with the subordinate EDU being a constituent of the superordinate EDU. There are various ways of distinguishing between and identifying coordination and subordination, a common one being by means of the explicit indicators of both types of linking constructions, i.e. coordinators and subordinators respectively. Another indicator, which holds at least for Danish, is word order: in most subordinate clauses, a negation must be placed between the subject and the verb and not between the verb and the object in sentences and other types of clauses (Heltoft & Hansen, 2011, p. 314).

The main difference between coordination and subordination, then, is a matter of the degree of syntactic dependency (see Givón, 2001, p. 327). Below, in 54), we see an example of how two coordinated EDUs are syntactically independent of each other, each having their own subject and main verb. The second EDU in 55), by comparison, exhibits a higher degree of dependency, because it depends on the first EDU for the expression of one argument, namely the subject, and lacks illocutionary force of its own. The same phenomenon is apparent in non-finite subordinate constructions like the one found in 57), which exhibits a higher degree of dependency than the subordinate finite construction in 56), since it lacks tense and an illocutionary force of its own. Lehmann (1988, p. 184ff) refers to this phenomenon as the 'hierarchical downgrading' in which subordinate EDUs integrated into the matrix EDU are found at one end of the downgrading continuum.

54) [The plane arrived late in London,] [**and** John went straight home.]

55) [John arrived late in London] [**and** went straight home.]

56) [John went straight home,] [**because** he arrived late in London.]

57) [**Having arrived** late in London,] [John went straight home.]

In most cases of coordination, it is also possible to add another connective to the coordinator. In example 58), the content from 55) has been added a connective (*therefore*) after the coordinator *and*. In this way, there is a more explicit highlighting of the consequential rhetorical relation between the two EDUs than could be seen 55).

58) [John arrived late in London] [and **therefore** went straight home.]

The notion of realisation has already been introduced in the previous chapters, but it has not been properly defined yet. Realisation, in the context of this thesis, is mainly related to subordinate EDUs and to whether these contain a finite verb form, a non-finite verb form or no explicit verb form at all. The non-finite verb forms, including nominalisations, do not mark tense, mood or aspect at all (Korzen, 1998, pp. 68–69). This is shown in the following examples; in 59) by means of an English *–ing* form (*having arrived*) and in 60) by a nominalisation (*arrival*).

59) [**Having arrived** late in London,] [John went straight home.]

60) [After **his late arrival** in London,] [John went straight home.]

This lack of subject marking in non-finite constructions generally entails an inherent subject and *topic* continuity, which means that the situation or event in question is processed and interpreted as related to the on-going topic but as less important than the situation or event of the matrix EDU, realised with a finite verb. Therefore, the rhetorical relation between an EDU in which the main verb is non-finite is entirely dependent on the matrix EDU, and such structures express a particularly strong integration of the EDU in question into the matrix EDU, as also stated by Lehmann (1988, p. 214):

> [A]dvanced hierarchical downgrading of the subordinate clause implies a low syntactic level for it. We will thus be justified if in the following we take advanced downgrading as a sufficient condition for high integration. High integration of the subordinate into the main clause correlates positively with its desententialisation.

Lastly, it is also possible to find an EDU expressed as a subordinate clause or phrase in which the verb has been omitted. Here, the verb is completely absent, but can, nevertheless, be inferred from the co-text of the adjacent EDU: [Once *{John arrived}* in London], [he went straight home.].

61) [Once in London,] [John went straight home.]

The example above also shows how subordinate EDUs can be placed before the matrix EDU. This is not the case in coordinated EDUs, except when these are of the same syntactic class and not linked by any discourse cue; in this case, the EDUs may shift position without altering the meaning (Ferrari & Zampese, 2000, p. 237).

4.2 The deverbalisation scale

As argued in the previous sections, the characteristics of the dependency and realisation of EDUs can be combined and summarised in a so-called deverbalisation scale, see Figure 4.1. The scale will serve as the analytical framework for studying text structure, and the idea behind the scale is to show how various degrees of deverbalisation can be illustrated hierarchically. The scale is based on a number of similar proposals (Foley & van Valin, 1984, p. 267; Givón, 2001;

Hopper & Thompson, 1984; Korzen & Gylling, 2012a; Korzen, 2009a; Lehmann, 1988) and on the different categories of EDUs presented in Chapter 3. Cosme (2008, p. 110) proposes a scale of EDU dependency in a similar fashion for Dutch, English and French. The scale below includes only those types of EDUs found in the corpus employed in this study. Other possible constructions such as free predicatives and absolute constructions, which were not found in my corpus, have been omitted from the present scale. Ideally, the examples of the different textualisations should express the same semantic content, but as this proved to be rather difficult without jeopardising the clarity of the examples, the content of the EDUs has been altered slightly throughout the scale.

The arrow in the outer left column illustrates how the deverbalisation of the first coordinated and later subordinated EDU increases from level (a) to level (j), together with its integration into the main or matrix EDU. The underlined EDUs in the first three upper levels (a)-(c) are more or less independent of the adjacent EDU in terms of grammatico-semantic values and can express independent speech acts. Thus, verb features on these levels have previously been referred to as prototypical (Hopper & Thompson, 1984; Lehmann, 1988; Skytte & Korzen, 2000, p. 84). EDUs at level (d) are slightly more integrated into the matrix EDU and have lost independent tense, mood and probably also illocution. This also holds true for EDUs at level (e), which, in addition, typically are part of a noun phrase, if these can be described as restrictive relative clauses. EDUs at level (f)-(h) all have non-finite verb forms, thereby losing all temporal, modal and aspectual values. Furthermore, the fact that these EDUs do not normally render explicit their subject underlines their increased syntactic integration and rhetorical subordination into the matrix EDU. It has been argued by some scholars that the non-finite constructions at level (f)-(g) always contain presupposed information (e.g. Lambrecht, 1994), whereas other assert that this is not true for all languages, e.g. Italian (Skytte & Korzen, 2000, pp. 89–90). Besides sharing the characteristics of EDUs at level (f)-(g), EDUs at level (h) are syntactically integrated into a noun phrase, typically the subject or the object of the matrix EDU. This entails further constraints on the EDUs which are almost completely embedded in the matrix EDU as a second order entity, using Lyons' (1977, p. 442ff) terminology. The EDU has lost all its verbal-morphological characteristics as its valency complements are syntactically reduced to secondary positions or simply left out. EDUs at the last two levels (i)-(j) share most of the characteristics of level (h) with the exception of being related to the entire event or situation conveyed by the matrix EDU. The EDUs under scrutiny have been underlined in the scale, and the morphological realisations of the main verb highlighted in bold-faced type.

In other words: the further down on the deverbalisation scale an EDU is textualised, the fewer grammatico-semantic features are expressed by the verb, the more deverbalised it is, and the more pragmatically and rhetorically subordinated and incorporated in the matrix EDU the EDU usually is.

Another interesting aspect of the deverbalisation scale is that the further down the scale, the more obligatory it becomes for the subordinate EDU to make explicit the rhetorical relation to the matrix clause through discourse cues such as coordinators, subordinators and prepositional phrases. With the exception of level (h), all levels including subordinate EDUs, (d)-(j), are more likely to contain some kind of discourse cue, whereas it is more optional for independent and coordinate EDUs, levels (a)-(c), to explicate the rhetorical relation to their adjacent EDUs (cf. the discussion of the ambiguous coordinator *and* in the previous section). And it is exactly this overlap between syntax and discourse which leads us from the domain of text structure into the domain of discourse structure, which is the topic of Chapter 5, cf. also Matthiessen & Thompson (1988, p. 301): 'Clause combining in grammar has evolved as a grammaticalisation of the rhetorical units in discourse [i.e. EDUs] defined by rhetorical relations'.

It could be argued that the scale does not take into account the various tense, aspect and mood distinctions found at levels (d)-(e), which ultimately do integrate the subordinate EDU more into the matrix EDU. In cases where a coordinated or subordinated EDU is in the subjunctive mood, the subordinate EDU is clearly more dependent on its matrix than subordinate EDUs in indicative mood (Ferrari & Zampese, 2000, p. 155). However, the subjunctive no longer exists in Danish. In the context of this thesis, differences in mood will not be taken into account in the following sections, as they are not valid parameters of comparison for the three languages under investigation. In addition, it could also be argued that the scale does not take into consideration the various types of relative clauses (level e), which ultimately differ in some aspects. See Section 3.5.2.3 for a discussion of why relative clauses have been grouped together.

An EDU can be textualised as a	Textualisation examples
a. independent sentence (including sentence fragments)	[Our group supports this proposal.] [We **will vote** in favour of it.]
b. main or matrix clause that is part of a sentence	[Our group supports this proposal;] [we **will vote** in favour of it.] [Because our group supports this proposal,] [we **will vote** in favour of it.]
c. coordinate main or matrix clause (including ellipsis)	[Our group supports this proposal,] [and we **will vote** in favour of it.]
d. subordinate finite adverbial clause	[Because our group **supports** this proposal,] [we will vote in favour of it.]
e. relative clause	[Our group supports this proposal] [of which we **will vote** in favour.]
f. infinitival clause	[Our group supports this proposal] [in order to **have** it approved.]
g. gerund, present or past participle construction	[Our group supports this proposal,] [**voting** in favour of it.] [As **said** by my colleague,] [we support this proposal.]
h. present, past participial or adjectival modifier	[Our group supports this proposal] [**involving** EU citizens.] [Our group supports the proposal] [**made** by the commissioner.] [Our group supports the proposal] [**present** in today's agenda.]
i. nominalisation	[Our group will vote in favour] [without **support** from all members.]
j. verbless construction	[By now **[Ø]** in favour of the proposal,] [our group supports it.]

Figure 4.1: The deverbalisation scale

4.3 Text structure in Danish, English and Italian

A review of the cross-linguistic literature dealing with Danish, English and Italian reveals a number of contrasts in the way the three languages typically structure texts, and in particular, deal with syntactic dependency and realisation. Since most cross-linguistic studies examine only two languages at a time, I have not found any literature that deals with Danish, English and Italian simultaneously. As a result, most of the studies referred to here are comparisons of only two of the languages involved. To support the bases of the hypotheses presented in the end of this section, I have also taken into consideration a number of studies dealing with languages closely related to the three languages under investigation in terms of typology. It should be noted that the differences outlined in this chapter, in general, are considered to be varying frequencies – all other things being equal – of universal features, where one language may show a preference for one feature, which the other two languages do not.

4.3.1 Danish versus Italian

Danish and Italian have been claimed to be markedly different in a number of ways. The most important contributions of Danish-Italian comparisons are found in the works by Skytte & Korzen (Skytte et al., 1999; Skytte & Korzen, 2000) and in later works of Korzen. Here, it has been found that Italian to a larger extent than Danish uses subordination and thus displays a more hierarchical text structure. The findings are partly based on narrative text types (retelling of movie sequences), where Danish texts have a stronger tendency to view events on a temporal axis by presenting them in a chronological sequence through coordination, while the Italian retellings tend to have events linked more tightly together by means of an intensive use of subordination. This more frequent use of subordination is not only due to more subordinate finite clauses but also due to a more frequent use of subordinate non-finite clauses. The studies also base their findings on argumentative texts, which display many of the same differences as the narrative texts. Korzen (Korzen & Gylling, 2012b; Korzen, 2009a) also reports on a selection of related studies including other text genres and types such as legal texts, websites and technical texts in Danish and Italian, all confirming the previous findings of Italian texts tending to have a more hierarchical text structure than Danish, that is, Italian uses more subordination. Since this claim has been tested on a wide range of text genres and types, we would expect the same pattern to be found in the texts of the Europarl corpus.

Apart from a more intensive use of subordination and a more frequent use of subordinate non-finite clauses, the studies mentioned above also provide evidence for more lexical variation and longer sentences in Italian than in Danish. The lexical variation is relevant for this study as

it has previously been argued that anaphora and discourse structure are closely related concepts, see Chapter 2, and that the interpretation of rhetorical relations also depends on the presence of an anaphor; if an anaphor in one EDU varies lexically from the antecedent in the previous EDU, the rhetorical relation between the two EDUs could be interpreted as some kind of elaboration. The notion of longer sentences, which is dealt with in Chapter 6 on information structure, is also important for this thesis because there may be differences in the distribution of rhetorical relations between texts with short sentences and texts with longer sentences. In addition, there could also be differences between the signalling of the relations between the EDUs as clauses are more frequently linked by a discourse cue than sentences. This phenomenon is also dealt with in Chapter 6.

4.3.2 English versus Italian

The literature on English-Italian differences and similarities in text structure is much less abundant empirically speaking. However, previously cited Pierini (2004) points out many of the same differences between English and Italian text structure that Skytte & Korzen (Skytte et al., 1999; Skytte & Korzen, 2000) also found for Danish and Italian. That is, English text structure is more linear than Italian text structure. English uses shorter sentences and does not subordinate to the same extent as Italian. Thus, English can be expected to be closer to Danish than to Italian, which from a typological perspective would make sense as both Danish and English are Germanic languages, whereas Italian is a Romance language. It also makes sense if we consult previous studies of Germanic versus Romance languages. Cosme (2008), for example, showed from a translational perspective that coordination in English L1 texts is translated with subordination in French (Romance) L2 texts. She also showed that Germanic languages vary internally: for instance, that subordination in English L1 texts is translated with coordination in Dutch (Germanic) L2 texts. As for the latter translation shift, the main portion of the changes consisted of English –*ing* clauses that were changed into coordinated constructions in Dutch L2. Fabricius-Hansen (1998) also compares Germanic languages internally from a translational viewpoint, and shows that English, German and Norwegian prefer different clause linking patterns. In particular, it is argued that German favours a hierarchical or vertical text structure, while English prefers an incremental, that is, horizontal or linear, text structure. Norwegian, then, is argued to prefer an even more linear text structure than English, as EDUs are organised in smaller chunks. As Danish and Norwegian in many ways are very close to each other, it shall be assumed that Danish exhibits some of the same patterns as Norwegian. Finally, other studies

of French versus Germanic languages show evidence of the same patterns in French as in Italian, with more subordination and non-finite constructions in French than in English and other Germanic languages (e.g. M. Ballard, 1995; Chuquet & Paillard, 1987; Korzen & Lundquist, 2003; Lundquist, 2005; Vinay & Darbelnet, 1995).

4.3.3 Hypotheses

As argued and demonstrated by the above-mentioned scholars, there seem to be noteworthy contrasts between Danish, English and Italian. Danish and English texts are more similar to each other in terms of text structure, by using more coordination than Italian. By contrast, subordination seems to occur more frequently in Italian texts. These two text structure styles have previously been described by Fabricius-Hansen (1996, p. 558) as the two poles of a continuum ranging from incrementality (coordination) to hierarchy (subordination), and by Chafe (1982, p. 38) as 'fragmented' style (coordinating structures) versus 'integrated' style (subordinating structures). Following the ideas presented above, Danish should be located at the incremental and fragmented end of the continuum, and Italian at the hierarchical and integrated pole. Owing to its more productive spectrum of non-finite constructions, English should be located somewhere along of along the two extremes of the continuum, somewhat closer to Danish than to Italian.
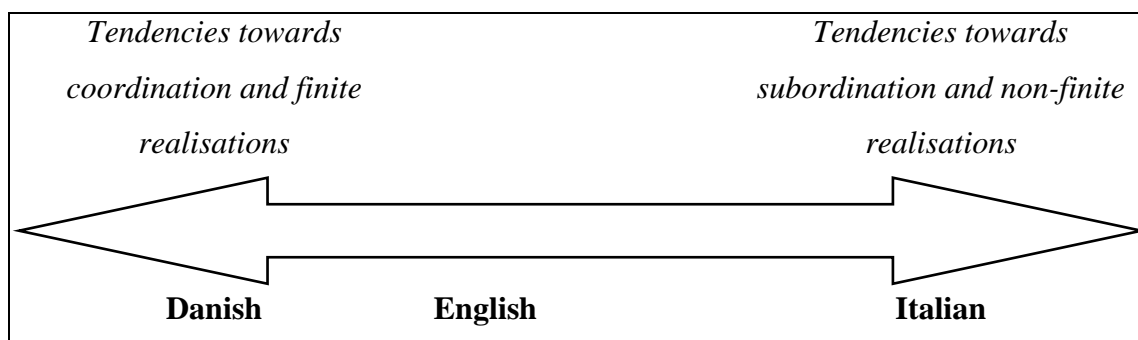

| *Tendencies towards coordination and finite realisations* | | *Tendencies towards subordination and non-finite realisations* |
| **Danish** | **English** | **Italian** |

Figure 4.2: Text structure hypothesis

Figure 4.2 is not to be interpreted in absolute terms, meaning that Danish only uses coordination, and Italian only subordination. Rather, the differences in text structure are to be considered as relative concepts, positing that Danish and English, as compared with Italian, prefer more coordination and finite realisations of EDUs, and that English, as compared with Danish, favours slightly more subordination and non-finite realisations of EDUs, all other things being equal. In relation to the deverbalisation scale, Danish should have relatively more textualisation of EDUs pertaining to the first three levels, whereas English and Italian should

have relatively more textualisations pertaining to the levels below. In the following sections, I explain the analyses of these two parameters and present the results.

4.4   Dependencies of EDUs

This section explains the analysis of dependencies of EDUs in the corpus. For this analysis, I first took the EDUs (see Chapter 3 for an overview of syntactic constructions considered) as units of analysis. EDUs can either be related to other EDUs inside the same sentence or to EDUs in other sentences. The EDU relations analysed in this chapter are all intrasentential, as it is the syntactic type of coordination and subordination that we are interested in here. The EDUs linked to other EDUs outside the sentence boundary are analysed in Chapter 5 and 6.

Table 4.1 shows the number of EDU relations considered for the analysis of text structure in each of the three languages, and the total number of EDUs in the corpus. The discrepancies between the languages are due to significantly longer sentences in Italian than in Danish and English, resulting in more intrasententially related EDUs and in a lower number of sentences in the Italian texts than in the texts of the other two languages (see also Table 6.1 in Chapter 6). These differences are adjusted in the statistical analysis.

	Danish		**English**		**Italian**	
EDUs in corpus	1443		1325		1469	
	N	**%**	**N**	**%**	**N**	**%**
Intersentential EDU relations	680	47	608	46	437	30
Intrasentential EDU relations	763	53	719	54	1032	70

Table 4.1: Numbers of EDUs in the corpus

4.4.1   Three examples from the corpus

Three examples are provided of syntactic dependencies in co-text, one for each language. None of them are full texts but meaningful excerpts of the first ten EDUs in a text. Just as in the remainder of the chapter, syntactic dependencies are represented in the right columns according to the various levels in the deverbalisation scale presented in Figure 4.1. The letters in parentheses refer to the levels of the scale. Embedded EDUs are shown in square brackets with an indication of their identification number (EDU#) in the text; see EDUs #2-4 in Table 4.3. As for the Danish and Italian examples, English translations are provided in the middle column;

grammatical characteristics are highlighted between square brackets in capital letters; see EDU #9 in the Danish excerpt.

#	EDU	English translation	Syntactic relation (dependency)
1	Hr. formand, lad mig først slå fast, at for os at se er Amsterdam-traktaten en langt bedre traktat end Maastricht-traktaten.	*Mr President, let me first establish that in our view the Amsterdam Treaty is a much better treaty than the Maastricht Treaty.*	independent sentence (a)
2	Det synes jeg er vigtigt at holde fast i.	*I think it is important to bear that in mind.*	independent sentence (a)
3	Der er i Amsterdam-traktaten lagt vægt på, hvad vi i Danmark har kaldt vores mærkesager,	*The Amsterdam Treaty places emphasis on what we in Denmark have called our important issues,*	main clause as part of a sentence (b)
4	men det er også mærkesager for mange andre.	*but they are also important issues for many others.*	coordinate main clause (c) to #3
5	Eksempler herpå er bedre miljøbeskyttelse, samarbejde om bekæmpelse af arbejdsløshed, åbenhed og større indflydelse til Europa-Parlamentet.	*Examples include better environmental protection, cooperation in the fight against unemployment, transparency and greater influence of the European Parliament.*	independent sentence (a)
6	Jeg vil dog sige, at på nogle områder har vi noget besvær med betænkningen om Amsterdam-traktaten.	*However, I would like to say that we have some difficulties with the report on the Amsterdam Treaty regarding some points.*	independent sentence (a)
7	Vi er nødt til at stemme imod nogle af områderne.	*We will have to vote against some of the points.*	independent sentence (a)
8	Vi mener, at det er meget beklageligt, at man vil forsøge at lægge hindringer i vejen for udvidelsen	*We think it is very regrettable that people should want to put obstacles in the way of enlargement*	matrix clause (b)
9	ved at kræve, at alle institutionelle reformer skulle være på plads,	*by [TO DEMAND] that all institutional reforms should be in place*	infinitival clause (f) to #8
10	inden udvidelsen påbegyndes.	*before the enlargement begins.*	subordinate finite adverbial clause (d) to #9

Table 4.2: Danish excerpt with annotations <ep-97-11-19.txt:161>

In this example from the Danish part of the corpus, we can see how five out of ten EDUs have been textualised as independent sentences (level a), and only two EDUs (#9 and #10) are textualised at the more dependent, subordinate levels (d) and (f). The English example in Table

4.3 shows similar patterns with respect to syntactic relations: four EDUs are independent sentences and three different EDUs are found at the other top levels (b)-(e) of the deverbalisation scale, EDU#2+4+8. The main difference with the Danish example is the presence of a participial modifier (level h) in EDU#3.

#	EDU	Syntactic relation (dependency)
1	Mr President, firstly, I find it quite incredible that Mr Tindemans is being hailed here as a visionary.	independent sentence (a)
2	I do not know how someone [EDU#3+4) can be considered a visionary.	matrix clause (b)
3	living in the Dark Ages,	present participial modifier (h) to *someone* in # 2
4	who believes in military blocs, global militarisation and imperialism	relative clause (e) to *someone* in #2 and #3
5	It is a ludicrous thought.	independent sentence (a)
6	On the issue of imperialism, it is quite clear in paragraph 4 of this document that 'the purpose of a common defence policy is to protect the Union's interest in all areas, including security of supply'.	independent sentence (a)
7	That is disgraceful.	independent sentence (a)
8	You are saying that you want a defence policy and defence system,	matrix clause (b)
9	not to protect yourself against potential aggressors,	infinitive clause (f) to #8
10	but to protect your fundamental selfish interests.	coordinate infinitive clause (f) to #9

Table 4.3: English excerpt with annotations <ep-98-05-13.txt:201>

The Italian example in Table 4.4 exhibits many of the textualisations found in the Danish and English examples. However, there are no EDUs shaped as whole independent sentences, because all intrasentential EDUs have subordinate EDUs attached; in some cases such as EDUs#4-6 and EDUs#8-10, even subordinate EDUs have subordinate EDUs attached, yielding a deeper text structure consisting of more degrees of subordination (see Ferrari & Zampese, 2000, p. 143), as shown in Section 4.4.2.1. This phenomenon is feasible because the Italian text has longer and fewer sentences (see Table 4.1), and because it contains embedded insertions such as the ones in EDUs#5+6 and EDUs#9 and relative clauses in EDU#9+10. In the next

section, I shall elaborate on the differences observed in these examples and describe the overall dependency patterns of the three languages.

#	EDU	English translation	Syntactic relation (dependency)
1	Signora Presidente, desidero informare il Parlamento di una grave iniziativa	*Madam President, I would like to inform Parliament of a serious initiative*	matrix clause (b)
2	che è stata presa la settimana scorsa, in Italia, dal governo regionale della Regione Lazio.	*which has been undertaken last week in Italy by the regional government of the Lazio region.*	relative clause (e) to *iniziativa* #1
3	Questo ha infatti costituito una commissione per il controllo dei libri di testo di storia nelle scuole, commissione	*The regional government has established a committee to monitor the history textbooks used in schools, committee*	matrix clause (b)
4	che deve controllare fino a che punto [EDU#5+6] questi testi sono inficiati di marxismo.	*which is to ascertain the extent to which [EDUs#5+6] these texts are influenced by Marxism*	relative clause (e) to *commissione* in #3
5	- come dicono i neofascisti	*- as the neo-fascists say*	subordinate finite adverbial clause (d) to *punto* in #4
6	che sono al governo nel Lazio -	*who are in power in Lazio -*	relative clause (e) to *neofascisti* in #5
7	Il fatto è che questi testi presentano soltanto la storia d'Italia, della Seconda guerra mondiale, del nazismo e del fascismo	*The fact is that these texts merely present the history of Italy, of the Second World War, of Nazism and Fascism,*	matrix clause (b)
8	assumendo, come ovvio, che non si possono mettere sullo stesso piano i valori [EDU#9] e le idee [EDU#10]	*assuming that it is clear that the values [EDU#9] and the ideas [EDU#10] cannot be regarded as equal.*	gerund (g) to #7
9	per cui si sono battuti i resistenti e gli antifascisti	*for which the resistance and the antifascists were fighting*	relative clause (e) to *valori* in #8
10	che hanno ispirato Hitler e Mussolini.	*which motivated Hitler and Mussolini*	relative clause (e) to *idee* in #8

Table 4.4: Italian excerpt with annotations <ep-00-11-13.txt:31>

4.4.2   Overall dependency patterns in Danish, English and Italian

As mentioned in the introduction to this chapter, all the EDUs considered in this section are intrasententially related EDUs, that is, EDUs that are syntactically related to other EDUs within the same sentence. Although, in the three tables above, a number of independent sentences (that is, EDUs related to EDUs in other sentences) were observed, these will not be considered in the

following analyses, because they cannot be attributed any syntactic relationship of coordination or subordination to other EDUs, which is one of two parameters under investigation of the present chapter. Instead these will be dealt with in the following chapters.

Regarding the Danish and English texts, EDUs that are intrasententially related to other EDUs account for approximately 50 % of the total number of EDUs of the entire corpus, whereas they comprise 70 % of the total number of Italian EDUs, cf. Table 4.1. Table 4.5 shows the intrasentential distribution of syntactically coordinated EDUs and the syntactically subordinated EDUs in each of the three languages. Subordinated EDUs that are coordinated to another subordinated EDU such as the English coordinated infinitival clause in EDU#10 in Table 4.3 above are included in the category subordinated EDUs.

	Danish		English		Italian	
	N	**%**	**N**	**%**	**N**	**%**
Coordinated EDUs	242	32	167	23	267	26
Subordinated EDUs	521	68	552	77	765	74

Table 4.5: Overall dependency distribution

As seen above, subordination between EDUs is the most common dependency for all languages with approximately 75 % of the total intrasententially related EDUs. This does not come as a surprise, as coordination mainly is restricted to verb phrase plus verb phrase constructions. By comparison, subordination covers a wide range of subordinate constructions, namely seven finite and non-finite constructions, see the deverbalisation scale in Figure 4.1. What is perhaps more surprising, if we recall the hypotheses presented in the previous section, is that English has the lowest number of coordinated EDUs of the three languages. One of the reasons for this is that the English sentences are shorter than the Italian sentences in terms of EDUs. Instead of coordinating EDUs syntactically inside the same sentence, the English writers have used rhetorical coordination to link EDUs intersententially. But this type of coordination is, as mentioned, not part of the dependency analysis. In contrast to this, the Italian writers have used longer sentences in terms of EDUs, some of which contain both subordinate and coordinate EDUs. Compare the English example in 62) with the Italian example in 63). In both examples, the main EDU is shown in bold, and the coordinated EDUs have been underlined.

62) **[Yet,** [as Mr Tindemans says,] **the Amsterdam Treaty makes a common defence policy a more credible prospect**] [given that it strengthens the organic bond between

the European Union and the Western European Union,] [created by Maastricht.] <ep-98-05-13.txt:201>

63) [**Non so se quella è la società civile**] [a cui pensa,] [ma mi fa un po' paura] [perché, [come ha detto il collega Swoboda,] oggi lei ci ha fatto un discorso molto pro-parlamentare] [ma, nel testo, tanti riferimenti al Parlamento io non li ho visti!] <ep-01-09-04.txt:150>

*[**I do not know whether that is the civil society**] [to which you are referring,] [but I am somewhat fearful] [because, [as Mr Swoboda said,] today, you have delivered a very pro-Parliament speech,] [but in the text itself, I cannot find many references to Parliament!]*

The English sentence in 62) contains four EDUs of which all the intrasententially related ones are subordinated to the matrix EDU or other subordinated EDUs. In the Italian sentence in 63), which contains six EDUs, two of the internally related EDUs are coordinated: the first is coordinated with the main EDU, the second coordinated with the subordinate EDU starting with *perché* (because).

Another possible explanation for why English seems closer to Italian than to Danish can be found in the segmentation principles adopted in this study (see Section 3.5.2.6 in Chapter 3). Following a number of other scholars (Carlson & Marcu, 2001, p. 11; da Cunha & Iruskieta, 2010, p. 570), I take colon and semicolon to mark clause boundaries, not sentence boundaries; see also Huddleston & Pullum (2002, p. 1735f), who state that like comma, colon and semicolon 'normally mark boundaries within a sentence, and hence can be regarded as secondary boundary marks' located between comma and full stop in a hierarchy of 'relative strength' (see also Fabricius-Hansen & Ramm, 2008, p. 8; Ferrari & Zampese, 2000, p. 236). In the Italian texts, 43 instances of colon and 35 instances of semicolon were observed as compared with only two colons and nine semicolons in the English, and one single colon and one single semicolon in the Danish. This more extensive use of colon and semicolon allows Italian sentences to include more information than Danish and English sentences, and to coordinate EDUs by means of colons and semicolons – EDUs that in Danish and English would probably have been separated by full stops[7]. This means that the overall frequency of coordinated EDUs in the Italian texts is higher than in English texts. Consider, for instance, the Italian example in 64) consisting of nine EDUs, two of which are coordinated by a colon and a semicolon respectively.

---

[7] This claim is supported by a tendency in the official Danish and English translations of the Italian texts of transforming a colon or a semicolon in the L1 into a full stop in the L2.

64) [A Cork si è svolta una conferenza sull'agricoltura, sul mondo rurale e sull'agricoltura compatibile con l'ambiente,] [nel corso della quale si è detto che le misure agroambientali dovranno diventare centrali nella nuova PAC**:**] [eppure qui si va verso decisioni] [che avallano un aumento massiccio del dosaggio di una certa tossina, tossina] [che comporta il rischio di sviluppare una nuova generazione di insetti] [resistenti ad un insetticida amico dell'ambiente**;**] [oppure, si rischia di utilizzare su ampia scala un erbicida, con le conseguenze] [che possono esserci per le acque,] [in cui questo erbicida è solubile.] <ep-97-04-07.txt:81>

[A conference was held in Cork on agriculture, the rural world and environmentally compatible agriculture,] [during which it was said that agrienvironmental measures should become central to the new CAP:] [and yet here we are moving towards decisions] [which will endorse a massive increase in the dosage of a particular toxin, a toxin] [which involves the risk of developing a new generation of insects] [resistant to an environment-friendly insecticide;] [or we are in danger of making extensive use of a herbicide, with the consequences] [that can entail for the water] [in which that herbicide is soluble.]

In a similar case from the Danish part of the corpus, the same number of EDUs have been distributed across seven independent sentences of which only two contain more than one EDU. Rather than continuing the information flow with relative clauses like the Italian example above, the Danish text in 65) repeats the relevant anaphoric noun or verb phrase as the subject (a definite noun or a pronoun) of the preceding sentence (*glyfosat er trængt ned/glyphosate has penetrated → det/that*; *glyfosat/glyphosate → sprøjtemidlet/the crop spray*; *Roundup → Roundup*; *Monsanto → Monsanto*).

65) [Hr. formand, i Danmark er det i weekenden kommet frem, at **glyfosat er trængt ned** i de øvre grundvandsmagasiner.] [**Det** har stor betydning i et land som Danmark,] [hvor vi drikker vandet direkte fra vandhanen.] [**Sprøjtemidlet** hedder **Roundup**] [og bliver produceret af Monsanto.] [**Roundup** er udbredt og godkendt i resten af EU.] [For mig er det ingen overraskelse, at systematisk spredning af gift ender i vores dyrebare grundvand.] [I en sådan sag må der kunne lægges et klart ansvar hos **Monsanto** for det danske grundvand.] [**Monsanto** står også bag gensplejsning af planter i stor stil.] <ep-03-05-13.txt:28>

*[Mr President, it emerged in Denmark at the weekend that **glyphosate has penetrated** into the upper groundwater reservoirs.] [**That** is very significant in a country like Denmark] [where we drink water directly from the tap.] [**The crop spray** is called **Roundup**] [and is manufactured by Monsanto.] [**Roundup** is approved and used*

extensively in the rest of the EU.] [For me, it is no surprise that systematically dispersed poison ends up in our precious groundwater.] [In a case like this, it has to be possible to hold **Monsanto** *clearly liable for Danish groundwater.]* [**Monsanto** *is also behind the large-scale genetic modification of plants.]*

Another frequent text structure pattern in Danish consists of a series of coordinated EDUs, as seen in example 66), where two coordinated EDUs occur within the same sentence. The relevant EDUs have been underlined.

66) [Det tilkommer ikke flertallet i Parlamentet at korrigere valgresultaterne] [ved at tvinge samtlige vælgere til at bidrage til finansiering af fem overnationale EU-partier,] [og det er utroligt, at Kommissionen som lovlighedens vogter kan fremsætte et så åbenlyst ulovligt forslag,] [og det er ufatteligt, at Parlamentet i dag vil vedtage et forslag,] [som efter den eksisterende retspraksis kan erklæres ulovligt ved EF-Domstolen.] <ep-01-05-17.txt:43>

[It is not for the majority in Parliament to correct the election results] [by forcing the entire electorate to contribute to the financing of five supranational EU parties,] [and it is incredible that the Commission, as the guardian of legality, is able to submit such a clearly unlawful proposal,] [and it is inconceivable that Parliament should today adopt a proposal] [which, in accordance with current legal practice, can be declared unlawful by the EC Court of Justice.]

4.4.2.1 *Degrees of subordination*

As mentioned earlier, Ferrari & Zampese (2000, p. 143) point out that subordinated EDUs can be subordinated to other subordinate EDUs, entailing a more hierarchical text structure. An example of this phenomenon is found in 67), where a participial modifier (*demanding*) is subordinated to a relative clause (*which avoids*), that again is subordinated to another participial modifier (*having*) which also is subordinated to a relative clause (*which, in fact, only allows*) that is finally subordinated to the matrix EDU of the sentence.

67) [The present interpretation was adopted] [which, in fact, only allows for one initial signature to be on a Rule 48 resolution and for people] [having to sign it in a special office,] [which avoids having to walk a line of people] [demanding that they actually sign up to particular Rule 48 resolutions.] <ep-98-06-15.txt:42>

In this way, we can refer to different degrees of subordination, where the EDUs in example 67) reach four degrees of subordination, because the matrix EDU has four subordinated EDUs

attached. In Figure 4.3, example 67) has been visualised in a stair-step diagram which shows how the EDUs are subordinated to each other, resulting in a deep and dependent text structure.
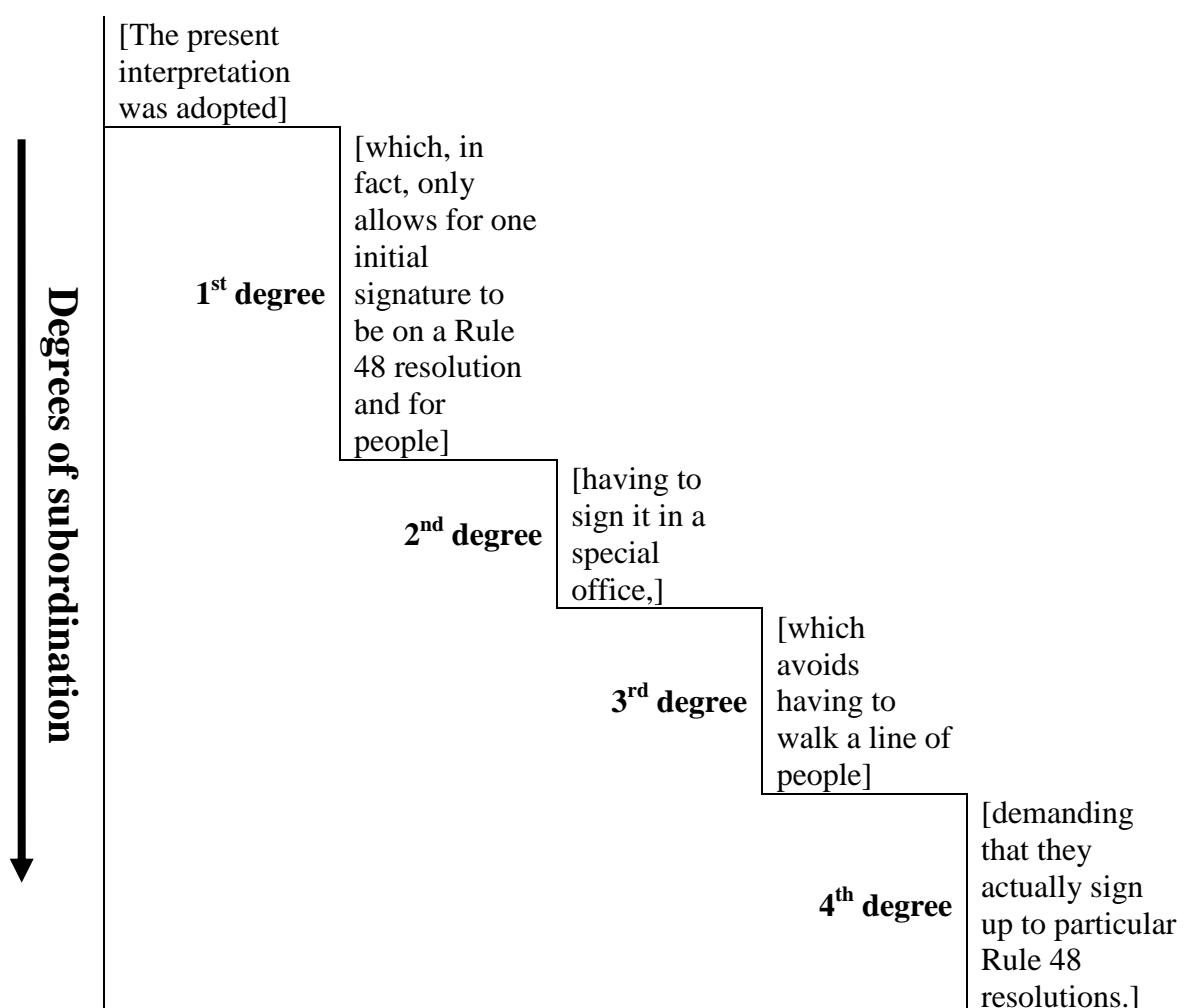

Figure 4.3: Degrees of subordination

A high degree of EDU subordination is usually caused by a number of relative clauses, or, in English and Italian, modifiers, as shown is the example above. But other subordinate constructions, such as infinitive clauses and finite adverbial clauses, can also be observed here. If we explore the cross-linguistic differences of this phenomenon between Danish, English and Italian, the numbers in Table 4.6 reveal strong indications of Italian diverging from Danish and English. Whereas more than 70 % of the Danish and English subordinate EDUs do not have other subordinated EDUs attached (first degree), only 45 % of the Italian subordinate EDUs fall inside the first degree. This again is due to the longer Italian sentences, and due to the higher numbers of relative clauses and modifiers in general. As can be seen in the table, one of the Italian sentences contains up to the ninth degree of subordination; Danish stops at the sixth degree and English at the seventh degree. Example 68) shows how relative clauses in Italian can

be used to create a highly hierarchical text structure, in which each underlined EDU elaborates on the previous EDU. The example also shows how Italian coordinates subordinate constructions to each other, in this case with coordinated relative clauses.

68) [In questo senso si propone un emendamento] [che fa riferimento alla classificazione dell'OCSE] [e che dovrebbe quindi escludere la flotta dell'Arabia Saudita dai benefici] [che derivano, appunto, dal provvedimento] [che riguarda gli aerei inquinanti,] [che non fanno rumore] [e che hanno più di 25 anni di età.] <ep-98-01-12.txt:43>

[In this connection, an amendment is tabled] [that refers to the OECD classification,] [and that should therefore exclude the Saudi Arabian navy from the benefits] [which result from the measure] [which relates to polluting aircraft] [that do not make a noise] [and that are more than 25 years old.]

	1st	2nd	3rd	4th	5th	6th	7th	8th	9th
■ Danish	75,5	17,9	5,0	1,2	0,2	0,2	0,0	0,0	0,0
■ English	70,8	21,3	6,1	1,1	0,3	0,0	0,5	0,0	0,0
■ Italian	45,2	27,8	16,2	6,2	1,8	1,8	0,5	0,3	0,3
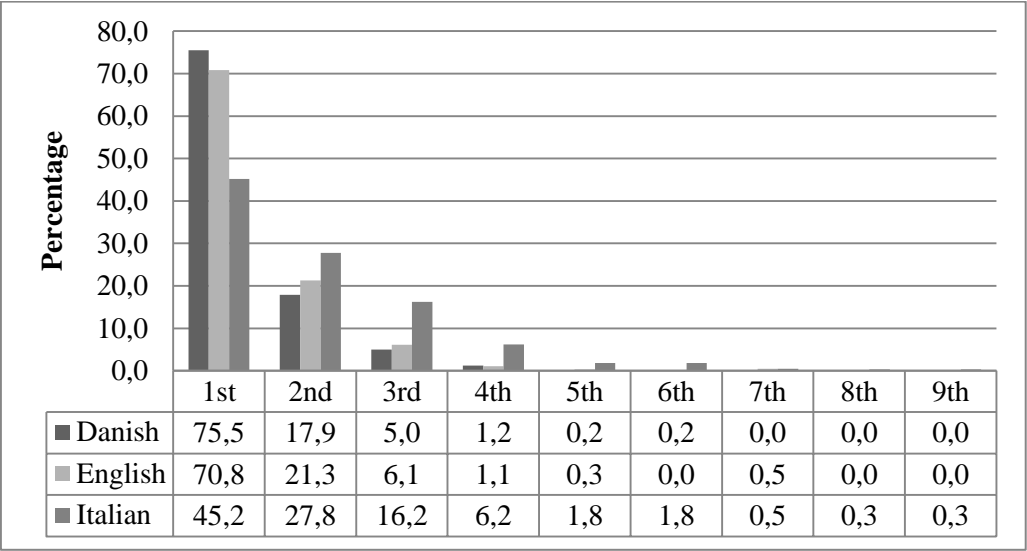
Table 4.6: Degrees of subordination in Danish, English and Italian

The numbers in Table 4.6 illustrate that although English and Italian text structure on the surface seems to converge with regard to the overall numbers from Table 4.5 between coordination and subordination, the two languages do not use subordination in the same manner. In Italian, subordination is more frequently used to create longer chains of subordinated EDUs, whereas subordination in English is mainly used to pair a matrix with a subordinate EDU, in almost two-thirds of all instances. Thus, the use of subordination in English is more similar to that found in Danish. As already mentioned, the explanation for this is mainly found in the discrepancies between the intersentential and intrasentential linkage of EDUs, where Italian turned out to include more EDUs inside the sentence boundary.

4.5   Realisations of subordinate EDUs

In this second analysis of text structure characteristics, the focus is again on the subordinated EDUs. The analysis of the corpus proceeds as follows. I consider all the subordinate EDUs in the three languages, and, for each of them, examine whether the main verb of each EDU is realised as a finite or non-finite verb form. Phrasal EDU constructions where the verb has been omitted were in this analysis also considered as verbless constructions. As we saw from the different levels of the deverbalisation scale in Figure 4.1, the first two levels that include subordinate EDUs (d-e) contain finite verb forms, while the next three levels (f-h) contain non-finite verb forms. The last two levels (i-j) comprise nominalisations and the complete omission of verbal elements. The three examples below each show a syntactic construction belonging to the above-mentioned realisations; in 69), from the Danish part of the corpus, the main verb (*mødes/meet*) of the adverbial clause is finite, in 70), from the Italian corpus, the subordinate clause begins with a gerund (*riconoscendo/recognising*) and in 71), from the English corpus, the EDU consists of a nominalisation (*appointment*) of the verb *appoint*. All relevant EDUs are underlined, and the main verb of the EDUs in question shown in bold-faced type.

69) [Jeg synes, statsministrene skulle studere det danske eksempel,] [<u>når de **mødes** til</u> <u>økonomisnak i Lissabon.</u>] <ep-00-03-13.txt:27>
[I believe that the EU's prime ministers should study the Danish example] *<u>[when they</u>* *<u>meet for their tête-à-tête on the economy in Lisbon.]</u>*

70) [Aspettiamo quindi con attenzione che la prossima Conferenza intergovernativa valorizzi definitivamente il ruolo di associazioni e fondazioni,] [**<u>riconoscendo</u>** <u>organizzazioni non</u> <u>profit e organizzazioni non governative come soggetti fondamentali per la crescita</u> <u>sociale comune europea.</u>] <ep-98-07-02.txt:29>

[Let us therefore hope that the next Intergovernmental Conference finally promotes the role of organisations and foundations,] *<u>[**recognising** non-profit organisations and non-</u>* *<u>governmental organisations as fundamental bodies for common social growth in</u>* *<u>Europe.]</u>*

71) [<u>Commissioner, since your **appointment**</u>] [you have adopted a very good habit of consulting and informing the European Parliament,] [for which we are grateful.] <ep-03-03-10.txt:60>

The overall results of this analysis can be found in Table 4.7. As expected, most subordinated EDUs are realised with finite verb form, as finite adverbial clauses or relative clauses. Again,

we can see an overall tendency of English and Italian to be closer to each other in terms of text structure, with around 40 % non-finite and verbless realisations. As we saw above in the degrees of subordination, Danish tends to prefer the same type of realisation for a vast majority of the subordinate EDUs found in the corpus. In the table, we can see that almost 90 % of the Danish subordinate EDUs have been realised as finite clauses as compared with approximately 60 % of the English and Italian EDUs.

	Danish		English		Italian	
	N	**%**	**N**	**%**	**N**	**%**
Finite verb forms	453	87	337	61	482	63
Non-finite verb forms	38	7	178	32	217	29
Verbless constructions	32	6	35	7	66	8

Table 4.7: Overall realisation of subordinate EDUs

These results can more easily be compared in Table 4.8, where the different realisations have been divided in a more detailed manner so that they correspond to the various levels of the deverbalisation scale including subordinate EDUs (d-j). Table 4.8 shows one striking difference between English and Italian on the one side, and Danish on the other, namely the higher usage of participial constructions (level g) and of non-finite modifiers (level h). This might appear to be an unfair parameter of comparison since the Italian gerund does not have any equivalent constructions in Danish, but as seen in the example below, Danish is in fact able to realise EDUs at levels (g-h). However, what is interesting is that Danish does not use these levels as frequently as English and Italian. The example in 72) constitutes the only instance from the corpus of a Danish past participial EDU (*begrundet/motivated*) along with a series of nominalisations (*hensyntagen/need*, *efterforskning/investigation*, *retsforfølgning/prosecution)* which, however, are not EDUs. In general, my impression of this particular text is that it is probably a speech that was originally written in a language other than Danish. This could explain the higher number of nominalisations and longer sentences (33 words per sentence on average) as compared with the other texts in the Danish corpus.

72) [Formandskabet har fremsat forslag til rådskonklusioner vedrørende forholdsregler i relation til informationsteknologi] [**begrundet** i hensyntagen til efterforskning og retsforfølgning af organiseret kriminalitet.] <ep-02-10-23.txt:238>

*[The Presidency has tabled draft Council conclusions concerning measures in relation to information technology,] [**motivated** by the need for investigation and prosecution of organised crime.]*
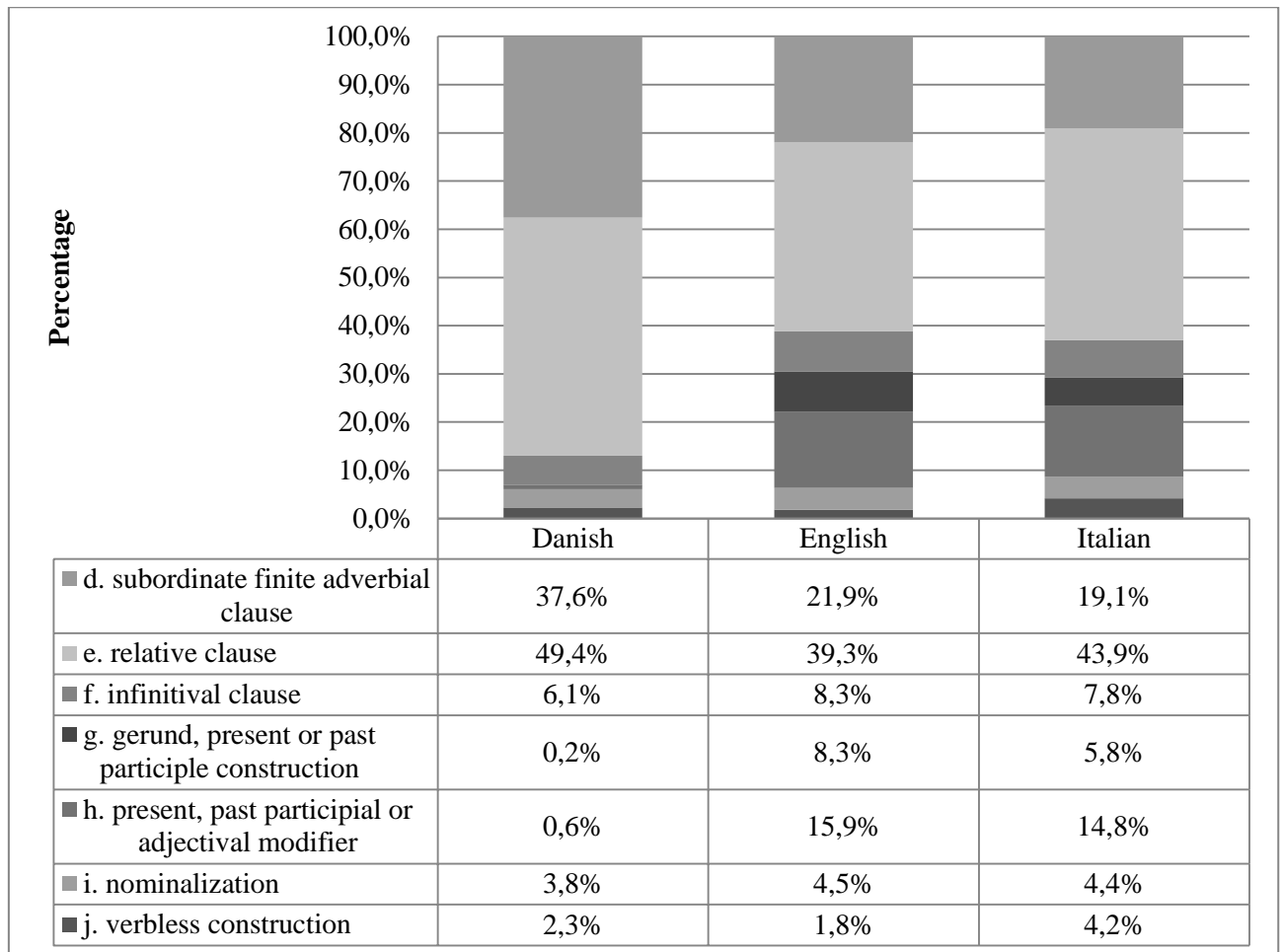
	Danish	English	Italian
■ d. subordinate finite adverbial clause	37,6%	21,9%	19,1%
■ e. relative clause	49,4%	39,3%	43,9%
■ f. infinitival clause	6,1%	8,3%	7,8%
■ g. gerund, present or past participle construction	0,2%	8,3%	5,8%
■ h. present, past participial or adjectival modifier	0,6%	15,9%	14,8%
■ i. nominalization	3,8%	4,5%	4,4%
■ j. verbless construction	2,3%	1,8%	4,2%

Table 4.8: Detailed distribution of realisations

No other textualisations of EDUs were found at these two levels (g-h) in the Danish texts, but in the official Danish translation of the two following English L1, a past participle modifier (*sold*) in 73) and a present participle construction (*starting*) in 74) have been translated into similar syntactic constructions in Danish (*solgt* and *startende*). However, these were some of the only instances in the corpus; usually all textualisations of levels (g-h) in English and Italian L1 were translated into relative clauses in Danish L2.

73) [What fellow Members may not know is that the main source of the disease was contaminated paté] [**sold** from three Belgian plants.] <ep-99-07-22.txt:25>

Official Danish L2: *[Hvad de andre medlemmer måske ikke ved, er, at hovedkilden til sygdommen var kontamineret paté] [**solgt fra tre belgiske fabrikker**.]*

74) [Thus, we would say: by all means have an enquiry,] [but first let us have a wide-ranging review of all Member States' systems for dealing with food and health scares,] [**starting with an evaluation of the British system.**] <ep-99-07-22.txt:25>

Official Danish L2: *[Derfor vil vi sige: Lad os for alt i verden få en undersøgelse,] [men lad os først få en vidtrækkende gennemgang af alle medlemsstaters systemer til håndtering af fødevare- og sundhedskriser,] [**startende** med en evaluering af det britiske system.]*

Since subordinate finite adverbial clauses and gerund participial constructions are essentially able to express the same content, I find it reasonable to believe that English and Italian textualisations at level (g) (gerunds and participles) in many cases correspond to Danish level (d) realisations (subordinate finite adverbial clauses). This could explain the higher number of Danish level (d) realisations. Again, if we consult the Danish official translations of the English and Italian L1 texts, this claim is corroborated, see example 75), where an Italian gerund (*augurandomi/(me) hoping*) is translated with a Danish subordinate finite adverbial clause (*idet jeg håber/as I hope*).

75) [Per queste ragioni voglio ringraziare per il lavoro [da loro svolto] i parlamentari [che hanno partecipato al comitato di conciliazione,] [**augurandomi** che questo sia il punto di partenza per una nuova politica nel campo delle tossicodipendenze.] <ep-96-12-09.txt:29>

Litt. English L2 of Italian L1: *[For these reasons, I should like to thank for the work [that has been done] those honourable Members] [who formed part of the Conciliation Committee] [hoping that this is the starting-point for a new policy on drug dependence.]*

Official Danish L2: *[Jeg vil derfor sige tak for det arbejde,] [der er udført af de parlamentsmedlemmer,] [der har deltaget i Forligsudvalget,] [**idet jeg håber**, at dette bliver udgangspunktet for en ny politik inden for narkotikamisbruget.]*
Litt. Danish L2: *[as I hope that this will be the starting-point for a new policy on drug dependence.]*

Similarly, I also find it reasonable to assume that some of the English and Italian realisations at level (h) correspond to the relative clauses (e) in Danish. The percentage differences at level (e) are not as considerable between Danish, English and Italian (49.4 % – 39.3 % – 43.9 %) as those between the languages at level (d) (37.6 % – 21.9 % – 19.1 %), but the difference between Danish and English is nevertheless noteworthy. Just as relative clauses, the different types of modifiers are typically used to elaborate on noun phrases, contributing some kind of additional information. In 76), an English L1 sentence is contrasted with a Danish L1 sentence in 77) to

illustrate this point. The English example contains a participial modifier (*proposed*), which in a comparable Danish L1 construction, has been realised as a relative clause (*der er foreslået/which has been proposed*).

76) [It is to say the least disappointing, that some Member States were prepared to conceal the real incidence of BSE in their national herds and equally unwilling to implement the control measures] [**proposed** by the Commission] ... <ep-01-02-01.txt:12>

77) [Vi er enige i den tidsplan,] [**der er foreslået**,] ... <ep-99-11-18.txt:143>
   *[We agree to the timetable] [**which has been proposed**] ...*

This tendency is also found when translating from English into Danish: English (and also Italian) modifiers are transformed into relative clauses in the Danish L2 corpus, see example 78) where *applied* turns into *der anvendes* (*that are applied*). Here, the Danish L2 is the official EU translation.

78) [We also fully support the view that the governing principle must be that of international obligations,] [fairly and humanely **applied**.] <ep-02-09-23.txt:62>
   *[Vi støtter også fuldt ud det synspunkt, at det ledende princip skal være internationale forpligtelser,] [**der anvendes** på en retfærdig menneskelig måde.]*
   *Danish L2: ... [**that are applied** in a fair and humane manner.]*

As mentioned in Chapter 3, restrictive relative clauses and participial modifiers are not always treated as EDUs in the literature, because they rarely express rhetorical relations other than Elaboration. And since they have very low frequencies in the Danish texts, it could be argued that these two types of constructions should not have been included in the present analysis. However, as we shall see in Chapter 5, a small number of restrictive relative clauses across the three languages express rhetorical relations other than Elaboration of their matrix EDUs. Therefore, it was argued that it is relevant to consider both restrictive relative clauses and participial modifiers as EDUs. Furthermore, EDUs with modifiers do not always correspond to restrictive relative clauses but also to appositive relative clauses, which suggests that at least modifiers should be included in the EDU segmentation. In the following example from the Italian part of the corpus, a past participial modifier (*firmato/signed*) elaborates on a noun phrase in the matrix EDU (*un appello/a petition*). In the official English L2 reproduced below, the

original Italian L1 participial modifier has been translated with an appositive relative clause (*which also bears*), as argued above.

79) [Prego i colleghi di firmare un appello] [che sto facendo circolare,] **[firmato** anche da altri intellettuali europei.]** <ep-00-11-13.txt:31>

*[I would ask the Members to sign a petition] [which I am circulating,] [**which also bears the signatures** of other European intellectuals.]*

4.6 Summing up

This chapter has addressed the linguistic structure of parliamentary discourse from a text structure point of view. Text structure, defined as the syntactic relations between intrasententially related EDUs, was studied employing two parameters. The first was dependency, understood as syntactic coordination and subordination of EDUs. The second was morphological realisation, understood as the explication of the main verb in an EDU, which can be realised as either finite, non-finite or not realised at all. To deal with these two parameters, a so-called deverbalisation scale was introduced covering the various syntactic constructions found in the corpus, from independent sentences with all grammatico-semantic characteristics at their disposal to highly integrated non-finite and verbless constructions with hardly any syntactic independence left. A hypothesis was presented predicting that Danish would have more independent textualisations than Italian, and that English would be closer to Danish than to Italian.

The results of the analysis of the first parameter, dependency, indicated that the proposed hypothesis had to be revised. English converged more towards Italian than towards Danish with regards to coordination and subordination. In fact, the English texts contained fewer coordinated constructions than Italian. It was argued that one reason for this could be the significantly longer Italian sentences which open up the possibility for more EDUs to include both coordination and subordination, whereas the English sentences are so short that they rarely contain both types of dependencies. As this discovery came as a surprise, I decided to examine the subordinate constructions in more detail. It turned out that there was actually a difference between English and Italian, as the Italian subordinated EDUs more frequently displayed a higher degree of subordination than their English counterparts, meaning that a higher number of subordinated EDUs had other subordinated EDUs linked to them. In this regard, English and Danish were more closely related. Thus for the first parameter, my hypothesis was only partially confirmed.

In the analysis of the second parameter, realisation, the results showed, again, a very close relationship between English and Italian. Both languages tend to realise a larger number of EDUs as non-finite constructions than Danish, in particular as gerund, present and past participles and different types of modifiers. Again, English exceeded Italian in terms of frequencies of non-finite textualisations. However, all three languages had more or less the same numbers of verbless constructions.

To sum up, Danish typically uses coordination and finite textualisation of EDUs. English and Italian tend to do the opposite, through a higher frequency of subordination and non-finite textualisations of EDUs. Italian text structure, in contrast to English, stands back as the language with the most hierarchical text structure, as the combination of more subordinated EDUs and non-finite realisations inside the same sentences entails a deeper text structure than those of English and, in particular, Danish. In Chapter 5, it shall be considered what these differences in text structure entail for the discourse structure of the three languages.

5   The discourse structure of parliamentary discourse

This chapter furthers the study of linguistic structures in parliamentary discourse through the analysis of discourse structure and, in particular, rhetorical relations. The previous chapter examined the syntactic dependencies of the EDUs and the realisations of subordinate EDUs on the surface of the syntax. The present chapter tackles underlying relations between the EDUs which together constitute the discourse structure of a text.

As already mentioned at the very beginning of this thesis, I will refer to the underlying relations of a text as rhetorical relations. I do this because rhetorical relation is the term used in Rhetorical Structure Theory (RST) (Mann & Thompson, 1988), the theory adopted here for analysing discourse structure. Relations of this type have been the topic of studies for decades and have been given different names. Longacre and colleagues (D. L. Ballard, Conrad, & Longacre, 1971) were some of the first to propose a taxonomy for the 'deep structure of interclausal relations' (cf. also Longacre, 1983). Later, competing taxonomies were created and other terms for the relations were proposed. Grimes (1975) refers to the relations as 'rhetorical predicates', others (Conte, 1999; Hobbs, 1985; Kehler, 2002; Sanders, Spooren, & Noordman, 1992; Stede, 2011) as 'coherence relations', Hoey (1983) as 'clause relations', Halliday & Hasan (1976) as 'logico-semantic relations', Martin (1992) or 'conjunctive relations', and, as stated above, Mann & Thompson (1988) use the label 'rhetorical relations'. More recently, and especially within the field of computational and formal linguistics, the more general term of 'discourse relations' is used as a theory-neutral descriptor (even when relations are part of a specific discourse theory) (Asher & Lascarides, 2003; Asher, 1993; Buch-Kromann & Korzen, 2010; Prasad et al., 2008; Renkema, 2009b). As noted by Rigotti (2009, p. 432), the choice of terminology seems to be much motivated by the nature of the study and the theoretical stances applied, exactly as we saw with the definitions of EDUs in Chapter 3. Roughly speaking, rhetorical relations are used when discourse analysts study the global structure and functions (hence the term rhetorical) of a text. Coherence relations when analysts examine how the clauses in a text have been combined to form a whole. Conjunctive relations are used when the syntactic description is extended to the discourse level, and discourse relations when discourse structure is viewed from a discourse semantics perspective. Stede (2011, p. 85), who uses the term coherence relations, defines the relations as a 'specific relationship, holding on the semantic or the pragmatic level of description, between adjacent units of text.' This definition broadly covers the use of rhetorical relations in the present study, although the term 'adjacent' should be

not interpreted in a narrow sense meaning two neighbouring units, but rather as two units within the same text.

The notion of rhetorical relations has evolved from a theory of discourse structure which is based on a number of assumptions that would appear to be rather intuitive. Hobbs (1985, p. 1), for example, states in the very first lines of his paper on coherence and discourse structure that discourse, indeed, has structure.

> Let us begin with a fact: discourse has structure. Whenever we read something closely, with even a bit of sensitivity, text structure leaps off the page at us. We begin to see elaborations, explanations, parallelisms, contrasts, temporal sequencing, and so on. These relations bind contiguous segments of text into a global structure for the text as a whole.

As such, three basic assumptions are widely recognised by discourse analysts. Firstly, that 'all work on discourse starts from the premise that discourse meaning is more than the sum of its parts' (i.e. the EDUs) (Forbes et al., 2002, p. 261). Secondly, that the structure of this discourse is 'a hierarchical structure' of relations binding the units together (Asher & Vieu, 2005, p. 591). And thirdly, that these relations can be assigned different semantic or pragmatic sources of coherence (Mann & Thompson, 1988). It is not within the scope of this thesis to test whether these assumptions actually hold, but I shall discuss them in the sections below in terms of RST.

There are at least three important challenges related to the notion of rhetorical relations. The first issue regards the very nature of rhetorical relations: what is their actual purpose? Different views can be found in the literature. Hobbs (1990, p. 23) argues that rhetorical relations serve the purpose of 'text-building strategies' that writers use to make the reader's comprehension easier. A similar view is found in Knott & Sanders (1998, p. 138), who see rhetorical relations as cognitive mechanisms that writers draw upon to join pieces of text together, and that readers recognise when interpreting those pieces. Thus, they perceive the determining of rhetorical relations in a text as 'part of the process of understanding it'. A different view is found in Grosz & Sidner (1986), who do not regard rhetorical relations as essential elements for a successful writer-reader interaction. Instead they propose that the reader does not need to understand which rhetorical relations are present in a text in order to understand it. They argue that rhetorical relations are primarily used as analytical tools to describe discourse structure. This viewpoint has later also been found in Traum (1993, p. 133), who points out that readers do not need to

recognise the specific rhetorical relations between all units, as long as the reader and writer agree on the overall purpose of the text. In this thesis, I follow the former group of scholars by viewing rhetorical relations as psychological or cognitive entities that are used to make the communication successful. This view is also found in RST and is in line with the discourse-as-process view presented in Chapter 2.

The second issue concerns the number of rhetorical relations to include in an analysis. Stede (2011, pp. 82–83) lists six perspectives from the literature devoted to creating inventories of rhetorical relations, all of which seem to be related to each other: (1) be sceptical and parsimonious, (2) resort to insights from philosophy, (3) be inspired by the lexicon of your language, (4) be motivated by syntax and semantics, (5) try to explain human cognition, (6) be inspired by authentic texts. Following one or a combination of these approaches has resulted in inventories ranging from sets with over 100 relations (Hovy & Maier, 1992; Martin, 1992) to sets with just two relations (Grosz & Sidner, 1986). However, the most dominant theories in the current literature typically contain 20-30 relations, e.g. Segmented Discourse Representation Theory (SDRT) (Asher & Lascarides, 2003) and the different versions of RST (Carlson et al., 2003; Mann & Thompson, 1988; Mann, 2005). How many relations should be included depends to a large extent on the nature of the study: if the aim is to create a corpus with high quality and reliable annotation, the number of relations must be limited, possibly grouped together in larger classes with subtypes and variants. This style is mainly found in computational approaches to discourse structure (Buch-Kromann & Korzen, 2010; Carlson et al., 2003; Kehler, 2002; Martin, 1992; Prasad et al., 2008). In the Penn Discourse Treebank (Prasad et al., 2008), for example, the Contingency class covers the subtypes of Cause and Condition in their semantic and pragmatic variations. In this way, the analyst can annotate the relation class instead of the relation subtype if in doubt which one to choose. Classes also exist in RST, although with a different purpose which has nothing to do with obtaining high inter-annotator scores. In RST, the idea behind the relation set is to create a manageable and flexible number of relations which occur in natural language, independent of text types and genres. Different approaches to RST have yielded different lists of rhetorical relations. The 'classical' set from Mann & Thompson (1988) included 24 relations, but since the relation inventory was not intended as a closed set of relations (Mann & Thompson, 1988, p. 250), more recent applications of RST have extended this to 31 relations in what has been termed the 'modern' or 'extended' RST version (Mann, 2005; O'Donnell, 2004). An even further extension of the RST inventory can be found in the RST Treebank project (Carlson et al., 2003; Carlson & Marcu, 2001) which comprises a total of

78 relations including Attribution; see the discussion in Section 3.5.2.1. Here, the relations are not explicitly divided into semantic and pragmatic sources of coherence (though this easily could have been done) but into classes as mentioned above, and into multinuclear (25) and mononuclear (53) relations. In this thesis, I apply the extended version of RST relations from the RST website (Mann, 2005), as an earlier pilot study of mine employing the relation set of the RST Treebank project (Carlson et al., 2003) showed that annotation difficulties increased and inter-annotator agreement weakened. However, the annotation manual of the RST Treebank (Carlson & Marcu, 2001) has been used as an additional source for the following analyses, as the manual contains more exhaustive relation definitions than the RST website does, decision strategies in cases of troublesome relations, and numerous clear-cut examples. Figure 5.1 shows the rhetorical relations used in the analyses of the following sections. The upper part of the figure contains the subordinate mononuclear relations where the satellite activates the semantic or pragmatic relation to the nucleus. The lower part of the figure includes all the coordinate multinuclear relations. While some of the relation names are self-explanatory, other relation names are less straightforward. In the following sections, I discuss the most central rhetorical relations to this study. Moreover, all relation definitions can be found in Appendix B and on the RST website (Mann, 2005).

The third issue has already been introduced in Chapter 3 and regards the problem of identifying rhetorical relations. We saw that it is somewhat difficult to reach high agreement numbers in this area of linguistics, probably due to the two above-mentioned issues. While I shall refrain from discussing this issue more in detail, it is important to mention that the results considered in this chapter are based on RST annotations, which essentially must be seen more as 'plausible judgements' (Mann & Thompson, 1988, p. 246) than as definite, objective observations, see also the discussion in Section 3.4.1 on inter-annotator agreement and own annotations.

	Semantic *(subject matter)*	Pragmatic *(presentational)*
Mononuclear	Circumstance	Antithesis
(Nucleus and satellite)	Condition	Background
	Elaboration	Concession
	Evaluation	Enablement
	Interpretation	Evidence
	Means	Justify
	Otherwise	Motivation
	Purpose	Preparation
	Solutionhood	Restatement
	Unconditional	Summary
	Unless	
	(Non-)Volitional Cause	
	(Non-)Volitional Result	
Multinuclear	Disjunction	
(all semantic)	Conjunction	
	Contrast	
	Joint	
	List	
	Multinuclear Restatement	
	Sequence	

Figure 5.1: Rhetorical relations employed

As a result of the many competing theories of discourse structure and inventories of rhetorical relations, some scholars have tried to combine different theories. Bateman & Rondhuis (1997), who also provide a very interesting discussion and comparison of Discourse Representation Structure (DRS) (Asher, 1993), RST (1988) and Martin's (1992) conjunctive relations, argue in favour of using a stratified version of DRS with RST and Martin's conjunctive relations to account for discourse structure in a better way than the individual theories appear to be capable of on their own. However, the combined approach is not without complications and has only found little use in subsequent literature. Similarly, the work of Hobbs (1979) has been continued by various other scholars such as Kehler (2002) and Wolf & Gibson (2005), who have all altered

Hobbs' original relations. Finally, Stede (2008a, 2008b) combines Martin's conjunctive relations with a thematic structure (i.e. a hierarchical structure showing (sub-)topic shifts), a referential structure (i.e. co-reference relations) and an intentional structure (i.e. 'deep' support relations between texts segments and their illocutions) in a multi-level analysis, which he argues is easier to use and more descriptive than RST. In this thesis, the aim is not so much to assess and test RST as to employ the ideas and methods of the theory. In addition to using the ideas of RST, I expand the study of discourse by investigating the two related linguistic structures of text structure and information structure and by adding to English two other languages: Danish and Italian.

5.1 Rhetorical Structure Theory

Rhetorical Structure Theory defines itself as 'a descriptive theory of a major aspect of the organisation of natural text' (Mann & Thompson, 1988, p. 243). The theory constitutes a functional approach devised to explain and analyse textual coherence and text organisation in general. As any other theory of discourse structure, the theory accounts for how the discourse units are identified and represented, which rhetorical relations are considered to hold between the units, and how the structure of discourse is represented graphically, that is, as a tree structure or the like.

As for the first element, we have already presented the notion of EDUs as employed in this thesis in Chapter 3. In RST, discourse units are seen as linguistic entities and are considered essentially to be clauses and, in some cases, also phrases without an explicit subject or verbal element. This definition of EDUs corresponds in many ways with that of the original proposal by Mann & Thompson (1988), although I have added certain syntactic constructions like restrictive relative clauses.

As for the second element, Figure 5.1 shows the rhetorical relations included in this analysis of discourse structure. Even though the set of relations, as argued above, is not a closed set, in order for an analysis to be operational, the set must be closed while the texts under consideration are annotated. For this study, the aim has not been to increase or reduce the number of relations, but to investigate how the extended set of rhetorical relations in RST is used in terms of frequency and textualisations across three different languages.

To understand the concept of rhetorical relations in RST, we need to recognise the notion of nuclei and satellites as the two basic types of parts in texts. A nucleus and a satellite are two non-overlapping text spans between which a rhetorical relation exists. The nucleus is essential

and cannot be omitted while the satellite is optional and is claimed to be dispensable (see Stede, 2011, p. 115; Taboada & Mann, 2006a, p. 427). The rhetorical relations in Figure 5.1 all have definitions consisting of four fields (Mann & Thompson, 1987, p. 8, 1988, p. 249):

- constraints on the nucleus
- constraints on the satellite
- constraints on the combination of nucleus and satellite
- the effect

In order to specify each field for the rhetorical relations encountered, the discourse analyst must make plausible judgments, as argued in the previous section, based on the co-text and the intentions of the writer. This means that it is the analyst that assesses whether it is plausible that the writer has such intentions or desires that 'effect' (see the fourth field above) when creating the text. To understand the fields given above and how they combine to define and characterise a relation, let us have a closer look at three definitions for rhetorical relations in RST: Purpose (semantic) in Figure 5.2; Concession (pragmatic) in Figure 5.3; and Contrast (multinuclear) in Figure 5.4.

As we can see from Figure 5.2, the rhetorical relation between the two EDUs is interpreted through the presence of a discourse cue *in order to* which typically expresses a purpose relation. The nucleus and satellite also correspond to a matrix clause and a subordinate clause. In this case, there is, then, correspondence between syntactic subordination and rhetorical subordination.
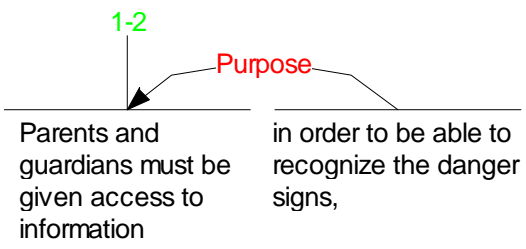
Relation name	*Purpose*
Constraints on nucleus	The nucleus is an activity.
Constraints on satellite	The satellite is a situation that is unrealised.
Constraints on the combination of nucleus and satellite	The satellite presents a situation to be realised through the activity in the nucleus.
The effect	The reader recognises that the activity in the nucleus is initiated in order to realise the satellite.
Text example	[Parents and guardians must be given access to information] [in order to be able to recognise the danger signs,]... <ep-96-12-09.txt:33>
Graphical representation	

Figure 5.2: Definition of the Purpose relation
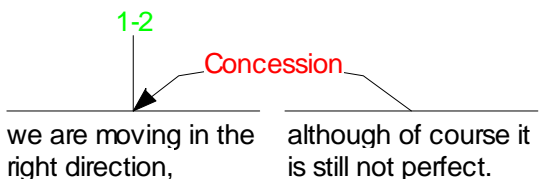
Relation name	*Concession*
Constraints on nucleus	The writer has positive regard for the situation presented in the nucleus.
Constraints on satellite	The writer is not claiming that the situation in the satellite does not hold.
Constraints on the combination of nucleus and satellite	The writer acknowledges a potential or apparent incompatibility between the situations presented in the nucleus and satellite. The writer regards the situations presented in the nucleus and the satellite as compatible; recognising the compatibility between the situations presented in the nucleus and satellite increases the reader's positive regard for the nucleus.
The effect	The reader's positive regard for the situation presented in the nucleus is increased.
Text example	… [we are moving in the right direction,] [although of course it is still not perfect.] <ep-01-11-12.txt:56>
Graphical representation	1-2 Concession we are moving in the right direction, although of course it is still not perfect.

Figure 5.3: Definition of the Concession relation

As mentioned, relations are divided into two main types, semantic and pragmatic relations, depending on the intended effect on the reader. In semantic relations such as the Purpose relation in Figure 5.2, the intended effect is that the reader recognises the relation between the nucleus and satellite in question. In pragmatic relations such as the Concession relation in Figure 5.3, the intention is to reinforce some inclination in the reader, *in casu* a positive regard for the statement that *we* (the EU Parliament) *are moving in the right direction*.
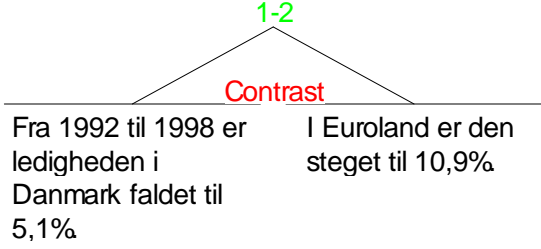
Relation name	Contrast
Constraints on nucleus	Multinuclear.
Constraints on the combination of nucleus and satellite	No more than two nuclei. The situations in these two nuclei are (a) comprehended as the same in many respects (b) comprehended as differing in a few respects and (c) compared with respect to one or more of these differences.
The effect	The reader recognises the comparability and the difference(s) yielded by the comparison being made.
Text example	[Fra 1992 til 1998 er ledigheden i Danmark faldet til 5,1%.] [I Euroland er den steget til 10,9%.] <ep-00-03-13.txt:27> *[Between 1992 and 1998, unemployment in Denmark fell to 5.1%.] [In the Eurozone, it increased to 10.9%.]*
Graphical representation	

Figure 5.4: Definition of the Contrast relation

The Contrast relation is one of the few symmetric relations in RST, referred to as multinuclear relations. These relations connect two or more nuclei which are considered to be equally important for the writer's purpose. The relations are not categorised as semantic or pragmatic by RST, but fall mainly within the semantic source of coherence, as they tend to connect the content of the EDUs in question rather than influence the beliefs and desires of the reader. Unlike the Purpose and Concession text examples, the Contrast text example shows that the notions of nucleus and satellite are not necessarily related to any syntactic factors or the signalling of a discourse cue; it is only the logical and propositional transition from one sentence to another which signals the relation. In this example, the contrast becomes clear when opposing the verbs of the two EDUs, *er faldet* (fell) and *er steget* (increased). All definitions of RST relations are based on functional and semantic criteria, not on morphological or syntactic signals as no reliable signals for any relations were found. As such, the apparently contrastive coordinator *but* can signal Contrast, Concession and Antithesis (see Stede, 2011, p. 104), along

with many other relations, especially in cases where the relation spans across several EDUs and sentences.

Concerning the last of the three elements presented at the beginning of this section, the form of the discourse structure, the three figures above also contain a graphical representation of the text examples, in the final row. The form of the discourse structures in RST is called a 'schema' when the relation connects a small number of text spans; in the examples above, these have been numbered 1 and 2 in the graphical representations. The straight lines in the schemas represent nuclear spans, and the curves denote the rhetorical relation. There are different kinds of schemas, which can all be found in Mann & Thompson (1988, p. 247). A text, which typically consists of more than two EDUs, is said to have a structure that can be built up in a recursive manner by connecting the EDUs with rhetorical relations in text spans. In this way, a text annotated with RST relations represents a tree structure. The tree accounts for the hierarchical character of the text and 'keep[s] record of the history of generated discourse structure' (Irmer, 2011, p. 133), permitting a structural distinction between rhetorically coordinated and subordinated EDUs. Tree-like structures are also used in a number of other theories (Buch-Kromann & Korzen, 2010; Carlson et al., 2003; Grosz & Sidner, 1986; Polanyi, 1988). Other graphical representations are sequences, stacks and graphs (see Irmer, 2011, pp. 131–141 for a discussion of these). The graphical representations of discourse structures play a minor role in this thesis, and consequently no further detail will be provided.

I chose RST as the theoretical framework for analysing discourse structure for three reasons. The first can be found in the cross-linguistic nature of the thesis. Even though RST was originally developed and applied to English, the theory has been applied to other languages and it has been claimed to be language-independent. Taboada & Mann (Taboada & Mann, 2006b) cite a number of languages other than English to which RST has been applied, including Chinese, Portuguese, French, Dutch, German, Arabic, Finnish, Japanese, Russian and Spanish. This means that not only was RST expected to be capable of analysing discourse structure in Danish and Italian, but also that the results of this thesis can be compared to results from studies in other languages. I shall do so in the following sections.

The second reason for viewing RST as an appropriate theory for investigating discourse structure cross-linguistically is that the theory, as argued by Irmer (2011, p. 127), is more interested in the various forms discourse structure can take than in the meaning of texts. This holds true not only for the *global* structure of the text, that is, the overall form of all the discourse units (the macrostructure), but also for the form of the *local* structure of the individual

discourse units, as we saw in Chapter 4 on text structure and syntactic relations. In this way, Irmer (ibid.) categorises RST within the group of 'discourse syntactical' accounts together with Forbes et al. (2002) and Polanyi (1988), to whom we may add Halliday & Hasan (1976), Martin (1992) and Sanders et al. (1992). The other groups of discourse theories are labelled 'informational' accounts when rhetorical relations are related to the information contained in discourse units (Asher & Lascarides, 2003; Hobbs, 1985; Kehler, 2002; Wolf & Gibson, 2005), 'intentional' accounts when theories take the intentions of writers into consideration (Grosz & Sidner, 1986; Poesio & Traum, 1997), and 'discourse topic' accounts when discourse structures are derived from questions assumed to be underlying texts (Klein & Stutterheim, 1987; Kuppevelt, 1995). A discourse syntactical account is essential for this thesis, because one of the aims of the study is to examine different linguistic structures simultaneously, from lower level features to higher level features.

The third reason is found in one of the early papers of RST (Mann & Thompson, 1987, p. 3), which states 'RST lays the foundation for studies in contrastive rhetoric', currently known as intercultural rhetoric. In a retrospective paper by Taboada & Mann (2006a, p. 424), intercultural rhetoric is also cited as one of the disciplines where RST has been employed successfully. Consequently, it is assumed that the RST analyses are able to reveal linguistic, rhetorical and contextual differences between the three languages under investigation.

RST is probably one of the most used theories for analysing discourse, but it has also received considerable criticism. I have already discussed some of the concerns raised by scholars in the previous section: the number of rhetorical relations to include, the subjectivity of the analyses and the issues concerning graphical representation. In the following two sections, I shall return to the shortcomings of the theory as they become relevant for my analyses.

5.1.1 RST in argumentative texts

This study follows a number of other studies that have used RST for analysing the discourse structure of argumentative text types. The theory has been claimed by some scholars to be the most adequate text theory for analyses of argumentative texts (Azar, 1999, p. 97). Inside RST, Mann & Matthiessen (1991, p. 235) also state that the nucleus-satellite distinction is 'crucial to a study of persuasion texts', and Abelen et al. (Abelen et al., 1993) see RST as 'especially useful for the analysis of persuasive language, because it allows, and even forces, the analyst to consider the intended communicative effects expressed in, or plausibly inferable, from the text.'

The use of the word 'rhetorical' in the name of the theory suggests that the theory is much concerned with the argument structure of texts. Azar (1999, p. 112) concludes his paper by stating that five of the pragmatic rhetorical relations in RST – Evidence, Justify, Motivation, Antithesis and Concession – are congruent with the concept of argument relation, that is, the four types of arguments: supportive, incentive, justifier and persuader. In the same manner, Taboada (2004b, p. 115) discusses the view of seeing RST as a "theory of influence", by drawing attention to the pragmatic relations, 'whose intended effect is to increase some inclination in the reader' (Mann & Thompson, 1988, p. 257); cf. Mann et al. (1992, pp. 44–45):

> Text structuring relations are functional; the character that they all share can be stated in terms of the categories of effects that they produce. They can be described in terms of the purposes of the writer, the writer's assumptions about the reader, and certain propositional patterns in the subject matter of the text. The text structuring relations reflect the writer's options of organisation and presentation; it is in this sense that a RST structure is "rhetorical".
>
> In contrast, one could assume that text structuring relations simply represent relations on the subject matter (e.g. of succession, cause or conditionality).

However, some scholars do not entirely accept the use of the word rhetorical, as they consider that 'this is a somewhat skewed notion of "rhetoric", though, and prefer the more neutral term "coherence-relational structure". After all, many of the relations involved are semantic or "informational" in nature and have little or no connection to rhetorics' (Stede, 2011, p. 83 footnote 2). I shall show that the writers of the texts in the corpus employed here have built up the texts in different ways to achieve an effect and that RST is capable of showing this, thus justifying the label 'rhetorical'.

The rhetorical strategies and styles of the texts studied in this thesis may not be the same as they would have been in longer texts, due to the time constraints and the occasional impossibility to revise the material 'on the spot', but we can still observe interesting choices of organisation on the part of the writer. The main argument, or arguments in cases where the writer presents more than one, is considered the main nucleus (cf. Azar, 1999, pp. 111–112); typical purposes are reflection, understanding, agreement, engagement, change of attitude, or call for action. The technical RST term for this is the 'Comprehensive Locus of Effect', see Mann et al. (1992, p. 61): the 'portion of the text that represents the essence of the text as a

whole'. In the next section, I shall present the overall results of the analysis of discourse structure in parliamentary discourse, focusing particularly on the rhetorical hierarchy of the EDUs, which can be seen as a continuation of the discussion of syntactic versus rhetorical subordination of EDUs from the previous chapter, and also concentrating on the two different sources of coherence in RST relations: semantic and pragmatic.

5.2   Rhetorical relations in parliamentary discourse

This section describes the results of the analyses of rhetorical relations in each of the three languages. The focus is first on the sources of coherence the relations express and subsequently on the rhetorical hierarchy between EDUs. These two foci shall be elaborated on through a listing of the relations found in the corpus together with appropriate examples. The units of analysis are again the EDUs which were also used in the previous chapter on text structure and initially defined and discussed in Chapter 3. The following listing of relations contains the number and types of rhetorical relations that connect EDUs. As each individual text in the corpus contains one main EDU that is not subordinated to the other EDUs,[8] see the section above, the total number of relations found is smaller than the total number of EDUs, cf. Table 4.1.

Table 5.1 shows the rhetorical relations found in the corpus, for Danish, English and Italian respectively. As can be seen from the table, the results for all three languages are strikingly similar in terms of relations and their distribution. Elaboration shows the highest number of occurrences, followed by Conjunction and a number of other relations with a frequency around 5 % of the total distribution. In the next few sections, I provide some details and examples of the ten most frequent relations from the three languages. The sections also anticipate the next chapter on information structure as I briefly comment on the signalling of the relations.

---

[8] In cases where the main EDU is rhetorically coordinated with another EDU, the first of these in terms of chronology is considered the main EDU.

		Danish		English		Italian	
		N	**%**	**N**	**%**	**N**	**%**
SEMANTIC RELATIONS	Circumstance	73	5.2	82	6.5	48	3.4
	Condition	36	2.6	25	2.0	28	2.0
	Elaboration	327	23.4	334	26.3	377	26.5
	Evaluation	23	1.6	22	1.7	31	2.2
	Interpretation	23	1.6	28	2.2	21	1.5
	Means	21	1.5	26	2.0	28	2.0
	Non-Volitional Cause	64	4.6	50	3.9	77	5.4
	Non-Volitional Result	21	1.5	25	2.0	27	1.9
	Otherwise	0	-	0	-	1	0.1
	Purpose	44	3.2	41	3.2	60	4.2
	Solutionhood	44	3.2	26	2.0	26	1.8
	Unconditional	3	0.2	6	0.5	1	0.1
	Unless	2	0.1	7	0.6	1	0.1
	Volitional Cause	61	4.4	35	2.8	60	4.2
	Volitional Result	27	1.9	12	0.9	25	1.8
PRAGMATIC RELATIONS	Antithesis	17	1.2	21	1.7	15	1.1
	Background	70	5.0	77	6.1	71	5.0
	Concession	71	5.1	56	4.4	82	5.8
	Enablement	2	0.1	4	0.3	2	0.1
	Evidence	60	4.3	43	3.4	45	3.2
	Justify	48	3.4	45	3.5	70	4.9
	Motivation	16	1.1	20	1.6	24	1.7
	Preparation	9	0.6	6	0.5	3	0.2
	Restatement	11	0.8	2	0.2	5	0.4
	Summary	26	1.9	24	1.9	19	1.3
MULTINUCLEAR RELATIONS	Disjunction	9	0.6	1	0.1	5	0.4
	Conjunction	177	12.7	169	13.3	181	12.7
	Contrast	27	1.9	21	1.7	31	2.2
	Joint	15	1.1	21	1.7	3	0.2
	List	37	2.7	13	1.0	22	1.5
	Multinuc. Restatement	0	-	2	0.2	2	0.1
	Sequence	31	2.2	25	2.0	30	2.1

Table 5.1: Number and percentages of rhetorical relations

5.3 The five most frequent semantic relations

In RST, 22 of the 32 rhetorical relations come from semantic sources of coherence. As already mentioned, semantic relations connect the situations described in the EDUs, in contrast to the pragmatic relations that to a greater extent include the writer-reader relationship. Scholars across theories have found that, in most text types, the semantic source of coherence is the predominant one, even though argumentative text types exhibit higher percentages of pragmatic relations than narrative and expository types (Sanders, 1997, p. 138; Stede, 2004).

The relations found in the present corpus confirm these previous observations, with more than 75 % of the total number of relations emanating from a semantic source of coherence, in all languages. Again, the similarities in the distribution across Danish, English and Italian are striking. The only minor difference is between Danish/English and Italian mononuclear and multinuclear relations, where Italian has a higher number of mononuclear relations (57.1 %) than Danish/English (55.1/56.7 %). This could be due to the difference between intrasententially and intersententially related EDUs found in Table 4.1.

	Danish	English	Italian
Semantic	76.3 %	76.6 %	76.4 %
- *Mononuclear*	*55.1 %*	*56.7 %*	*57.1 %*
- *Multinuclear*	*21.2 %*	*19.9 %*	*19.3 %*
Pragmatic	23.7 %	23.4 %	23.6 %

Table 5.2: Distribution of sources of coherence

The reason for the overall distribution of semantic versus pragmatic relations is not only that 22 out of the 32 relations in RST are semantic, but also that semantic relations to a larger extent than pragmatic relations are motivated by discourse cues from the syntax such as subordinators; an EDU containing a subordinator such as *because* often expresses a semantic causal relation. As could be seen in Table 4.1 containing the overall distribution of EDUs in the corpus, we saw that all three languages showed higher tendencies to use intrasententially than intersententially related EDUs. As a result, it makes good sense that all languages exhibit more semantically motivated relations, as semantic relations more often than pragmatic relations are found inside the sentence boundary in relative clauses and subordinate finite adverbial clauses; cf. Table 5.6, 5.7 and 5.8. Furthermore, it also makes sense from a rhetorical perspective: all texts in the corpus build up one or more arguments, and for this to be done in a persuasive and effective way, the writers also need to elaborate on the subject matter of their arguments through the causal relations of RST (e.g. Cause, Condition, Purpose and Result) together with attitude relations (Evaluation and Interpretation). However, the two relations that skew the distribution between semantic and pragmatic relations are Elaboration and Conjunction (cf. also Knott & Sanders, 1998; Knott, 1996; Korzen & Müller, 2011; Korzen, 2010, 2011). Elaboration and Conjunction together account for almost 40 % of all the relations found in the corpus. While one could argue that this overrepresentation of Elaboration and Conjunction does not tell us much about the discourse structure of a text because the two relations can be considered quite 'empty'

in terms of semantics, I will argue that the intensive use of Elaboration and Conjunction plays an important – perhaps even universal – role in text-building strategies in supporting and conjoining parts of arguments. Below, these two relations will be discussed in more detail, together with other frequently used relations. In all examples, the relevant satellites are underlined and relevant linguistic material is shown in bold-faced type.

5.3.1 Circumstance

Circumstance satellites are mostly found inside the sentence boundary in the texts of the corpus. According to the definition of the relation (see Appendix B), they set a temporal or spatial framework within which to interpret the nucleus. The first two examples show that the Danish temporal subordinator *når* (when) in 80) and the English subordinator *when* in 81) clearly express the circumstantial framework for their respective nuclei. In both cases, the nucleus-satellite distinction is highly visible, as both nuclei contain the writers' desires (*keep the Danish krone* and *we need action on jobs*).

80) [Min gruppe håber, at de danske vælgere beslutter at bevare den danske krone,] [**når** de nu skal til folkeafstemning om ØMU-tilslutning den 28. september i år.] <ep-00-03-13.txt:27>

*[My group hopes that Danish voters will decide to keep the Danish krone] [**when** they take part in the referendum on membership of EMU on 28 September of this year.]*

81) [**When** the European finance ministers meet next week,] [we need action on jobs.] <ep-96-07-03.txt:75>

Although most of the Circumstance satellites are signalled by temporal subordinators in the corpus employed here, Circumstance has no reliable signals, as noted by Mann & Thompson (1987, p. 49). In the examples below, it can be observed that the relation can also be found in syntactic constructions with non-finite verb forms; in examples 82) and 83) as nominalisations (*meddling* and *opposition*) accompanied by discourse cues (*After* and *Ever since*).

82) [**After** 30 years of **meddling** by the Commission and in-fighting by Member States,] [what we now have is a world-class ecological disaster and the near destruction of many parts of the British fishing industry.] <ep-02-05-29.txt:37>

83) [**Ever since the opposition** to the THORP nuclear plant in 1993,] [the level of radioactive discharges from nuclear operations in Cumbria has substantially increased.] <ep-99-09-13.txt:59>

By textualising the Circumstance satellite as a nominalisation, or below in the two Italian examples as past participial clauses (*detto/said*) in 84) and modifiers (*svoltasi/held*) in 85), the writer to a greater extent highlights the satellite status of an EDU, cf. the deverbalisation scale in Figure 4.1., as compared with the lesser dependent subordinate finite clauses in 80) and 81).

84) [**Detto** questo,] [chiediamo che tale comitato rifletta, nel titolo e nel mandato, entrambi i concetti della sicurezza marittima e della prevenzione dell'inquinamento.] <ep-01-02-12.txt:64>

[That said,] [we would wish both the title and the mandate of such a committee to encompass the two concepts of maritime safety and the prevention of pollution.]

85) In occasione di una riunione dell'UNESCO, [**svoltasi** qualche anno fa a Casablanca,] si era posto il problema di immaginare che l'intero arco alpino diventasse patrimonio dell'umanità. <ep-01-01-15.txt:80>

[At a UNESCO meeting, [held some years ago in Casablanca,] the issue was raised of whether to make the entire Alpine Region a World Heritage Site.]

5.3.2 Elaboration

Elaboration is, by far, the most frequent relation: approximately 25 % of all relations found across the three languages were Elaborations. In contrast to other relations, Elaboration has no salient context, that is, it is present in many different co-texts (cf. Kapogianni, 2011, p. 65), from adding detail about the situation to specifying some element of the subject matter presented in the nucleus. In the literature, various approaches to dividing Elaboration into different subtypes can be found. Carlson & Marcu (2001), for example, add another three subtypes (Addition, Example, Definition) to those already included in the extended version of RST (Set-Member, Abstraction-Instance, Whole-Part, Process-Step, Object-Attribute, Generalisation-Specific). In contrast to expanding Elaboration, the CDT Treebanks (Buch-Kromann, Hardt, & Korzen, 2011; Korzen & Müller, 2011) propose Elaboration as a subtype of Conjunction, hereby merging the two most frequent relations. Elaboration is rarely signalled by discourse cues, and may both elaborate on the content of a whole clause or sentence and on a single noun phrase. In example 86) from the Danish part of the corpus, we can observe how an Elaboration satellite is

textualised as an independent sentence, elaborating on the noun *beskyttelsesniveau* (level of protection).

86) [Det vil medføre, at lande med et højt beskyttelsesniveau kan risikere at blive oversvømmet med reklamer og markedsføring fra firmaer og fra lande med et lavere beskyttelsesniveau.] [Beskyttelsesniveauet i de nordiske lande er generelt højere end i andre lande,] [især når det gælder markedsføring over for børn.] <ep-99-05-06.txt:42>

[This will mean that countries with a high level of protection may risk being overrun with advertisements and marketing from firms and from countries which have a lower level of protection.] [The level of protection in the Nordic countries is generally higher than in other countries,] [especially as far as marketing in relation to children is concerned.]

Most of the instances of Elaboration in English and Italian are found in restrictive relative clauses and the various types of non-finite modifiers. As with the Circumstance relations, these highly syntactically dependent constructions underline the satellite status of Elaboration EDUs, cf. 87), where the relative clauses elaborate on the framework, and 88), where the modifier *denominato* (known) adds further information on the name of the committee.

87) [Council Regulation No 3760/92 established the framework] [upon which the CFP rests.] <ep-02-05-29.txt:37>

88) ... [la prima riguarda l'istituzione del comitato unico,] [**denominato** Comitato per la sicurezza marittima;] ... <ep-01-02-12.txt:64>

*... [the first concerns the establishment of the single committee,] [**known** as the Committee on Safe Seas;]...*

In Chapter 3, I discussed the criteria for EDU segmentation and argued that restrictive relative clauses and participial modifiers should be included as individual EDUs, because they are capable of expressing rhetorical relations other than Elaboration. This is still true, but this segmentation convention also results in the inclusion of many restrictive relative clauses and modifiers expressing Elaboration. I do not see this as an optimal solution for all studies of discourse structure, but in this thesis where one of the aims was to investigate linguistic structures, e.g. the textualisation of EDUs and rhetorical relations across different languages, it was considered necessary to adopt this convention.

5.3.3 Non-Volitional Cause

Instances of Non-Volitional Cause and its counterpart Non-Volitional Result are typically cases of deductive reasoning; the only difference between the two consists in nuclearity. In Non-Volitional Cause, the writer provides the cause of a situation that does not have a chosen outcome and in which the agent (if any) did not intentionally motivate the action caused by the situation in the nucleus. This is the case in the Italian example in 89), where the writer is prepared to postpone the report to another date because the agenda is already too full. In examples like this, we can see Non-Volitional Cause used as a politeness strategy (P. Brown & Levinson, 1987) so that the reader, *in casu* the President, does not lose face. Note also the polite conditional mood of the main verb in the nucleus (*sarei* translated with *I would be*) (see Maiden & Robustelli, 2007, p. 336), the polite address to the president using the formal third person singular (*lei*), and, last but not least, the embedded conditional adverbial clause (*se lei è d'accordo/if you agree*).

89) [Tuttavia, signor Presidente, [poiché mi rendo conto che l'ordine del giorno è estremamente denso,] sarei anche disposto, [se lei è d'accordo,] a chiedere uno slittamento della relazione a una prossima tornata] ... <ep-97-11-17.txt:24>
[However, Mr President, [as I realise that the agenda is extremely full,] I would be prepared, [if you agree,] to ask for the report to be deferred to a future sitting] ...

In this way, the cause of the situation in the nucleus becomes external to the writer; a situation that he would have avoided if he could. In the corpus, Non-Volitional Cause is both placed before and after the nucleus in subordinate finite adverbial clauses, but they can also be unmarked in coordinated clauses. Here, the cause is usually placed before the nucleus (which, thus, contains the consequence). In example 90), we can see that the first coordinated EDU is rhetorically subordinated and expresses the cause of the second EDU. Note also that, in cases like this, the causal relation is not marked in the satellite but in the nucleus (*altså/so* signals a result), which in the co-text of the given text is more central to the writer that the cause in the first EDU.

90) [Men pointen er jo, at en national beskatning muligvis vil være i strid med EU-retten,] [og problemet er **altså** ikke løst.] <ep-03-06-02.txt:67>
*[The point is, of course, that national taxation would possibly be in conflict with EU law,] [and **so** the problem has not been solved.]*

Non-Volitional Cause can also be textualised in non-finite constructions in examples such as 91) and 92), where the semantics of the verbs *given* and *caused* signal the causal relationship between the two EDUs.

91) [Mr President, [**given** the history of the United Kingdom in relation to the BSE crisis,] some delegates may consider that the UK is disqualified from giving an objective view.] <ep-99-07-22.txt:25>

92) [My constituents in London also suffer the problems of sleep disturbance] [**caused** by early morning flights at Heathrow Airport] ... <ep-97-06-09.txt:99>

5.3.4 Purpose

Contrary to Cause and Result, Purpose is not divided into volitional and non-volitional cases; instead it has been defined to include both instances (Mann & Thompson, 1987, p. 57). The relation is mostly intrasentential, triggered by discourse cues such as *to, so that* and *in order to*, with its Danish and Italian counterparts. A majority of the Danish EDU realisations with non-finite verb forms express Purpose relations, although the relation in the following example is expressed in a subordinate finite adverbial clause.

93) [Ordføreren lægger også stor vægt på, at der skal gives støtte til de unge landmænd,] [**så** de kan overtage produktionen.] <ep-98-06-16.txt:34>

*[The rapporteur also strongly emphasises the need to support young farmers] [**so they can take over production**.]*

The difference in terms of textualisation of the Purpose EDUs is related to the choice of discourse cue: if the cue employed is *to* or *in order to*, the Purpose EDU will be textualised as an infinitival clause. If the discourse cue *so (that)* is used, the Purpose EDU must be a subordinate finite adverbial clause. Purpose can also be found in relative clauses; in example 94), the relative clause (or, in fact, an 'echo apposition', see Korzen (2007)) actually signals its rhetorical relation to the nucleus using the word *purpose*.

94) [It is quite clear that what you are doing here is militarizing the EU,] [turning it into a military alliance,] [the main **purpose** of which is to protect its selfish interests] [and get easy access to global resources.] **<ep-98-05-13.txt:201>**

Purpose can also be found between intersententially related EDUs. In example 95), the second of the two clusters of EDUs, that is, the two text spans, expresses the Purpose of streamlining

the two Commission proposals mentioned in the first text span. Again, we can observe how the Purpose relation is explicitly signalled by the word *scopo (*purpose*)*. Note that the first text span also contains an intrasentential Purpose textualised in the infinitival clause *per introdurvi detto comitato* (in order to incorporate the committee).

95) [Signor Presidente, signora Commissario, onorevoli colleghi, [come ha ricordato il relatore,] con il documento di oggi si considerano due proposte della Commissione esecutiva:] [la prima riguarda l'istituzione del comitato unico,] [denominato Comitato per la sicurezza marittima;] [la seconda riguarda una direttiva] [che modifica le direttive esistenti] [per introdurvi detto comitato.] [Lo **scopo** di tutto ciò [- è stato già ricordato dai colleghi -] è razionalizzare l'opera di aggiornamento della legislazione comunitaria in vigore alle disposizioni e alle convenzioni internazionali sulla sicurezza marittima e la protezione dell'ambiente marino, nonché al progresso tecnico.] <ep-01-02-12.txt:64>

*[Mr President, Commissioner, ladies and gentlemen, [as the rapporteur has said,] this document examines two Commission proposals:] [the first concerns the establishment of the single committee,] [known as the Committee on Safe Seas,] [and the second concerns a directive] [which amends the existing directives] [in order to incorporate the committee.] [The **purpose** of this [- as Members have already pointed out -] is to streamline the operation of bringing the Community legislation in force in line with the international provisions and conventions on maritime safety and the protection of the marine environment, and with technical progress.]*

5.3.5 Volitional Cause

Volitional Cause is defined as an action or situation which may cause the agent of the situation in the nucleus to perform that action. As such, Volitional Cause typically contains an agent which controls an action that yields the nuclear situation; in the following example, the agent is the writer's political group that has voted against a report, because they are concerned about a number of issues.

96) [We did, however, vote against the report in committee] [**because** of a number of outstanding concerns,] [which we will try to address tomorrow in a series of amendments and requests for split votes.] <ep-02-07-01.txt:103>

Volitional Cause is used to express the writer's reason for acting or arguing the way she does. In the following two Danish examples in 97) and 98), we can see how different syntactic constructions highlight the salience of the causal relation between the satellite and the nucleus.

In 97), the causal satellite is signalled by the subordinator *for* (because), with the discourse particle *jo* (translated with *as you know*) supporting the truthfulness of the satellite content. Thus, in this example, we can see how the rhetorical relationship between two EDUs may display traces of both semantic and pragmatic sources of coherence.

97) … [det kan under ingen omstændigheder accepteres, at grænseværdierne ikke overholdes,] [**for** det er **jo** dokumenteret, at der er tale om fare for kvinder og børn,] [hvis der indtages for meget bly.] <ep-96-12-11.txt:268>

*… [under no circumstances can it be accepted that the limit values are not respected,] [**because**, **as you know**, it has been documented that there is a risk of harm to women and children] [if too much lead is absorbed.]*

In 98), the causal satellite relates two independent sentences. In the second sentence, the EDU *Det har vi gjort, fordi* (We have done so because) functions as a discourse cue expressing the causal relationship to the previous sentence and has therefore not been segmented as two separate EDUs.

98) [Som medlemmer af den socialdemokratiske gruppe har vi stemt sammen med vores partifæller for og imod en række ændringsforslag.] **[Det har vi gjort, fordi vi gerne vil præge udviklingen i Europa i en socialdemokratisk retning.]** <ep-99-11-18.txt:143>

[As members of the group of socialists, we have voted, together with our fellow party members, for and against a number of amendments.] **[We have done so because we want the development in Europe to move in a social democratic direction.]**

Both cases show that Volitional Cause to a larger extent than Non-Volitional Cause serves to articulate the writer's reasoning without taking the reader's face into consideration. Thus, when using Volitional Cause, the writer defends her arguments knowing that the reader might not agree with her.

As with Non-Volitional Cause and Non-Volitional Result, Volitional Cause is used more frequently than Volitional Result, because the EDU expressing the result often contains the most important information. This aspect of rhetorical hierarchy will be discussed in Section 5.5.2. The section above has presented the main findings, with examples, for the most frequent semantic rhetorical relations for all three languages. Now, we will turn to the findings on the pragmatic relations, and how the two types of sources of coherence compare.

5.4 The five most frequent pragmatic relations

The pragmatic relations in RST were originally called presentational relations, because they were claimed to 'facilitate the presentation process itself' (Mann & Thompson, 1987, p. 17). In this thesis, I use the term pragmatic instead of presentational, following van Dijk (1979, p. 449): 'pragmatic connectives express relations between speech acts, whereas semantic connectives express relations between denoted facts'. The original pool of pragmatic relations contained only seven relations, but it was later expanded to include ten relations whose intended effect is to increase some inclination in the reader. The relations and their intended effects are listed in Table 5.3, partially taken from Stede (2008b, p. 325), who, however, does not include Preparation, Restatement, Summary and Justify in his list, because they are not pragmatic 'on a par' with the other pragmatic relations.

Pragmatic rhetorical relations	Intended effect on reader
Antithesis	Encourage appreciating
Concession	Encourage appreciating
Evidence	Encourage believing
Justify	Encourage acceptance
Motivation	Encourage acting
Preparation	Encourage continued reading
Enablement	Ease acting
Restatement	Ease understanding
Summary	Ease understanding
Background	Ease understanding

Table 5.3: Pragmatic relations and their intended effects

The literature not only discusses the actual nature of pragmatic relations and the best-suited term(s), but also which relations to include as pragmatic. I shall not go into a discussion of whether Preparation, Restatement and Summary are more pragmatic than semantic, both because it is outside the scope of this thesis and because the frequency of these three relations in the corpus employed is relatively low (<2 %).

5.4.1 Background

The first pragmatic relation, Background, is mostly found at the beginning of the texts. The purpose of placing the background information at the very beginning, from the writer's

viewpoint, is to explain to the reader why the issue is important or to provide some information necessary for the reader to understand the situation presented in the nucleus. The relation is used similarly in all three languages. In example 99), the writer provides some background (*this debate was opened by Mr Da Costa*) so that the reader understands the main nucleus of the text (*we condemn that hypocrisy and the double standards*). In this case, there are several EDUs between the Background and the nucleus spans, indicated by means of blank square brackets [...]. As in many cases of Background, the relation is used to connect text spans of several EDUs across various sentences and not signalled overtly by any discourse cues.

99) [Madam President, this debate was opened by Mr Da Costa,] [who merely confirms that the EU presidency is trying to influence the composition of a government of one of the Member States.] [This debate, and the background to it, reinforced the point that the Council, the Commission and this Parliament itself are dominated by Socialists.]
[...]
[So we condemn that hypocrisy and the double standards, especially of the British government and the other largely socialist heads of EU governments] [and include that of the United States.] <ep-00-02-02.txt:32>

In many of the texts of the corpus, the first or first few sentences are an acknowledgment or expression of gratitude towards the president of the sitting or another MEP. This kind of rhetorical move (see Biber et al., 2007) was also analysed as Background, as the move is employed to increase the reader's ability to comprehend to whom the writer actually addresses her speech. Example 100) shows an instance of this use of Background, where the writer addresses the president of the sitting.

100)   [Signor Presidente, la ringrazio per avermi dato la parola su questo punto.] [Vorrei sottolineare che più volte, anche per lettera, come presidente della CERT, ho dovuto lamentare che le relazioni della commissione per la ricerca e l'energia finiscono per essere discusse in orari quantomeno particolari: o tardi nella serata o all'estremo limite della tornata.] <ep-97-11-17.txt:24>

[Mr President, thank you for allowing me to speak on this point.] [I would like to point out that I have had to complain several times, including in writing, as chairman of the Committee on Research, Technological Development and Energy, that its reports end up being discussed at very odd times: either late in the evening or at the very end of the sitting.]

5.4.2 Concession

Besides being a pragmatic relation, Concession, and its closely related companion Antithesis, is also one of the classic rhetorical persuaders (Azar, 1999, p. 106). The persuasive effect of Concession lies in the writer's presentation of a counter argument. The writer states in advance what may be seen as an unfavourable argument for her belief and thus eliminates a possibly unfavourable intervention from the reader. Furthermore, Concession also reinforces the credibility of the nucleus, as the reader is led to understand that the writer has already considered other arguments as possible valid counter arguments. In this way, Concession is a persuasive technique that could be described as psychologically manipulative. In the following example from the Danish part of the corpus, this effect is clearly shown, as the writer, in the first sentence, accepts the fact that people are tired of all the fraud in the EU and therefore demand a common public prosecutor's office. In the next sentence, the writer refutes the people's argument by stating that the actual problems are to be found elsewhere. Note also the appealing rhetorical manoeuvre *Jeg kan godt forstå* (I can well understand), which almost explicates the concessive nature of the EDU.

101)   [Jeg kan godt forstå, at folk er trætte af al den svindel] [og derfor kræver, at vi får en fælles anklagemyndighed.] [Det er imidlertid en stor mastodont at bygge op,] [når problemerne i virkeligheden ligger et andet sted.] <ep-00-05-16.txt:52>

[I do understand that people are tired of all the fraud] [and therefore demand that we have a common public prosecutor's office.] [It would, however, be a colossal enterprise to set up] [when the problems, in reality, are elsewhere.]

In this example, we observe again how the discourse cue (*imidlertid/however*) signalling Concession is located in the nucleus and not in the satellite, which usually is considered the typically location of discourse cues. According to Mann & Thompson (1987, p. 17), one of the reasons for this placement is that some relations have canonical orders of spans, that is, the satellite is placed before the nucleus or vice versa. Concession belongs to the former group, see example 102), where the Concession is expressed between two coordinated EDUs, the first being the satellite.

102)   [You may be curious] [but you are not going to work on that subject.] <ep-97-03-11.txt:36>

The only cross-linguistic difference in the use of Concession is a higher tendency in English to use the relation intersententially. The prototypical textualisation of Concession is a subordinate finite adverbial clause with the discourse cue *although*, which can be placed both before and after the nucleus. In example 103), the concessive satellite is placed after the nucleus. In this case, the persuasive effect is downgraded, putting more focus on the apparent incomparability between the nucleus and the satellite.

103)   [Hr. formand, mit udgangspunkt er Equal-betænkningen og ligestilling mellem mænd og kvinder,] [**selvom** jeg må indrømme, at det indimellem er træls at tale om dette spørgsmål,] ... <ep-00-02-14.txt:70>

*[Mr President, my starting point is the Equal report and equality between men and women,] [**although** I must admit that it is sometimes hard to talk about this issue,] ...*

5.4.3   Evidence

Evidence is also a classical rhetorical argument relation, as its intended effect is to increase the belief in what is said in the nucleus. This is done by putting forward a situation in the satellite that is acceptable in terms of content to the reader. In example 104), the nucleus (*Chainsaws are vital for managing woodlands*) functions as the conclusion, whereas the satellite (*Manufacturers say that...*) functions as the argument supporting the claim in the conclusion.

104)   [Chainsaws are vital for managing woodlands.] [Manufacturers say that [if this amendment goes through,] they will no longer be able to manufacture these goods.] <ep-01-10-01.txt:83>

In another example of the Evidence relation, it can again be seen how the relation essentially encompasses the very nature of argumentation: in 105), the satellite is presented in order to evidence the truthfulness of the nucleus so that the reader's propensity to believe the statement increases. Here, the writer uses another country (or union) to support the conclusion in the nucleus.

105)   [Set aside-ordningen er det bedste værktøj til at regulere kornproduktionen.] [Det har USA vidst i mange år.] <ep-99-01-13.txt:44>

[The set-aside scheme is the best tool for regulating the production of cereals.] [The USA has known this for many years.]

Finally, Evidence can also be used intrasententially, as in the Italian example 106). Cross-linguistically, the Italian texts used Evidence more often than the Danish and English texts to relate two or more EDUs within the same sentence, typically by referring to another person's statements, facts or figures.

106)   [Questa è una necessità obiettiva,] [se si continua a credere alla primazia del politico su un certo tipo di economia lobbista] [che certamente finora non ha dato, [come dimostrano i dati del rallentamento economico,] il risultato sperato.] <ep-02-05-14.txt:71>

[This is a real necessity] [if we still believe in the primacy of politics over a certain type of lobbying economy,] [which has certainly not yielded, [as the figures on the economic slowdown show,] the desired result thus far.]

5.4.4   Justify

Justify is used when the writer wants to increase the reader's readiness to accept the right to present the speech act of the nucleus. In example 107), the purpose of the nucleus is to question the attitude of France and Spain towards the Davos package. The satellite is interpreted as Justify because the writer uses her experience (*as a veteran observer of these four-year long negotiations*) for justifying the question presented in the nucleus.

107)   [Madam President, as a veteran observer of these four-year long negotiations I suppose I should not have been surprised by the attitude, particularly of France and Spain, at Monday's meeting of the General Affairs Council.] [How can they argue that the Davos package is too generous to President Mandela's South Africa?] <ep-99-02-24.txt:58>

The same argumentative technique is found in the Italian example 108), where the writer again uses his professional background to increase the reader's acceptance of the relatively complicated claim presented in the nucleus about animal experimentation. In fact, the writer repeats his scientific background three times (*come uomo di scienza/as a scientist; come scienziato/as a scientist; come uomo che ha lavorato 35 anni in laboratorio/as a man who has worked in laboratories for 35 years)*, leaving no doubt in the reader's mind that he is the most competent person to present this claim.

108) [Voglio fare, come uomo di scienza, anche un'altra considerazione, caro Commissario,] [e cioè che i metodi alternativi, [nonostante sia importante svilupparli, potenziarli, finanziarli,] non saranno mai, mai in grado [- e lo dico come scienziato, come uomo] [che ha lavorato 35 anni in laboratorio -] di sostituire la complessità di un organismo animale e di un organismo umano.] <ep-01-04-02.txt:42>

[As a scientist, I would like to make another point, dear Commissioner,] [and that is that, [although it is important to develop, improve and finance alternative methods,] they will never, never [- and I am saying this as a scientist, as a man] [who has worked in laboratories for 35 years -] they will never be able to reproduce the complexity of an animal or human organism.]

Finally, Justify is used in this study to account for all the instances where writers refer to colleagues who share the same beliefs. In example 109), the relative clause expresses the justification of the nucleus.

109) [I will tell my colleague, Commissioner Verheugen, of the views] [that have just been expressed,] [which I know are held by other Members of Parliament.] <ep-00-09-04.txt:46>

5.4.5 Summary

Summary differs from Restatement with respect to size only: Summary satellites are shorter in bulk than their respective nuclei; Restatement satellites are equal in size with their nuclei. Summary is typically found at the very end of the text and functions as a recapitulation of the main viewpoints expressed by the writer. In most texts containing Summary, the satellite recaps the main nucleus, or as seen in 110), the main nuclei of the text. In this way, Summary can be used to check one's analysis of nuclearity in the text (although this naturally has to be done with caution). In example 110), the writer presents three statements on the position of his political group in the main body of the text and concludes the text by summarising that these were the positions, in a text span consisting of five EDUs across two sentences.

110) (Three statements on the position of the writer's group.)
[That is the position of our group,] [subject to the arguments] [we are listening to now.] [However, [whatever happens to those particular amendments,] I am quite confident that there will be a successful outcome in the vote tomorrow.] <ep-01-11-12.txt:56>

In another example 111), from the Danish part of the corpus, Summary is used in the same way as above. Here, the main nucleus is presented in the first sentence of the text, and the summary in the last sentence. In between, we find 19 EDUs elaborating on why the writer is disappointed with the colleague's report.

111)   [Hr. formand, jeg må sige, at jeg er skuffet over resultatet af Rosado Fernandes' arbejde med denne betænkning.]

[19 EDUs elaborating on this]

[Jeg må sige, hr. formand, jeg er meget skuffet over denne betænkning.] <ep-98-06-16.txt:34>

[Mr President, I must say I find the results of Mr Rosado Fernandes' work on this report rather disappointing.]
[19 EDUs elaborating on this]
[Mr President, I must say I am most disappointed with this report.]

This and the previous section have presented the ten most frequent mononuclear relations, with examples from all three languages. Some examples contained coordinated EDUs, that is, syntactic coordination, even though the EDUs in question exhibited a subordinate relationship rhetorically speaking. The next section focuses on the multinuclear relation of Conjunction and on the difference between syntactic dependency and rhetorical hierarchy.

5.5   Rhetorical hierarchy and syntactic dependency

As mentioned in the Chapter 4 on text structure, this thesis also sets out to examine whether syntactic coordination and subordination are identical with rhetorical coordination and subordination. Matthiessen & Thompson (1988) argue that the notion of syntactic subordination very well could have arisen out of the textualisation of rhetorical relations. The same viewpoint, put somewhat more strongly, is found in the previously cited SDRT (Asher & Vieu, 2005), which assumes a one-to-one relationship between syntactic dependency and rhetorical hierarchy. In the previous sections, I have already put forward various cases where two syntactically coordinated EDUs were not rhetorically coordinated: the examples in Non-Volitional Cause, Evidence, Justify and Concession, to name a few. This section presents an investigation of rhetorical hierarchy and syntactic dependency, starting with a description of the multinuclear Conjunction relation as used in the corpus.

5.5.1 Conjunction

Conjunction is the second most frequent relation found in the corpus for all three languages, with slightly more than 10 % of the total number of relations. Conjunction is not found in the original RST inventory, but was embedded in the List relation. With the extended RST relation set (Mann, 2005), Conjunction was added to account for the many coordinated clauses that could not be interpreted as EDUs under the List relation but as juxtapositions. Ramm & Fabricius-Hansen (2005, n. 9) report on the basis of personal communication with the developer of the RSTTool (O'Donnell, 2004) that Conjunction 'is meant to cover constructions with "and" connectives'. This may seem like a very vague definition which essentially could contradict many of the analyses from the previous sections, but Conjunction is, as argued, a very empty relation in terms of semantics. In cases where the two conjoined EDUs are of equal rhetorical status, e.g. the Danish example in 112), Conjunction merely serves as an additive relation rather than as a real semantically enriched rhetorical situation, here signalled jointly by *og ... i øvrigt også* (and ... also).

112)   [Jeg synes, det er et direktiv, der går i den rigtige retning,] [**og** det går **i øvrigt også** på det overordnede plan i den rigtige retning.] <ep-98-09-15.txt:97>

*[I think it is a directive which moves in the right direction,] [**and** at a more general level it **also** moves in the right direction.]*

In many instances, two or more coordinated EDUs share some syntactic property. This can be the subject but also other grammatical elements, as shown in the English example in 113), where the object of the prepositional phrase (*a major impact*) is elided. The discourse cue in the example (*not only ... but also*) is very common in the texts of the corpus, which is probably owing to their argumentative nature and the cue's emphasising effect on the second EDU.

113)   [Mr President, the introduction of the euro will have a major impact **not only** for EU citizens] [**but also** on the international monetary and financial system.] <ep-98-12-02.txt:150>

In other instances, Conjunction is used to relate two EDUs or text spans across the sentence boundary. For example, this is the case in 114), which resembles the English example in 113), through the discourse cue *ikke blot ... også* (not only ... also).

114)   [Det er jo **ikke blot** grænsen mellem Danmark og Tyskland.] [<u>Det er **også** grænsen mellem de øvrige nordiske lande og Tyskland.</u>] <ep-96-06-19.txt:226>

*[This is **not only** the border between Denmark and Germany.] [<u>It is **also** the border between the other Nordic countries and Germany.</u>]*

It has been argued that rhetorically coordinated EDUs that are not in the same sentence differ in some respects from syntactically coordinated EDUs that belong to the same sentence. Blakemore (Blakemore & Carston, 2005; Blakemore, 1987) suggests that with two syntactically coordinated EDUs the writer signals to the reader that the two EDUs (or text spans, when several EDUs are included) should be processed as one unit, both units functioning as premises in the derivation of a joint cognitive effect. Conversely, two coordinated EDUs in two distinct sentences should be processed as two units. Blakemore (Blakemore & Carston, 2005) also states that another difference between intrasententially and intersententially coordinated text spans is that certain inferences are permitted regarding the semantic-pragmatic relations holding between them, the first unit always functioning as background to the processing of the second (*pace* Ramm & Fabricius-Hansen, 2005, p. 5). The section below discusses the idea of using syntactically 'symmetric' (multinuclear) constructions with 'asymmetric' (mononuclear) degrees of importance by looking at some of the causal relations in the corpus.

5.5.2   Cause or Result?

When a causal relationship is expressed between two EDUs, usually both a cause and a result are present. In RST, the task is to determine which of the EDUs is the most important in the given context. In the corpus studied here, we saw in the previous section that the cause in both its variants (Non-Volitional and Volitional) most frequently was annotated as the satellite. We also saw that typical causal constructions were subordinate finite adverbial clauses containing discourse cues such as *because, since* and *as,* together with other subordinate constructions. But causal relations are also textualised in coordinated constructions with the discourse cue *and*, see the English example in 115). Here, the first two underlined EDUs establish the causal relationship (Non-Volitional Cause) to the nucleus span consisting of the rest of the sentence, but they are syntactically coordinated with the other matrix clause in the sentence *(and I hope...)* through the semantically empty coordinator *and*.

115)   [<u>My constituents in London also suffer the problems of sleep disturbance</u>] [<u>caused by early morning flights at Heathrow Airport</u>] [**and** I hope Parliament's call for a ban on

night flights will be acted upon] [before my citizens' lives are made even more miserable] [with the increase to air traffic] [which is likely to result from Terminal 5.] <ep-97-06-09.txt:99>

This connection between rhetorical subordination and syntactic coordination in causal relationships between EDUs becomes much clearer in example 116), where the discourse cue *derfor* (therefore) has been added to the coordinator *og* (and), indicating that the first EDU provides the situation that has caused the situation in the second EDU (the nucleus).

116)　　[Forslaget har længe gået på omgang mellem institutionerne,] [**og** det er **derfor** vigtigt, at Rådet ikke lægger hindringer i vejen på dette vigtige område.] <ep-03-02-11.txt:38>

*[The proposal has been passed around between the institutions for a long time,] [**and** it is **therefore** important that the Council does not create obstacles in the way of this important area.]*

The extra discourse cue *derfor* (therefore) could have been omitted, but because the writer wanted to highlight the salience of rhetorical relations, the semantic relationship between the two EDUs has been overtly signalled. A similar way of signalling rhetorical subordination between two syntactically coordinated EDUs is found in 117). Here, the coordinator *og* (and) has been added to a clause which takes a sentential complement *dette medfører, at...* (this means that…), which, again, indicates that a causal relationship should be inferred between the two spans.

117)　　[Kun få fælles regler og procedurer er gældende i alle medlemsstaterne,] [**og dette medfører, at** der ikke eksisterer et legalt og administrativt værktøj,] [der kan bidrage til en bedre styring af migrationsstrømmene.] <ep-03-02-11.txt:38>

*[Only a few common rules and procedures apply in all the Member States,] [**and this means that** there is no legal and administrative tool] [that can contribute to a better management of migration flows.]*

Common to both examples in 116) and 117) is that the causal satellites could be rewritten into subordinate finite adverbial clauses; in example 117), as *[Because only a few common rules and procedures apply in all the Member States], [there is no legal and administrative tool that can contribute to better management of migration flows]*. Syntactically coordinated constructions with causal related EDUs are almost exclusively found in the Danish texts, although the

constructions are perfectly valid in English and Italian, too, both grammatically and semantically. The reason for this is not only the higher number of intrasententially coordinated EDUs in Danish, but also the tendency for Danish to be less constrained with respect to which rhetorical relations can be expressed in syntactically coordinated constructions, see Table 5.4. We shall return to this issue in Chapter 6 on information structure.

5.6   Cross-linguistic differences in discourse structure

The overall distribution of the rhetorical relations across Danish, English and Italian showed no major differences in terms of frequencies and usage, see Table 5.1. However, the last few sections have shown that the three languages display some dissimilarities in the way they textualise the relations intrasententially versus intersententially. In Table 5.4, a more detailed distribution of the various relations is presented, showing how often they connect EDUs inside the same sentence or across sentence boundaries. The two columns under each language contain the distribution of each relation in terms of percentage, divided into intrasentential linkage and intersentential linkage. In the first row, for example, we can see that Circumstance is used in Danish to link EDUs inside the same sentence in 98.6 % of the cases found. The percentages are based on the numbers from Table 5.1. In contrast to Table 5.1, where the overall distribution of the various rhetorical relations was seen to be relatively similar, Table 5.4 displays interesting cross-linguistic differences concerning the distribution of relations and their linkage.

First, we can see that some relations have a tendency to link EDUs inside the same sentence. This holds particularly for Circumstance, Condition, Means, Otherwise, Purpose, Unconditional, Unless and Disjunction, and is true of all three languages. The most frequent relation, Elaboration, in most cases, operates like these relations, especially in Italian, with 92.0 % of the instances relating EDUs inside the same sentence. However, the fact that elaborating satellites are textualised as independent sentences more frequently in Danish and English could be because the writer feels the need to split up the nucleus and the Elaboration satellite into two sentences to avoid too long sentences which could jeopardise the reader's comprehension.

On the other hand, relations such as Solutionhood, Background, Enablement, Preparation, Summary, Joint and Multinuclear Restatement most frequently link EDUs that are not located in the same sentence. Note that only one of these is semantic (Solutionhood), which is used for question-answer-like sequences. Four come from pragmatic sources of coherence, but apart from Background, which was discussed above, the occurrences of the other relations are so low that it is hard to conclude whether they are used intersententially because they typically link text

spans including several EDUs or because they do not have any explicit discourse cues to signal them. Disjunction is usually textualised in coordinated clauses through the coordinator *or*, and the last two multinuclear relations, Joint and Multinuclear Restatement, also have low occurrences, but I find it quite reasonable that these are usually textualised as independent sentences.

Another interesting observation is that the semantic relations exhibit a general tendency across the three languages to textualise semantic relations intrasententially more frequently than pragmatic relations. Pragmatic relations are typically textualised between sentences, with a few exceptions in the Italian texts (Concession and Justify). This could indicate a more universal tendency across Danish, English and Italian to prefer sentence shifts when relating speech acts. Table 5.5 shows the distribution of sources of coherence and the two types of linkage. From a cross-linguistic perspective, Danish and English tend to distribute the sources of coherence almost equally, whereas Italian in all cases displays a higher tendency to prefer intrasentential linkage of EDUs. For instance, the Italian pragmatic relations are more frequently textualised between EDUs inside the same sentence (44.9 %) than Danish (25.2 %) and English (20.1 %) pragmatic relations.

		Danish		English		Italian	
		Intra-	**Inter-**	**Intra-**	**Inter-**	**Intra-**	**Inter-**
SEMANTIC RELATIONS	Circumstance	98.6 %	1.4 %	97.6 %	2.4 %	100.0 %	0.0 %
	Condition	100.0 %	0.0 %	100.0 %	0.0 %	100.0 %	0.0 %
	Elaboration	70.0 %	30.0 %	77.5 %	22.5 %	92.0 %	8.0 %
	Evaluation	26.1 %	73.9 %	31.8 %	68.2 %	83.9 %	16.1 %
	Interpretation	21.7 %	78.3 %	17.9 %	82.1 %	66.7 %	33.3 %
	Means	95.2 %	4.8 %	88.5 %	11.5 %	96.4 %	3.6 %
	Non-Vol. Cause	32.3 %	67.7 %	34.0 %	66.0 %	28.6 %	71.4 %
	Non-Vol. Result	17.4 %	82.6 %	44.0 %	56.0 %	55.6 %	44.4 %
	Otherwise	-	-	-	-	100.0 %	0.0 %
	Purpose	93.2 %	6.8 %	100.0 %	0.0 %	98.3 %	1.7 %
	Solutionhood	2.3 %	97.7 %	7.7 %	92.3 %	3.8 %	96.2 %
	Unconditional	100.0 %	0.0 %	100.0 %	0.0 %	100.0 %	0.0 %
	Unless	100.0 %	0.0 %	100.0 %	0.0 %	100.0 %	0.0 %
	Volitional Cause	85.2 %	14.8 %	85.7 %	14.3 %	93.3 %	6.7 %
	Volitional Result	74.1 %	25.9 %	75.0 %	25.0 %	96.0 %	4.0 %
PRAGMATIC RELATIONS	Antithesis	35.3 %	64.7 %	57.1 %	42.9 %	80.0 %	20.0 %
	Background	5.7 %	94.3 %	0.0 %	100.0 %	9.9 %	90.1 %
	Concession	52.1 %	47.9 %	35.7 %	64.3 %	68.3 %	31.7 %
	Enablement	0.0 %	100.0 %	0.0 %	100.0 %	0.0 %	100.0 %
	Evidence	13.3 %	86.7 %	9.3 %	90.7 %	40.0 %	60.0 %
	Justify	41.7 %	58.3 %	44.4 %	55.6 %	68.6 %	31.4 %
	Motivation	18.8 %	81.3 %	15.0 %	85.0 %	29.2 %	70.8 %
	Preparation	0.0 %	100.0 %	0.0 %	100.0 %	33.3 %	66.7 %
	Restatement	45.5 %	54.5 %	50.0 %	50.0 %	20.0 %	80.0 %
	Summary	0.0 %	100.0 %	0.0 %	100.0 %	5.3 %	94.7 %
MULTINUCLEAR RELATIONS	Disjunction	100.0 %	0.0 %	100.0 %	0.00 %	100.0 %	0.00 %
	Conjunction	54.2 %	45.8 %	51.5 %	48.5 %	74.6 %	25.4 %
	Contrast	77.8%	22.2%	85.7 %	14.3 %	83.9 %	16.1 %
	Joint	0.0%	100.0%	0.0%	100.0 %	0.0 %	100.0 %
	List	35.1%	64.9%	61.5 %	38.5 %	86.4 %	13.6 %
	Mc. Restatement	-	-	0.0%	100.0 %	0.0 %	100.0 %
	Sequence	41.9%	58.1%	28.0 %	72.0 %	26.7 %	73.3 %

Table 5.4: Distribution of rhetorical relations and linkage

	Danish		English		Italian	
	intra-	inter-	intra-	inter-	intra-	inter-
Semantic relations	66.4 %	33.6 %	72.6 %	27.4 %	82.6 %	17.4 %
Pragmatic relations	25.2 %	74.8 %	20.1 %	79.9 %	44.9 %	55.1 %
Multinuclear relations	51.4 %	48.6 %	48.0 %	52.0 %	70.4 %	29.6 %

Table 5.5: Sources of coherence and linkage

Moreover, we can observe that four relations in particular tend to be textualised differently in Danish and English than in Italian: Evaluation, Interpretation, Antithesis and Evidence. These cross-linguistic differences are interesting because they reveal the interstructural connection between text structure, discourse structure and information structure. We have seen that rhetorical relations appear to be used in the same way in Danish, English and Italian, but that there are cross-linguistic differences in the way in which the relations are used syntactically to link EDUs. In the cases of the four relations mentioned above, we can again observe that Danish and English prefer to textualise satellites in independent, or new, sentences. See the examples below of how Evaluation is textualised differently across the three languages: in Danish 118) and English 119) intersententially, and in Italian intrasententially as embedded clauses – a coordinate clause insertion in 120) and an appositive relative clause in 121).

118)   [Fru formand, på vegne af den liberale gruppe vil jeg gerne hilse betænkningen af Terrón i Cusí velkommen.] [Det er en vigtig og nødvendig betænkning.] <ep-03-02-11.txt:38>

[Madam President, on behalf of the Group of the European Liberal, Democrat and Reform Party, I should like to welcome Mrs Terrón i Cusí's report.] [It is an important and necessary report.]

119)   [I therefore ask the Commissioner to take that on board.] [It is essential.] <ep-97-09-16.txt:102>

120)   ... [ricordo il fatto che nel testo si evochino [- e questo è positivo -] le lingue minoritarie regionali] [che hanno strumenti di protezione diversi.] <ep-01-01-15.txt:80>

... [I would point out [- and this is a positive fact -] that the text refers to regional minority languages,] [which are covered by a range of protective measures.]

121)   [L'esigenza, [che faccio mia,] di un'applicazione flessibile del patto di stabilità e di crescita e di un coordinamento delle politiche economiche, sociali e ambientali a partire dalla zona euro è stata però largamente disattesa in questi ultimi anni di recessione,] ... <ep-02-10-21.txt:49>

[The need, [which I fully endorse,] for flexible application of the Stability and Growth Pact and for coordination of economic, social and environmental policies, starting with the eurozone, has, however, been completely disregarded during the recent years of recession,] ...

Table 5.6, Table 5.7 and Table 5.8 summarise the findings of this chapter by pairing them with the different levels of the deverbalisation scale from Figure 4.1 in the previous chapter. In particular, the tables show in which types of textualisation the rhetorical relations are found, broken down according to each specific language: Table 5.6 contains all Danish relations, Table 5.7 all English and Table 5.8 all Italian. Above all the tables, the corresponding letter of the levels in the deverbalisation scale are found, specifying the syntactic construction employed. Table 5.6, Table 5.7 and Table 5.8 should be read as follows: if we look at the rhetorical relation Evaluation, we can see that the majority of textualisations occur at level (a) for Danish and English: in Danish, there are nine instances and in English, ten instances of the relation being textualised at level (a) as independent sentences. But if we look at Table 5.8 with the Italian distributions, we find only a single instance of Evaluation at level (a). Here, the majority of Evaluation textualisations are found at level (e), relative clauses, as exemplified above in 121).

Other cross-linguistic differences can be observed between the tendency in Danish to restrict the number of non-finite textualisations to express Elaboration, Means or Purpose and the tendencies in English and Italian to allow non-finite textualisations to express a wider range of rhetorical relations: Circumstance, Elaboration, Means, Purpose, Cause, Concession and Conjunction. In this way, we can see how the similarities in text structure patterns between English and Italian, in certain cases, can be refound in the discourse structure of the two languages.

What can also be deduced from the tables is that syntactic coordination does not always correspond to rhetorical coordination. By adding up the number of textualisations at level (c), coordinate main or matrix clauses for each language, we can see that, in Danish, 56.2 % of these express multinuclear relations, but that the other 43.8 % express mononuclear relations. The same patterns are found in English and Italian, with 65.5 % versus 34.5 % and 53.5% versus 46.5 % respectively, underlining the fact that syntactic coordination does not equal rhetorical

coordination. However, if we look at the number of syntactic subordinate textualisations (levels d-j) that express rhetorical coordination, the numbers clearly indicate a tendency towards a closer correspondence between syntactic and rhetorical subordination: in Danish, only 18.5 % of the multinuclear relations are textualised at level (d)-(j), in English 16.4 %, while in Italian we find 31.5 %. These percentages would have been even lower if I had not regarded coordinated subordinate constructions such as two coordinated relative clauses as two subordinate constructions but as one subordinate and one coordinate, see Section 4.4.2.

A – independent sentence E – relative clause I – nominalisation

B – main or matrix clause F – infinitival clause J – verbless const.

C – coordinate main or matrix clause G – gerund, pres. or past part.

D – subordinate finite adverbial clause H – pres., past or adj. modifier

		A	B	C	D	E	F	G	H	I	J
SEMANTIC RELATIONS	Circumstance	1	1		62	2	3			2	2
	Condition				29	3				3	1
	Elaboration	29	64	13	3	209	1		2	2	4
	Evaluation	9	7	2		5					
	Interpretation	14	5	1	1	2					
	Means	1	1	6	1		5			7	
	Non-Vol. Cause	10	35	7	8	2					
	Non-Vol. Result	7	10	4	2						
	Otherwise										
	Purpose	2		2	16	8	14	1		1	
	Solutionhood	18	24	2							
	Unconditional				3						
	Unless				1		1				
	Vol. Cause	3	15	2	30	9				1	1
	Vol. Result	2	5	12	4	1				2	1
PRAGMATIC RELATIONS	Antithesis	6	7	2		1					
	Background	26	36	8							
	Concession	15	35	11	10						
	Enablement	2									
	Evidence	19	25	13	3						
	Justify	14	13	7	9	4					1
	Motivation	5	5	5		1					
	Preparation	5	2	1	1						
	Restatement	3	2	3	1		1				1
	Summary	10	13	2							1
MULTINUCLEAR	Disjunction			5			4				
	Conjunction	27	49	91	6	4			1		
	Contrast	4	4	9	7		2			1	
	Joint	5	8	2							
	List	14	9	11	1	2					
	M. Restatement										
	Sequence	3	12	14		1	1				

Table 5.6: Textualisation of rhetorical relations in Danish

A – independent sentence   E – relative clause   I – nominalisation

B – main or matrix clause   F – infinitival clause   J – verbless const.

C – coordinate main or matrix clause   G – gerund, pres. or past part.

D – subordinate finite adverbial clause   H – pres., past or adj. modifier

		A	B	C	D	E	F	G	H	I	J
SEMANTIC RELATIONS	Circumstance	1	7	2	37	5		11	4	12	3
	Condition				24				1		
	Elaboration	25	49	11	4	171		5	71		
	Evaluation	10	3	4		5					
	Interpretation	10	13	2		3					
	Means	1	2		1			14	2	5	1
	Non-Vol. Cause	10	27	1	3	3	1	4			
	Non-Vol. Result	4	10	3	3	4			1		
	Otherwise										
	Purpose			1	2	4	25	1	2	5	
	Solutionhood	7	16	2	1						
	Unconditional				4			1			1
	Unless				2						1
	Vol. Cause	4	5	11		6		3	4		2
	Vol. Result	2	2			3	1	2			
PRAGMATIC RELATIONS	Antithesis	3	12	3			1	1		1	
	Background	29	47	1							
	Concession	16	23	5	7	1		1		1	2
	Enablement	2	2								
	Evidence	7	28	4	1	2			1		
	Justify	12	11	5	14	2			1		
	Motivation	5	13	1		1					
	Preparation	3	3								
	Restatement			1							
	Summary	13	8	2	1						
MULTINUCLEAR	Disjunction			1							
	Conjunction	25	49	82		4	3	2	2		1
	Contrast		4	10	4		3				
	Joint		16	4	1						
	List	2	3	7			1				
	M. Restatement		2								
	Sequence	3	13	8		1					

Table 5.7: Textualisation of rhetorical relations in English

A – independent sentence E – relative clause I – nominalisation

B – main or matrix clause F – infinitival clause J – verbless const.

C – coordinate main or matrix clause G – gerund, pres. or past part.

D – subordinate finite adverbial clause H – pres., past or adj. modifier

		A	B	C	D	E	F	G	H	I	J
SEMANTIC RELATIONS	Circumstance		1		22	4	5	8	1	6	1
	Condition				27				1		
	Elaboration	3	28	21	6	215		3	87	2	12
	Evaluation	1	1	12		16					
	Interpretation	2	3	5		9			2		
	Means			3		2	1	9	1	10	1
	Non-Vol. Cause	8	42	10	7	5	1	1	1	1	1
	Non-Vol. Result	5	6	2		9		4		1	
	Otherwise			1							
	Purpose	1	1		2	8	40		2	5	
	Solutionhood	6	15	5							
	Unconditional				1						
	Unless									1	
	Vol. Cause		1	6	24	11	1	2	3	4	1
	Vol. Result		1	5	1	12	1	4			1
PRAGMATIC RELATIONS	Antithesis		9	4		1					1
	Background	10	49	9		1			1		1
	Concession	5	31	9	9	13	1	3		1	6
	Enablement		2								
	Evidence	5	19	11	4	3		2	1		
	Justify	9	14	12	19	10		3	2	1	1
	Motivation	2	13	2	1	3			1		
	Preparation		2								1
	Restatement	2	2	2							
	Summary	3	13	2							
MULTINUCLEAR	Disjunction			1	2				2		
	Conjunction	5	34	98	11	11	4	5	7	2	4
	Contrast	1	3	19	4	2	1		1		
	Joint		3								
	List	1	2	13			5				1
	M. Restatement		1								
	Sequence	4	16	8	1		1				

Table 5.8: Textualisation of rhetorical relations in Italian

5.7 Summing up

This chapter has taken us one step further in the description of the linguistic structures of parliamentary discourse in Danish, English and Italian. After having observed the patterns in syntactic structures of texts in Chapter 4, we have here looked at the realisations of underlying structures through rhetorical relations. The description model was Rhetorical Structure Theory (RST).

First the two parameters of interest in relation to discourse structure were defined, namely that of semantic versus pragmatic source of coherence and that of rhetorical hierarchy. This was followed by a discussion of both RST and its competing theories. Despite the fact that RST is widely used in different fields of linguistics, it was necessary to elaborate on various aspects relevant to the scope of this thesis.

The study was divided into two types according to the two parameters. Firstly, the analyses showed that RST can be used to analyse argumentative texts in all three languages without changing the original concepts of the theory. Secondly, the analysis of sources of coherence showed that Danish, English and Italian tend to resort to the same relations, with almost the same distribution of semantic and pragmatic relation types. Thirdly, the analysis of rhetorical hierarchy examined to which degree syntactic coordination and subordination corresponded to those of rhetorical. The results provided information about how, in particular, syntactic coordination does not always correspond to rhetorical coordination, whereas syntactic subordination more often means that the EDUs are also rhetorically subordinated. Cross-linguistically, the differences observed were that English tended to have more rhetorically subordinated EDUs that were syntactically coordinated with other EDUs as compared with Danish and Italian. It was also observed that Italian tended to have more rhetorically coordinated EDUs that were syntactically subordinated to other EDUs. However, many of these were coordinated subordinate constructions. All in all, the results of this chapter demonstrate to which extent discourse structure patterns are affected by text structure patterns and thus underline the argument presented at the very beginning of this thesis of linguistic structures being intertwined and that it is beneficial to study them simultaneously. The results also show that even though the frequency of rhetorical relations may not differ across languages, textualisations of EDUs may. In the next chapter, I will complete the study of linguistic structures by investigating information structure, relating it to the two structures of text and discourse which have been discussed so far.

6 The information structure of parliamentary discourse

This chapter presents the results of the analysis of information structure in the corpus. My approach to information structure is based on two parameters, both focusing on how information in the texts is packaged. As in the previous two chapters, the units of analysis are Elementary Discourse Units (EDUs), which, in the context of this chapter, are investigated from an information structural point of view. The first parameter under consideration is linkage, which involves an investigation of how the EDUs of the texts are linked to each other both intersententially and intrasententially. Thus the notion of linkage is closely related to the concept of text structure presented in Chapter 4, where the syntactic dependencies and realisations of EDUs were studied. The second parameter is signal, which examines the implicit and explicit marking of rhetorical relations between EDUs. I refer to the signalling items as discourse cues and investigate how these are used cross-linguistically and how they enter into specific rhetorical relations. In this way, the notion of signal is closely related to the concept of discourse structure presented in Chapter 5. The present chapter is organised as follows: Section 6.1 provides the results of the analysis of linkage, and Section 6.2 presents the results of the signal analysis. Section 6.3 summarises the results and discusses the cross-linguistic similarities and differences. This is the last chapter of the three devoted to the analysis of linguistic structures in parliamentary discourse and thus completes the examination of how linguistic structures interact and differ cross-linguistically from each other.

6.1 Linkage of EDUs

In Section 2.1.4, the concept of information structure was related to the notions of information packaging, inspired by the works of Chafe (1976), Clark & Haviland (1977), Fabricius-Hansen (1999), Loock (2010), Prince (1986) and Vallduví & Engdahl (1996). Among these scholars, Chafe (1976, p. 28) refers to the use of packaging as follows:

> I have been using the term *packaging* to refer to the kind of phenomena at issue here, with the idea that they have to do primarily with how the message is sent and only secondarily with the message itself, just as the packaging of toothpaste can affect sales in partial independence of the quality of the toothpaste inside.

In subsequent applications of the term by functional and cognitive linguists, packaging has been related directly to the notion of information structure as can be seen from the following quotes from Clark & Haviland (1977, p. 5) and Prince (1986, p. 208):

To ensure reasonably efficient communication, … [t]he speaker tries, to the best of his ability, to make the structure of his utterances congruent with his knowledge of the listener's mental world.

Information in a discourse does not correspond to an unstructured set of propositions; rather, speakers seem to form their utterances so as to structure the information they are attempting to convey, usually or perhaps always in accordance with their beliefs about the hearer: what s/he is thought to know, what s/he is expected to be thinking about.

Accordingly, the concept of information structure is considered to constitute the structural realisation of information packaging, that is, the structuring of sentences and clauses by morphological and syntactic means (see Vallduví & Engdahl, 1996, p. 460). As mentioned above, two parameters of information packaging are considered in the sections of this chapter. Under the first parameter of investigation, information packaging is used to account for the linkage of EDUs, that is, whether the EDUs of a text are related to other EDUs inside the same sentence or to EDUs in other sentences. As we saw in Sections 2.1.2 and 2.1.4, it is widely assumed across various fields of linguistics, stylistics and rhetoric that texts containing long sentences in terms of words and EDUs are considered more difficult to understand, less reader-oriented and denser than texts with shorter sentences (cf. Fabricius-Hansen, 1996; Gabrielsen & Juul Christiansen, 2010; Piemontese, 1998; Renkema, 2004; Thornbury, 2005). However, we have also seen that texts with too short sentences or unvaried sentence lengths run the risk of being perceived as fragmented or incoherent, in particular in an EU context (*How to write clearly/Scrivere chiaro/Skriv klart*, 2011, *Writing for translation*, 2010). Here, the advice to politicians and administrative staff is a mean sentence length of 20-25 words and not to include too many insertions and embedded clauses, at least in Danish and English texts. Although there is no definite correlation between the number of words in a sentence and the number of EDUs in the texts of the corpus employed in this thesis, there is a clear tendency that the higher the number of EDUs, the higher the number of words. Consequently, the mean sentence lengths of the texts should merely be seen as indicators of cross-linguistic differences which need to be supported by qualitative investigations. As can be seen in Table 6.1, the mean numbers differ for Danish, English and Italian in some respects (e.g sentence length), whereas they are very similar in other respects (e.g. words per EDU). The total numbers of words and EDUs have been

presented previously in this thesis, see Chapter 3. It was shown that they are more or less constant across the three languages, with almost 15,000 words and 1,400 EDUs in each language. However, when calculating the total number of sentences, the Italian texts show a strong tendency to include longer sentences than the Danish and English texts. In terms of percentages, the mean Italian sentence is 56.1 % longer than the mean Danish and 40.2 % longer than the mean English. By contrast, English and Danish mean sentence lengths are more similar, English being 11.3 % longer than Danish sentences. Chapters 4 and 5 discussed how this tendency in Italian to use longer sentences affects text structure and discourse structure, resulting in a higher percentage of intrasententially related EDUs in the Italian texts than in those of the other two languages, as also revealed in the last row of Table 6.1.

	Danish	English	Italian
Number of words	14,737	14,666	14,781
Number of sentences	680	608	437
Number of EDUs	1,443	1,327	1,469
Mean words per sentence	21.67	24.12	33.82
Mean words per text	294.74	293.32	295.62
Mean words per EDU	10.21	11.05	10.06
Mean sentences per text	13.60	12.16	8.74
Mean EDUs per sentence	2.12	2.18	3.36

Table 6.1: Basic numbers and selected means

It must be emphasised that the mean numbers in Table 6.1 are less conclusive than they appear because of typological differences between the three languages. In Section 3.3, it was noted that the three languages do not signal definite articles and explicate subjects in the same manner. For instance, Italian, in contrast to Danish and English, is a pro-drop language that is capable of including the subject in the conjugation of the verb. Among other things, Italian is also able to attach personal pronouns that function as objects to infinitives and imperatives, e.g. *call him* becomes *chiamalo (call-him).* These differences could explain some of the differences between the numbers in Table 6.1, but it is nevertheless interesting that Italian, a language that is

characterised as a pro-drop language and that in many cases does not indicate the subject of a verb by means of a separate word, has a mean sentence length that is approximately 50 % longer than two languages in which the explication of constituents is obligatory.

6.1.1 A brief comparison of the L1 and L2 texts

As also mentioned in Chapter 3, I have chosen to include the translated L2 texts of the Danish, English and Italian L1 texts in Europarl. One of the reasons was that the translations in some instances proved to be useful in the annotation of rhetorical relations. Another reason was that the L2 texts have been translated in a way that is very close to the original L1 text in terms of lexis and grammar. As noted, some morphological and syntactic constructions were altered into other types, e.g. Italian gerunds are transformed into Danish coordinate or subordinate clauses, as seen in example 122), where the Italian gerund *riconoscendo* (recognising) is transformed into a subordinate finite adverbial clause *idet man anerkender* (as one recognises).

122) [Aspettiamo quindi con attenzione che la prossima Conferenza intergovernativa valorizzi definitivamente il ruolo di associazioni e fondazioni,] [**<u>riconoscendo</u>** <u>organizzazioni non profit e organizzazioni non governative come soggetti fondamentali per la crescita sociale comune europea.</u>] <ep-98-07-02.txt:29>

*Official Danish L2 [Vi venter således med interesse på, at man på den næste regeringskonference definitivt fremhæver foreningernes og fondenes rolle,] [**<u>idet man anerkender</u>** nonprofitorganisationernes og de ikke-statslige organisationers vigtighed for den fælles samfundsudvikling i Europa.]*

*Litt. translation of Italian L1 [Let us therefore hope that the next Intergovernmental Conference finally promotes the role of organisations and foundations,] [**<u>recognising</u>** non-profit organisations and non-governmental organisations as fundamental bodies for common social growth in Europe.]*

*Litt. translation of Danish L1 [Let us therefore hope that the next Intergovernmental Conference finally promotes the role of organisations and foundations,] [**<u>as one recognises</u>** non-profit organisations and non-governmental organisations as fundamental bodies for common social growth in Europe.]*

My general impression of the official L2 translations was, however, that the translators had tried to reproduce the L2 text as closely to the original as possible, lexically and grammatically speaking. Table 6.2 shows transformation patterns between the original L1 texts and the L2 translations. The differences in percentages between the L1 source text and the L2 translated texts are shown in parentheses in the two L2 columns. In the English and Italian L2 texts of the

Danish L1, we can see that the number of words is higher than in the Danish L1[9], that the number of sentences is almost similar in the L1 and L2, and that the mean sentence length is slightly higher (around 10 %) in L2 than in L1 texts. Similar differences are found in the L2 translations of the English and Italian L1 texts, where the mean sentence lengths are approximately 10 % higher or lower than those of the L1 texts. The differences in average sentence length observed between the L1 texts in Danish, English and Italian in Table 6.1 can therefore not only be caused by language typological divergences between the three language systems.

	Danish L1	English L2	Italian L2
Number of words	14,737	16,732 (+13.5 %)	15,799 (+7.2 %)
Number of sentences	680	699 (+2.8 %)	672 (÷1.2 %)
Mean words per sentence	21.67	23.94 (+10.5 %)	23.51 (+8.4 %)
	English L1	**Danish L2**	**Italian L2**
Number of words	14,666	13,718 (÷6.9 %)	14,456 (÷1.5 %)
Number of sentences	608	618 (+1.6 %)	566 (÷7.4 %)
Mean words per sentence	24.12	22.20 (÷8,6 %)	25.54 (+5.9 %)
	Italian L1	**Danish L2**	**English L2**
Number of words	14,781	15,569 (+5.3 %)	15,909 (+7.6 %)
Number of sentences	437	496 (+13.5 %)	482 (+10.3 %)
Mean words per sentence	33.82	31.39 (÷7.7 %)	33.01 (÷2.5 %)

Table 6.2: Comparison of L1 and L2 corpora

While I shall refrain from commenting on the possible translation strategies which could have resulted in the discrepancies between the L1 and L2 texts, it is nevertheless worth noting that the

---

[9] In modern translation studies concerned with universal patterns, it is assumed that translated texts are longer than their source texts because of the need for 'explicitation' (Becher, 2011; Pym, 2005; Toury, 1995), so this result was expected.

Danish and English L2 of the Italian L1 contain over 10 % more sentences than the source texts. This is not the case in the translations of the Danish and English L1 texts, and a qualitative examination of this reveals that it is typically the long Italian sentences with colons and semicolons that have been broken up into two or three independent sentences in the Danish and English translations. In example 123), an Italian sentence consisting of five EDUs with one colon and one semicolon has been transformed both in the Danish and English official L2 texts into three sentences, replacing the colon and semicolon with full stops.

123)   [Ricordo solo che la rivoluzione culturale esige un tempo più lungo:] [qui abbiamo solo cambiato alcuni posti] [che era vitale cambiare] [perché bisognava dar l' esempio di mobilità, per paesi e per persone;] [bisognava dimostrare che i posti non sono appannaggio fisso di nessuno.] <ep-99-10-05.txt:135>

Official Danish L2 [Jeg vil blot minde om, at den kulturelle revolution kræver længere tid.] [Her har vi jo kun ændret nogle stillinger,] [som det var vigtigt at ændre,] [da det var nødvendigt at gå foran med et godt eksempel på mobilitet for lande og personer.] [Det var nødvendigt at vise, at der ikke er nogen, som har en fast fortrinsret til stillingerne.]

Official English L2 [I would only remind you that the cultural revolution needs more time.] [Here we have only changed some posts] [which it was essential to change] [because we had to give an example of mobility, for countries and people.] [We had to show that the posts are not anyone's fixed prerogative.]

6.1.2 EDUs per sentence

Although it is difficult to see from the mean number of EDUs per sentence row in Table 6.1, there are noteworthy differences in the distribution between the numbers in the three languages. As we can see in Table 6.3 showing the distribution of EDUs per sentence in all the L1 texts, the typical pattern in Danish texts suggests that one EDU in many cases (40.3 %) corresponds to one sentence. In English, the distribution shows the same percentages for the occurrences of one and two EDUs per sentence (34.5 % and 34.3 % respectively), followed by a considerable number of sentences with three EDUs (19.3 %). In Italian, by contrast, the highest concentration of EDUs is found in sentences with three EDUs. In fact, each triangle in Table 6.3 that represents Italian sentences with one, two, three and four EDUs constitutes approximately 20 % of the total number of EDUs. This entails that sentences with up to four EDUs in Italian amount to 78.1 % of all EDUs, whereas the equivalent numbers in Danish and English are much higher: 93.0 % and 96.0 % respectively. As argued above, we can also see that the Italian sentences contain more EDUs than the Danish and English counterparts. In fact, there are a small number

of Italian sentences with more than nine EDUs; the Danish and English texts contain only very few sentences with more than seven EDUs (0.3 % in Danish and 0.4 % in English).

EDUs per sentence	1	2	3	4	5	6	7	8	9	10	11	12	13
Danish	40,3	31,7	14,6	6,4	4,5	1,2	1	0,3	0	0	0	0	0
English	34,5	34,3	19,3	7,9	2,1	0,8	0,7	0,2	0,2	0	0	0	0
Italian	17,3	21	21,6	18,2	9,3	4,8	3,4	2,5	1,1	0,2	0	0,2	0,2
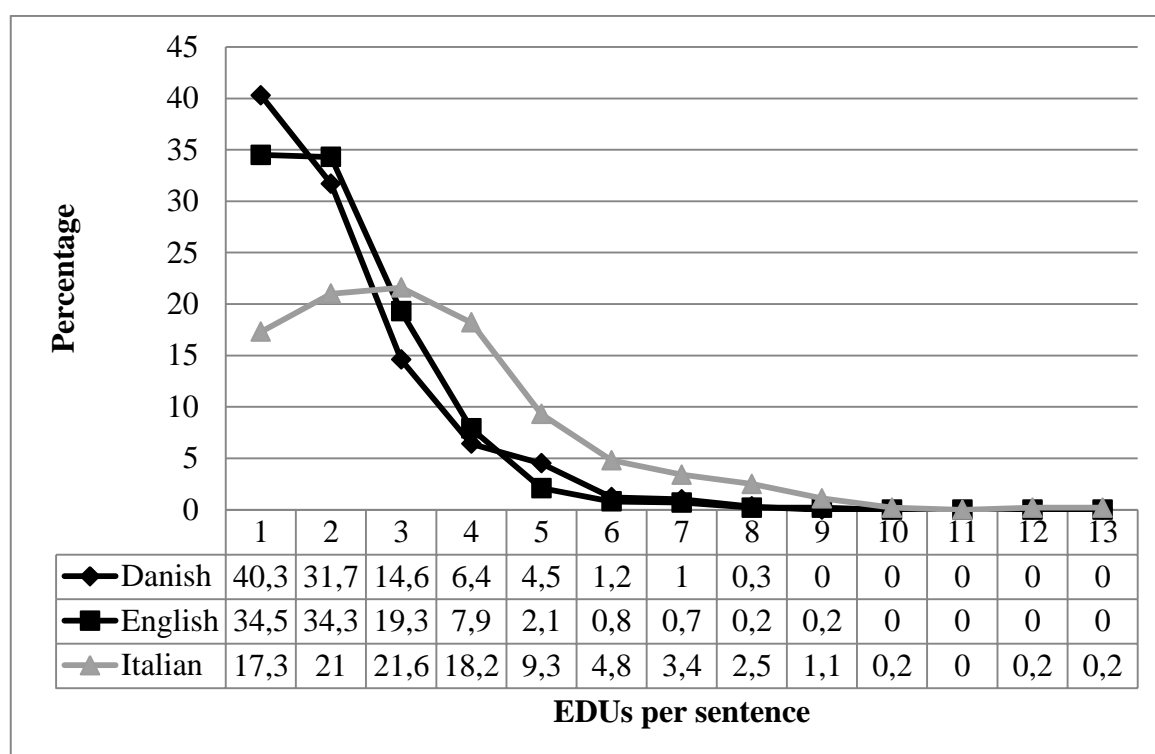
Table 6.3: Distribution of EDUs per sentence

These numbers confirm that the longer sentence lengths found in the Italian texts are not only due to differences in typological language characteristics between Danish, English and Italian, but also due to differences between the mean numbers of EDUs per sentence in the three languages. For instance, Italian sentences contain more EDUs and thus more information than Danish and English sentences. Example 124), containing an Italian sentence with nine EDUs, illustrates how a relatively large amount of information can be packaged inside the same sentence. The interesting thing about this example is the way the underlined satellites are structured linguistically as subordinate EDUs. The first two (*quando si tratta/when it is a matter* and *come io sostengo/as I point out*) are subordinate finite adverbial clauses, and the three following (*sottoscritto/supported*, *contententi/containing* and *testate/tested*) are postmodifiers. Thus the satellites are integrated in their respective nuclei and in various ways dependent on them, see the deverbalisation scale presented in Chapter 4. This means that cognitively they play a less salient role in the discourse structure of the sentence than if the satellites had been textualised as independent sentences (cf. Renkema, 2004, p. 147). Furthermore, satellites textualised as subordinate clauses are typically marked by discourse cues, here *quando* (when)

and *come* (as), revealing their argumentative functions in relation to the main claim of the sentence found in the nucleus (*it is essential to carry out an initial set of experiments on animals*).

124) **[Ma** [quando, signor Commissario, [come io sostengo nell'emendamento] [sottoscritto da oltre cinquanta parlamentari,] si tratta di nuovi cosmetici] [contenenti ingredienti nuovi,] [mai testati sperimentalmente prima] [al fine di caratterizzarne il profilo tossicologico negli animali da laboratorio,] **in tali condizioni io sono convinto, da scienziato, che sia obbligatorio procedere ad una prima fase di sperimentazione animale]** [prima dell'uso nell'uomo] [e prima dell'immissione in commercio.] <ep-01-04-02.txt:42>

[However, [when, Commissioner, [as I point out in an amendment] [supported by over 50 Members of Parliament,] it is a matter of new cosmetics] [containing new ingredients] [never tested in the past] [in order to establish their toxicological profile in laboratory animals,] ***then, I am convinced, as a scientist, that it is essential to carry out an initial set of experiments on animals]*** *[before the use on human beings] [and before the emission on the market.]*

In contrast to the strategy above of downgrading satellites into subordinate EDUs of various types, example 125) from a Danish text with eleven EDUs applies an upgrading strategy. Here, the satellite EDUs are not always reduced to subordinate clauses. The first sentence introduces the main nucleus (*beef must be labelled with regard to provenance*), while the following sentences fulfil different rhetorical functions: the underlined independent sentence *Målet med denne ordning* (The purpose of the regulation) expresses a causal relationship (Volitional Cause) with the previous sentence. By comparing this way of upgrading the information of the sentence to an independent sentence with the downgrading of the Italian counterpart in example 124), where a causal relationship (Purpose) was textualised as an infinitve clause (*al fine di caratterizzarne/in order to establish)*, we can gain further insight into the typical information structure patterns of the two languages. As mentioned earlier, English in this respect resembles Danish more than it resembles Italian.

125) [Jeg går ind for, at oksekød mærkes med det eller de lande,] [kødet kommer fra,] [og jeg er stærkt imod at indføre et fælles EU-mærke,] [hvor forbrugerne ikke kan se, hvilke lande der er tale om.] [Målet med denne ordning må være at opnå, at forbrugerne har tillid til kød fra samtlige EU-lande.] [Indtil det er en realitet,] [har forbrugerne krav

på at vide præcis, hvilket land kødet kommer fra.] [Jeg har stillet et ændringsforslag om, at slagteriets autorisationsnummer ikke skal fremgå på etiketten.] [Med mit forslag ønskede jeg at sikre, at små og mellemstore slagterier fortsat kan levere kød til opskæring.] [Dette autorisationsnummer giver ikke forbrugeren nogen anvendelig information] [og har ingen konsekvens for sporbarheden.] <ep-00-04-11.txt:271>

[I am in favour of beef being labelled with the name of the country or countries] [it comes from] [and I am strongly against introducing a common EU label] [from which consumers cannot see which countries are involved.] [*The purpose of the regulation in question must be to achieve that consumers have confidence in meat from all EU countries.*] *[Until this is a reality,] [consumers need to know precisely which country their meat comes from.] [I have tabled an amendment to the effect that the slaughterhouse's authorisation number should not be shown on the label.] [In this way, I wished to ensure that small and medium-sized slaughterhouses could continue to supply meat to butchers.] [The authorisation number does not provide the consumer with any useful information] [and has no bearing upon traceability.]*

At the end of the previous chapter, in Section 5.6, I gave a number of examples of how the satellite status of the Evaluation relation was textualised differently across the three languages – in Danish and English as independent sentences and in Italian as coordinate insertions or relative clauses. From a rhetorical perspective, this difference shows how Danish and English use full stops to emphasise the importance of the satellite EDU – although the information contained in this EDU is still considered to be less important than the information in the nucleus, cf. the section on RST in Chapter 5. Furthermore, the more frequent use of full stops in Danish and English also indicates a higher tendency in these two languages to vary sentence length more frequently than in Italian, cf. the remarkable differences in Table 6.3 between the Danish, English and Italian sentences with only one EDU (40.3 – 34.5 – 17.3 %). In examples 126) and 127), we can see how this stylistic emphasis differs in English and Italian. The first example from English shows how a sentence containing seven EDUs is followed by a sentence with just one EDU, here underlined.

126)   [Not only does it transfer decision-making away from the Commission] [and bring it to the Council,] [but it does so via meetings] [for which there are no prior published papers or minutes,] [from which the European Parliament and all those [committed to openness and transparency in decision-making] are excluded,] [and where the need to write papers diverts already stretched Commission staff away from their proper job of managing humanitarian aid effectively.] [The delay in translating those papers into

eleven different languages can be measured in terms of extra deaths.] <ep-96-05-20.txt:33>

The last sentence of this English example is textualised as an independent sentence with one EDU in order to highlight the consequences (*more people die*) of the situation presented in the previous sentence (*a new committee diverts staff away from their proper job of managing humanitarian aid*). In this way, the writer is able to direct more attention to the serious consequences than if he had included more EDUs in the last sentence. In the Italian example in 127), we can see what effect the opposite strategy has. Here, the writer starts with a sentence consisting of 13 EDUs followed by a sentence with eight EDUs, again underlined. The main nuclei of this excerpt are found in two coordinate clauses (*e, quindi, noi siamo assolutamente d'accordo… e siamo convinti che…/and we are therefore totally in agreement… and we are convinced that*) located in the second sentence and surrounded by a series of satellite EDUs. This means that no emphasis is accorded by the writer, and it is up to the reader to separate the individual units of information from each other.

127) [In questi anni non abbiamo esitato ad indicare nella Repubblica di Serbia il responsabile principale delle drammatiche vicende, specialmente in Bosnia,] [e abbiamo spesso difeso con nostre risoluzioni il diritto della Croazia a vedersi restituito il territorio] [occupato dai serbi,] [ma proprio per questo, nel momento [in cui dopo gli accordi di Dayton, si avvia una situazione di maggior tranquillità nell'ex Jugoslavia] [- e si deve avviare questa situazione -] noi non possiamo, oggi, con la stessa forza [con cui abbiamo denunciato errori e crimini di altri paesi] non denunciare quello] [che avviene nella Croazia stessa,] [per quanto riguarda il rispetto dei diritti dell'uomo, il rispetto delle minoranze etniche, la libertà di stampa, il pluralismo della stampa e, non ultima, la questione [a cui ha accennato lei, onorevole Fassino,] dell'impegno] [che la Croazia ha di collaborare con il Tribunale dell'Aja] [e consegnare coloro] [i quali sono sospettati di gravi reati.] [Tutto questo non è avvenuto] [e, quindi, noi siamo assolutamente d'accordo con la decisione del Consiglio d'Europa e con la decisione dei governi dell'Unione] [che lei ci annunzia] [e siamo convinti che, [se in questi mesi [in cui vi sono le truppe dell'IFOR,] la comunità internazionale non riesce a determinare un clima di trasparenza e di rispetto dei diritti dell'uomo,] diventerà molto difficile per noi allontanarci da quelle

terre] [sapendo che [non appena l'IFOR va via,] c'è il rischio che tutto torni ad incendiarsi.] <ep-96-06-05.txt:85>

[In these years we have never hesitated to point the finger at the Serb Republic as primarily responsible for the tragedies, especially in Bosnia,] [and our resolutions have frequently defended the right of Croatia to have the territory] [occupied by the Serbs restored to it,] [but precisely because of this, at a time [when a more peaceful situation is developing in former Yugoslavia after the Dayton agreements] [- and this situation must develop -] we cannot fail to condemn today, as forcefully [as we have condemned mistakes and crimes in other countries,] what [that is going on in Croatia itself] [as regards respect for human rights, respect for ethnic minorities, press freedom, press pluralism and, last but not least, the point [that you mentioned, Mr Fassino,] about commitment [that Croatia has] to cooperate with the Hague Tribunal] [and hand over those] [who are suspected of serious crimes.] [None of this has happened,] [and we are therefore totally in agreement with the decision by the Council of Europe and the Union governments] [that you have announced] [and we are convinced that [if the international community fails to establish a climate of transparency and respect for human rights] [while the IFOR troops are there,] it will become extremely difficult for us to leave [knowing that [as soon as IFOR departs] there is the risk of everything going up in flames again.]

In this Italian example, we can also observe a high number of embedded EDUs textualised as relative clauses and subordinate finite adverbial clauses, which writers were discouraged from using in the previously cited EU style guides. Similarly, long sentences were to be avoided. The results presented in this section could be seen as indications of Italian not observing these rules, even though, as noted in the Italian version of the style guide (*How to write clearly/Scrivere chiaro/Skriv klart*, 2011), longer and more complex sentence structures are more acceptable in this language than in other languages. However, the results could also be interpreted as Danish and English writers preferring clear indications through full stops of information boundaries, whereas Italian writers do not necessarily perceive a one-to-one relationship between sentence and information boundaries. In Section 6.2, the investigation of the information structure of the three languages continues with an examination of the signalling of rhetorical relations.

6.2   Signalling rhetorical relations

The semantic or pragmatic source of coherence in a rhetorical relation may be signalled in a number of ways. The prototypical and perhaps most studied signal is that of discourse cues such as coordinators (128), subordinators (129), conjunct adverbs (130), prepositional phrases (131) and clauses which take sentential complements (132) (Quirk, Greenbaum, Leech, & Svartvik,

1972). Definitions of these five types of discourse cues are presented below. All examples are taken from the English part of the corpus.

128)   [It is important to sort out the structural problems in our economies] **[but** this will take time.] <ep-96-07-03.txt:75>

129)   [**If** he can get those types of concessions from the arbitration deal on 25 September] [he will have done the Windward Islands a tremendous favour.] <ep-97-09-16.txt:102>

130)   [It represents a single, simple, safe, harmonized regime] [that all sides of this House can support.] [**However**, I should like to focus the Commission's mind on another issue.] <ep-97-11-05.txt:183>

131)   [Of course, at its core is equal access to Community waters] [which was derogated until 1 January 2003.] [Only **at this point** will the CFP fully come into force.] <ep-02-05-29.txt:37>

132)   [The delay in translating those papers into eleven different languages can be measured in terms of extra deaths.] [**It seems to us that** the Council is more interested in playing the politics of Europe] [than in assisting people in distress.] <ep-96-05-20.txt:33>

The definitions of these five syntactic classes are as follows, see also Knott (1996, pp. 66–67):

- Coordinators: these always appear between the clauses they link; the clauses may occur in separate sentences or in the same sentence. If combined in a sequence with other discourse cues, coordinators always appear leftmost in the sequence. In English, typical coordinators include *and, but,* and *or*.
- Subordinators: these introduce subordinate clauses in complex sentences. The subordinate clause may occur to the left or the right of the main clause, but the subordinator is always to the left of the subordinate clause. In English, typical subordinators include *although, because,* and *if*.

- Conjunct adverbs: these relate whole clauses, and can appear at different points within them, although there is often a default position for particular cues. There are also syntactic constraints on exactly which positions conjunct adverbs can occupy: at the beginning of a clause, between subject and verb, between any auxiliary verbs, between auxiliary verb and main verb, after a copula if there is one, and before a sentential complement if there is one. In English, typical conjunct adverbs include *as a consequence, however,* and *thus.*
- Prepositional phrases: these often contain propositional anaphora referring back to the previous clause. In English, typical prepositional phrases include *at this point, for these reasons,* and *in this respect.*
- Clauses which take sentential complements: these often introduce a particular intentional stance with respect to the content of the clause they introduce. In English, typical phrases of this kind include *it follows that* and *it may seem that.*

Discourse cue is intended as a general term for those words that signal discourse coherence in texts. The term 'cue' is taken from Grosz & Sidner (1986) and corresponds essentially to other proposals found in the literature including 'discourse connectives' (Forbes et al., 2002), 'discourse particles' (Aijmer, 2002; Mosegaard Hansen, 1998; Siegel, 2002), 'discourse markers' (Schiffrin, 1987; Taboada, 2006) or simply 'connectives' (Bazzanella, 1990; Skytte & Korzen, 2000). Discourse cues have been studied from a range of perspectives and approaches, e.g. as signalling a sequential relationship between sentences (Fraser, 1999), with regard to gender (Erman, 1992) and age (Andersen, 2001), from a relevance-theoretic point of view (Blakemore, 2002), and as marking discourse coherence (Lenk, 1998; Schiffrin, 1987). The approach adopted in this thesis mainly falls within that of discourse coherence, following Schiffrin (1987, p. 49), who states that the 'analysis of discourse markers is part of the more general analysis of discourse coherence – how speakers and hearers jointly integrate forms, meaning, and actions to make overall sense out of what is said'. The analysis presented below can be considered as a formalisation of the informal definitions found in Schiffrin on markers like discourse coherence, when applying the set of rhetorical relations of Rhetorical Structure Theory described in Chapter 5. The section below explores how rhetorical relations are signalled by discourse cues, which, I argue, can be considered to be another aspect of how information is packaged in texts.

6.2.1 Previous studies on signalling rhetorical relations

In the previous chapter, we saw that there are rhetorical relations between the individual discourse units of a text. It was briefly discussed how these relations can be signalled by the writer to highlight the structure of a given text or alternatively remain implicit. In fact, the analysis of 150 texts of parliamentary discourse in Danish, English and Italian shows that in a majority of the instances the rhetorical relation is not signalled by any discourse cues. This finding is very much in line with the results of other studies (Knott & Dale, 1994; Schauer & Hahn, 2001; Taboada, 2009). In an RST context, Taboada & Mann (2006a, p. 441) estimate that 'over 50 per cent of the [RST] relations being unsignalled is representative' for studies of various text types. Accordingly, we must assume that even if a discourse cue is omitted between two discourse units, the coherence and the meaning of those two units are typically preserved. In the previous chapter, I referred to Traum's (1993) assumption that readers do not need to recognise every single rhetorical relation holding between the EDUs of a text to understand a text as long as the reader agrees with the writer on the overall purpose of the text. This assumption was originally thought to explain why discourse theorists cannot agree on which and on how many rhetorical relations should be included in the description of discourse structure. But it can also be used to account for why writers do not explicate rhetorical relations more than they apparently do. To understand the cognitive difference(s) between covert and overt signalling of rhetorical relations a number of scholars have carried out various experiments testing functions of discourse cues – typically coordinators and subordinators. Some of these (Haberlandt, 1982; Meyer, Brandt, & Bluth, 1980; Sanders & Noordman, 2000; Sanders et al., 1992) suggest that texts with many explicit discourse cues are easier to process and reproduce for readers than texts without discourse cues. For instance, Degand and colleagues (Degand, Lefèvre, & Bestgen, 1999; Degand & Sanders, 2002) show how the presence of discourse cues, in their case subordinators, increase the ability of readers to comprehend the content of texts. In an experiment on Italian discourse cues, Soria & Ferrari (1998) compared explicit marking of relations with a number of rhetorical relations. Their results suggest that explicit marking of rhetorical relations is not entirely optional but at least partially constrained by the type of relation that is signalled by the means of expression. Again, this corresponds to previous and subsequent findings in other languages, see Taboada & Mann (2006a, pp. 436–437). Studies in conversation analysis reveal the same tendencies. In spoken dialogue, it is argued that discourse cues are partly responsible for establishing the common ground between conversation participants: discourse cues guide the 'grounding', understood as a collaborative effort in which

the participants send signals of understanding to each other during the conversation (Byron & Heeman, 1998; Clark & Schaefer, 1989; Clark & Wilkes-Gibbes, 1986). All of the above-mentioned experiments and studies are reader-oriented in that they study how explicit marking of discourse cues affects the comprehension, the processing and the ability to recall the content of a text from the reader's point of view. This is also the view employed in the present study, even though nothing excludes discourse cues from helping the writer to organise her own ideas in a structured and coherent manner.

6.2.2 Issues related to the study of discourse cues

Apart from the general issue of discourse cues not being signalled more frequently than observed, at least two other problems can be identified. Schiffrin (1987, p. 52) recognises the first challenge in accounting for the scope of a discourse cue, namely how much of the preceding or following text is affected by the use of a discourse cue. In the context of RST as applied in this thesis, this problem is solved by the hierarchical tree structure that EDUs enter into when linked together by rhetorical relations. Firstly, a discourse cue has direct 'scope' over the text spans that constitute the same schema when it is found within one of them. Secondly, a discourse cue has indirect scope over the higher and lower levels of that schema, and thirdly, a cue has only marginal scope over the 'sister' relations of that schema. An exemplification of this is provided in Figure 6.1, which consists of three sentences with four EDUs in total. In the figure, the two discourse cues *because* and *therefore* have been represented in capital letters to ease identification. The figure shows three relations: the first is Elaboration between the nucleus in EDU#1 and the satellite span in #2-4. This relation is not signalled by any cues. The second relation is Volitional Cause, composed of two spans of text (labelled number #2 and #3) and signalled by the discourse cue *because*. The third relation is Volitional Result, whose nucleus is span #2-3 and whose satellite is #4. This relation is signalled by the cue *therefore*. By representing the rhetorical relations in this way, the scope of *therefore* is everything in those three spans (#2-4), whereas the scope of *because* is only spans #2 and #3, since the Volitional Cause relation is embedded within the Volitional Result relation. I shall refrain from elaborating more on this here, but refer the reader to Section 5.1 on RST tree representations.

[My group cannot support Amendment No. 1.] [We cannot support the amendment] [because it only addresses a minor part of the problem.] [Therefore, we will vote against it tomorrow.]
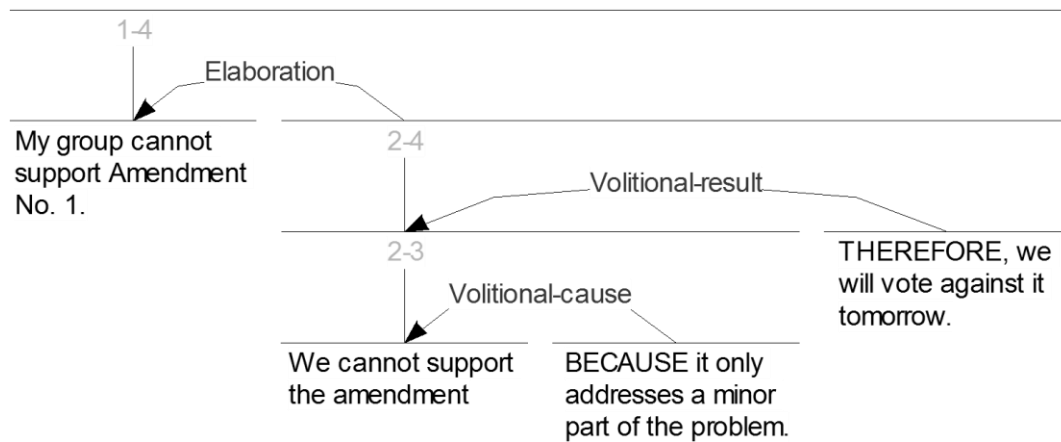
Figure 6.1: RST tree with discourse cues

The second problem in the study of discourse cues is related to the challenge of defining them. At the beginning of this section, some of the terms used in the literature were introduced, all referring to the notion of discourse cues. In spite of various arguments for the choice of using one particular term rather than another, there is little agreement as to which linguistic items constitute the objects of analysis. Suggestions range from multi-word expressions such as *to return to my original point* (Fraser, 1988) to single-word expressions such as *well* and *like* (Hasund, 2002), *or* and *but* (Schiffrin, 1987), *oh* and *mhm* (Jucker & Smith, 1998). In this thesis, my aim was include all discourse cues that in one way or another could mark the rhetorical relations presented in the previous chapter – a decision that by no means was an easy one, especially because the inclusion or exclusion needed to be valid for all three languages under investigation. However, the transcription conventions made by the EU administration of the parliamentary proceedings eased the analysis considerably: pause fillers such as *I mean, you know, oh, mhm* or similar items which are typically ascribed spoken discourse (see Muller, 2005, p. 19ff) were excluded from all texts, even though the speakers probably made use of these during their speeches. The choice was to focus on some specific cues that could be described as discourse connectives in their widest sense. I have taken into consideration discourse cues that are usually labelled as coordinators, subordinators, conjunct adverbs, prepositional phrases and clauses which take sentential complements (Quirk et al., 1972), as exemplified at the beginning of this chapter.

Again, the criteria were that they were capable of expressing one of the rhetorical relations in RST. Even though Forbes et al. (2002) argue that sentence adverbials such as *nevertheless, also, on the other hand* should not be treated as discourse cues since they do not contribute to establishing a structural connection between EDUs, but an anaphoric one, I have included in the present study. I have done so because I do not see structural and anaphoric connections to be two mutually exclusive types of linkage. In fact, many of the discourse cues found in my corpus also contain anaphoric expressions. Take, for instance, the following prepositional phrases found in the Italian texts which both indicate consequential content and anaphoric connections: *per tali motivi (for these motives), per tutte queste ragioni (for all these reasons), per tutti questi motivi (for all these motives), per l'insieme di queste riflessioni (for all of these reflections), come conseguenza di tale decisione (as a consequence of that decision)* and *per i motivi già esposti (for the motives already mentioned),* cf. also the discussion in Section 2.1.5 of the interplay between cohesion and coherence including that of anaphora and rhetorical relations.

I have previously discussed how other elements in a sentence or clause may function as discourse cues. In Chapter 5, we saw how the semantics of a verb (*cause*) indicated a causal relationship between a nucleus and a satellite, how the presence of the noun *purpose* pointed towards a purpose relationship between a relative clause and its matrix clause. These two types of discourse cues (verbs and noun phrases) have not been included in the following analysis of discourse cues, because it proved very difficult to define specific criteria of when to include these and when not to do so. As previously mentioned in Chapter 4, mood and tense are not considered as discourse cues either for the reasons stated above, even though they may ease the interpretation of which rhetorical relation holds between two EDUs. Furthermore, mood and tense do not constitute a valid parameter of comparison in this study, because the three languages involved here do not have at their disposal the same range of moods and tenses; the Italian language system contains more past tenses than Danish and English and also more moods, e.g. the conditional mood (see Maiden & Robustelli, 2007, p. 219ff).

6.2.3   Linguistic items used as discourse cues

Bearing in mind the basic differences between the textualisations in the three languages presented in Table 4.7 and Table 4.8, it would seem beneficial to conduct a survey of the linguistic items used as discourse cues in order to examine whether these differences affect the distribution of discourse cue items. Moreover, in order to understand which types of discourse cues will be considered in this chapter, I have chosen to introduce these before presenting the

distribution of implicit and explicit signalling. Table 6.4 shows the ten most frequently used discourse cues in Danish, English and Italian texts. After each discourse cue, the frequencies in relation to the total amount of discourse cues found are shown in parentheses; for the Danish and Italian cues, translations are provided in square brackets. As can be seen in the table, the two most frequent discourse cues in all three languages are the two coordinators *and* and *but* together with their cross-linguistic counterparts *og/men* in Danish and *e/ma* in Italian. The percentages indicate that Danish and English exhibit higher tendencies to use *og* (21.4 %) and *and* (22.4 %) than Italian uses the corresponding cue *e* (14.0 %). By looking at the distribution of the other nine discourse cues, we can see that many of the same items are found across the three languages, although their distributions differ slightly, e.g. *hvis* (2.8 %) in Danish compared to its cross-linguistic counterpart *if* (4.9 %) in English and *se* (3.8 %) in Italian.

	Danish	**English**	**Italian**
1	og (21.4 %) *[and]*	and (22.4 %)	e (14.0 %) *[and]*
2	men (10.1 %) *[but]*	but (7.1 %)	ma (6.2 %) *[but]*
3	når (6.1 %) *[when, as, if]*	as (5.5 %)	per (5.7 %) *[to]*
4	også (4.9 %) *[also]*	to (5.1 %)	anche (5.3 %) *[also]*
5	hvis (2.8 %) *[if]*	when (4.9 %)	come (4.8 %) *[as]*
6	derfor (2.7 %) *[therefore]*	if (4.9 %)	perché (4.3 %) *[because]*
7	fordi (2.7 %) *[because]*	also (3.8 %)	se (3.8 %) *[if]*
8	jo (2.5 %) *[as you know, of course]*	however (3.5 %)	e + *another cue* (3.4 %) *[and]*
9	så (2.5 %) *[so]*	firstly *(and similar)* (2.9 %)	quindi (3.2 %) *[therefore, so]*
10	men + *another cue* (2.2 %) *[but]*	so (2.9 %)	infatti & però (2.1 %) *[as you know etc.]* & *[however]*

Table 6.4: The ten most frequent discourse cues in the Danish, English and Italian texts

The differences between the textualisations of EDUs and between the distributions of discourse cues in relation to the deverbalisation scale mentioned above do not seem to have affected the linguistic items used as discourse cues found across Danish, English and Italian to any significant degree. However, it is interesting to observe that some types of discourse cues are present in some of the languages' top ten lists while absent from the other two. For instance, the

Danish discourse cue *jo*, often translated with *of course* in the official English EU translations, is a very typical discourse particle used to express a kind of presupposition from the writer's perspective. In the context of RST, this cue is used in satellites to increase the reader's belief of the situation presented in the nucleus (Evidence). In the same way, it is interesting to observe that among the ten most frequent discourse cues in English, we find a discourse cue that reveals textual organisation in the form of *firstly*, which I have grouped together with *secondly*, *thirdly*, *first of all* etc. The reason that it is interesting to find this discourse cue among the most frequently used must be seen in the light of the general tendency in English to use fewer discourse cues compared to Danish and Italian. An example of this discourse cue is found in 133).

133)   [There are three main points I would like to make.] [**Firstly**, events in Seattle and Prague show that …] […] [**Secondly**, like Mrs Maij-Weggen, I want to talk about Burma.] […] [**Thirdly**, this summit is being held on the Korean peninsula in Seoul.] […] <ep-00-10-02.txt:84>

Lastly, in the Italian list of discourse cues, the discourse particle *infatti* is worth mentioning. *Infatti* resembles in many aspects the Danish *jo* in that it in most cases expresses an Evidence relation between the satellite and nucleus, see example 134).

134)   [Questa scelta comporta necessariamente una severa convalida da parte della Comunità europea.] [Questa **infatti** deve poter verificare ex ante, tramite la Commissione esecutiva, la congruità e la conformità di questi investimenti all'obiettivo di consentire all'Unione europea di diventare, entro il 2010, l'area più competitiva e più dinamica di una società] [basata sulla conoscenza, sulla piena occupazione e sullo sviluppo sostenibile,] [favorendo altresì il loro coordinamento.] <ep-02-10-21.txt:49>

*[This decision requires, of necessity, strong support from the European Community.] [**Indeed**, the Community must be able, through the Commission, to assess, ex ante, the compatibility and conformance of the investments with the objective of making the European Union the most competitive and dynamic economy] [based on knowledge, full employment and sustainable development in the world by 2010,] [facilitating also the coordination of these investments.]*

6.2.4   Implicit or explicit signalling of rhetorical relations

As mentioned above, discourse cues can either be explicated or remain implicit. Table 6.5 displays the overall distribution of explicitly and implicitly marked discourse cues in the Danish,

English and Italian texts of the corpus. The percentages represent the number of times a discourse cue is present or absent between two text spans of EDUs. That is, in Danish, 46 % of the total number of rhetorical relations annotated in the texts has been interpreted through a discourse cue. The Danish and Italian texts exhibit very similar distributions of explicit and implicit signalling of rhetorical relations, whereas the English texts make less frequent use of explicitly marked discourse cues (33.2 %).

	Danish	**English**	**Italian**
Explicit signalling	46.0 %	33.2 %	45.9 %
Implicit signalling	54.0 %	66.8 %	54.1 %

Table 6.5: Distribution of explicit and implicit signalling of rhetorical relations

This difference between Danish/Italian and English is interesting from a number of perspectives. Firstly, because the differences adhere to the differences observed between English and Spanish spoken discourse in Taboada (2004b, p. 149), who reports on a tendency in Spanish (45 %) to mark rhetorical relations with discourse cues more frequently than in English (30 %). Accordingly, this suggests that the tendency found in the Europarl corpus of English marking rhetorical relations less frequently than the other two languages could be considered as a more general pattern across text types. Secondly, the differences are interesting because they reveal different degrees of underspecification across the three languages. Spooren (1997) refers to various degrees of underspecification of discourse cues such as ambiguous and implicit cues: an ambiguous cue could be the coordinator *and*. Following this idea of underspecification, the English texts investigated in the present thesis could be described within the highest degree of underspecification as they resort more frequently to implicit signalling of rhetorical relations than do Danish and Italian. Bearing in mind the previously cited style guides from the EU (*How to write clearly/Scrivere chiaro/Skriv klart*, 2011, *Writing for translation*, 2010) and the Gricean maxims (Grice, 1975) discussed in Chapter 2, we may regard English as observing quite strictly the advice of keeping it short and simple (i.e. the KISS principle found in the style guides) and Grice's maxims of 'quantity' and 'manner'; cf. also Spooren (1997, p. 150), who, in this regard, refers to Traugott & König's (1991) 'R-principle' of not saying 'more than necessary'. But the differences could also be interpreted in the inverse direction as indicators of Danish and Italian being more reader-friendly than English, because they provide explicit cues of how to decode the relations between EDUs to a higher extent than English, cf. the above-mentioned psycholinguistic experiments by Sanders and Degand, among others, at the beginning of this

section. Whichever interpretation is found to be the most plausible, the overall numbers in Table 6.5 call for further investigation.

One way of pursuing this further is to carry out a survey of the distribution between explicit discourse cues that are used to link EDUs inside the same sentence and discourse cues that link EDUs across the sentence boundary. In this way, we may be able to account for the lower frequency of explicit cues in English. As an initial comment, it should be noted that discourse cues that relate EDUs inside the same sentence are considered to be more essential than discourse cues that link EDUs across different sentences. Intrasentential discourse cues are in most cases obligatory between coordinated clauses and between matrix and finite subordinate adverbial clauses; exceptions are, of course, between EDUs separated by colons or semicolons and between matrix and relative clauses, where discourse cues are optional, along with other constructions. An example of two EDUs without any cues, separated by a semicolon is shown in 135).

135) [Linguistic and cultural diversity is not about defining or redefining boundaries;] [it underpins the cultural strength of our enlarging European Union.] <ep-03-01-13.txt:64>

Across the sentence boundary, discourse cues are also more optional, and when explicated, they are typically textualised as conjunct adverbs or clauses which take sentential complements. Table 6.6 displays the overall distribution of the scope of discourse cues in Danish, English and Italian.

	Danish	English	Italian
Discourse cues linking EDUs across different sentences	28.3 %	25.4 %	26.2 %
Discourse cues linking EDUs inside the same sentence	71.7 %	74.6 %	73.8 %

Table 6.6: Distribution of discourse cues in relation to linkage

As noted above, explicit discourse cues are mainly used to link EDUs inside the same sentence, with almost 75 % of the occurrences across the three languages. As such, no notable differences between English and the other two languages can be found in the distribution of discourse cues, although English displays the lowest frequency of discourse cues that link EDUs across the sentence boundary in terms of percentage (25.4 %) in comparison to Danish and Italian (28.3 %

and 26.2 % respectively). Thus, this survey does not reveal why English apparently uses fewer explicit discourse cues than Danish and Italian.

Another way of investigating the issue further is by consulting the linguistic items used as discourse cues in the three languages. Table 6.7 shows the distribution between the syntactic classes considered in this analysis. I have chosen to conflate the three last classes (conjunct adverb, prepositional phrases and phrases which take sentential complements) into a single group labelled 'Others'. This has been done because an actual distinction between conjunct adverbs and prepositional phrases is often hard to make and because the group of phrases which take sentential complements was rather small. Finally, these three classes are not always treated in the literature as discourse cues (e.g. Forbes et al., 2002).

	Danish	**English**	**Italian**
Coordinators	46.0 %	37.7 %	33.3 %
Subordinators	27.7 %	50.8 %	28.9 %
Others	26.3 %	11.5 %	37.8 %

Table 6.7: Distribution of linguistic items used as discourse cues

As seen in this table, there are no overall similarities across the three languages: in Danish, there is a tendency to prefer coordinators (46 %) as discourse cues; in English, subordinators represent the majority of discourse cues (50.8 %); and in Italian, the distribution of the three groupings of discourse cues is more equally divided (33.3 % – 28.9 % – 37.8 %). I have earlier mentioned that in some cases the presence of a discourse cue between intrasententially related EDUs determines the syntactic classes of two EDUs in question. In the two examples below, example 136) contains two coordinated EDUs due to *and*, whereas example 137) contains a subordinate and a matrix EDU due to *because*.

136)   Mary is ill and will not come tonight.

137)   Because Mary is ill, she will not come tonight.

This means that the numbers of coordinators, subordinators and other discourse cues partly reveal the distribution of syntactically coordinate and subordinate EDUs. This is most clearly indicated by the numbers of coordinating discourse cues, which as compared with the distribution of coordinate versus subordinate EDUs in Table 4.5 disclose a tendency towards a higher frequency in Danish (32 %) of coordinate EDUs than in English (23 %) and Italian (26

%). However, as reported above, discourse cues are in most cases not signalled between EDUs, and the numbers in Table 6.7 are therefore not entirely comparable with those of Table 4.5. For a better understanding in which type of EDUs the explicitly signalled discourse cues occur, we can investigate the distribution of these in relation to the different levels of the deverbalisation scale (see Figure 4.1). This distribution is reproduced in Table 6.8, from which we can observe, firstly, how discourse cues occur at the different levels within one language, and secondly how on the individual levels cues vary across the three languages. The numbers represent the occurrences of explicitly marked discourse cues at the given level, which in the case of level (a) means that 30.5 % of EDUs textualised as independent sentences in Danish contain an explicit discourse cue.

	Danish	English	Italian
a. independent sentence	30.5 %	18.5 %	41.0 %
b. main or matrix clause that is part of a sentence	26.2 %	18.1 %	40.7 %
c. coordinate main or matrix clause	89.3 %	81.4 %	71.9 %
d. subordinate finite adverbial clause	98.0 %	95.0 %	95.2 %
e. relative clause	6.2 %	2.3 %	10.1 %
f. infinitival clause	100.0 %	100.0 %	100.0 %
g. gerund, present or past participle construction	0.0 %	45.7 %	22.7 %
h. present, past participial or adjectival modifier	0.0 %	6.8 %	15.0 %
i. nominalisation	100.0 %	100.0 %	100.0 %
j. verbless construction	66.7 %	60.0 %	71.9 %

Table 6.8: Distribution of explicitly marked discourse cues in relation to the deverbalisation scale

In particular, the differences at levels (a) and (b) between English versus Danish and Italian are worth noting. The numbers of these two levels represent the discourse cues that express a rhetorical relation between two or more intersententially related EDUs and are thus related to the numbers presented in Table 6.6, in which no notable differences between the intersentential and intrasentential linkage of discourse cues were observed. However, by contrasting the number of independent sentences (a) and main or matrix clauses (b) that contain explicit

discourse cues with those that do not, we can see that the English texts to a much lower extent (18.5 % and 18.1 %) than Danish (30.5 % and 26.2 %) and especially Italian (41 % and 40.7 %) marks these two types of textualisations with discourse cues. In the examples below, we can see how the reader of the two English text segments has to infer the appropriate rhetorical relations on his own, whereas the reader of the Italian example is guided through the presence of a discourse cue in the second EDU. In both examples, the satellite EDUs have been underlined. In 138), the satellite expresses an Antithesis, and in 139), the satellite expresses a Concession of the situation presented in the nucleus. The English example could be paraphrased with *Although they already believe*, whereas the Italian example already contains an explicit discourse cue (*però/however*), here highlighted in bold-faced type.

138)   [<u>They already believe they are protected by the EU</u>.] [In some cases, alas, that is not so.] <ep-97-11-05.txt:183>

139)   [Credo sia giusto che la proposta sulle competenze dell'Unione non interferisca nell'ordinamento degli Stati nazionali.] [<u>Non possiamo, **però**, ignorare che alcune costituzioni nazionali attribuiscono importanti competenze legislative alle regioni,</u>]... <ep-03-01-13.txt:73>

*[I believe it is right for the proposal on the Union's powers not to interfere with the systems of the individual States.] [<u>We cannot, **however**, disregard the fact that some national Constitutions confer major legislative powers on the regions,</u>] ...*

In this case, English seems more writer-oriented than Danish and Italian, as it is often up to the reader to interpret the rhetorical relation between two intersententially related EDUs. Levels (a) and (b) are of particular interest here, because it is at these two levels that the use of discourse cues is more optional as compared with lower levels such as (c), (d) and (f). The lower percentage of explicit discourse cues at level (c) in the Italian texts (71.9 %) is mainly due to the higher number of coordinate EDUs separated by colons or semicolons, between which a discourse cue is often not explicated. The two following examples from the Italian part of the corpus show an example in 140) with semicolon and an explicit discourse cue (*però/yet*) and an example in 141) with a colon and no discourse cue.

140)   [Qui è stata evocata la collegialità della Commissione;] [<u>la sensazione è **però** che si lavori un pò troppo per compartimenti stagni.</u>] <ep-97-04-07.txt:81>

*[Mention has been made here of the collegiate nature of the Commission;] [**yet** one does get the feeling that it works rather too much in separate compartments.]*

141)   [Io credo che sul serio dovremmo, noi tutti, considerare gli anziani un patrimonio dell'umanità:] [la loro saggezza, la loro conoscenza, la storia [di cui sono viventi interpreti,] in realtà non sono fino in fondo valorizzate e utilizzate.] <ep-02-04-11.txt:43>

[I feel that we should all genuinely see elderly people as our human heritage:] [their wisdom, their knowledge and the history [of which they are the living exponents] are not fully valued or exploited.]

The next notable difference is found in the Italian relative clauses (level e) that exhibit a higher percentage of explicit discourse cues (10.1 %). Example 142) shows an instance of this. It seems difficult to pinpoint the exact reason for this stronger tendency in Italian to include discourse cues in relative clauses, but in Gylling (2012a), I argue that there may be a connection to the Latin distinction between typical (my translation of the Italian term *proprie*) and atypical (*improprie*) relative clauses (see Ogina, 2007, p. 271; Sensini & Roncoroni, 1997, p. 513). In atypical relative clauses in Latin, the mood was subjunctive, entailing that these relative clauses could express the same kind of rhetorical relations as adverbial clauses, no matter whether these were restrictive or appositive in nature (cf. also Blatt, 1946, p. 200ff; Ernout & Thomas, 2002, p. 336; Gast & Schäfer, 2012; Loock, 2010; Sausy, 1977, p. 394f). Latin and Italian are two closely related languages, but since Italian relative clauses with verb forms in the subjunctive mood are typically restricted to indicating the writer's intervention (Prandi, 2010), it could be argued that Italian uses discourse cues instead of the subjunctive to indicate that the relative clause expresses a semantic or pragmatic relation other than Elaboration of its matrix clause. Example 142) shows an instance of an Italian relative clause with the discourse cue *pur(e)* (however) signalling a concessive relation.

142)   [Tutto questo non deve certo farci dimenticare le restrizioni delle libertà civili] [che si registrano in quel paese,] [ma ritengo che non sia sufficiente ed efficace un atteggiamento di pura e semplice condanna,] [che **pur** non deve mancare.] <ep-00-06-14.txt:176>

*[Nevertheless, we clearly cannot disregard the suppression of civil liberties] [which takes place in that country,] [but I feel that it would not be sufficient or effective for us to adopt a purely condemnatory position,] [which we, **however**, must show.]*

The last two notable differences are found between the English and Italian textualisations at levels (g) and (h). As for the former, English uses discourse cues in present or past participle clauses (45.7 %) approximately twice as often as does Italian (22.7 %). With regard to the latter, Italian proves to include more explicit discourse cues in modifiers (15 %) than does English (6.8 %). In both cases, the Danish texts do not contain any explicit discourse cues. The examples below show typical textualisations with and without discourse cues in English and Italian. In 143) and 144), we can see two English participial clauses starting with a discourse cue (*when* and *as*), and in 145), we can observe an Italian gerund (*sottraendo*/subtracting) that has no explicit discourse cue attached.

143)   [I do not always agree with the Council,] [but I agree with its view that approximation of the criminal law of the Member States could be necessary for certain specific types of offences,] [but [**when** approximating criminal law] the specificity of the national systems has to be taken into account.] <ep-01-06-11.txt:59>

144)   [If BNFL has problems in discharging radioactive products such as technetium 99,] [it should store such waste materials on land in Britain] [**as** opposed to dumping them in the Irish Sea.] <ep-99-09-13.txt:59>

145)   [Per questa ragione, sostengo la proposta di assumere per il 2004 gli obiettivi di Lisbona e di Göteborg, ossia la programmazione di investimenti pubblici nei settori della ricerca, della formazione lungo tutto l'arco della vita e del risanamento ambientale, nonché la costruzione di una rete europea integrata nei trasporti e nelle telecomunicazioni, come obiettivo addizionale del Patto di stabilità e di crescita,] [**sottraendo** l'ammontare di questi investimenti dal calcolo del deficit dei bilanci dei governi nazionali.] <ep-02-10-21.txt:49>

*[I therefore support the proposal to include the Lisbon and Gothenburg objectives, in other words the programming of public investment in the areas of research, lifelong training and environmental improvement and the creation of integrated European transport and telecommunications networks, as an additional objective of the Stability and Growth Pact,] [**subtracting** the sum of these investments from the total budgetary deficit of Member States' governments.]*

6.2.5 Signalling rhetorical relations with discourse cues

Although an account of the relationship between the discourse cues found in the corpus and their corresponding rhetorical relations would seem appropriate in the context of this chapter, I have chosen not to compare these two parameters with each other. This is because a number of scholars have concluded that there are almost no safe correlations between relations and cues (Knott, 1996; Schauer & Hahn, 2001; Taboada & Mann, 2006, p. 439; Taboada, 2004, p. 150; Van der Vliet & Redeker, forthcom.). While some cues such as *if* (along with its Danish and Italian counterparts *hvis/se*) exclusively signal one rhetorical relation, namely Condition, in the examples found in my corpus, other cues such as *and* (along with *og/e)* are used to signal several relations: Elaboration, Conjunction, List, Purpose, Sequence, (Non)-Volitional Result. Consequently, I have decided not to include a table of the correlations between discourse cues and rhetorical relations but to account for the signalling of the individual rhetorical relations instead. Furthermore, a number of cues were only observed once or twice in the corpus, and it would be erroneous to conclude anything based on these.

In Table 6.9, the distribution of the explicitly signalled discourse cues in relation to the rhetorical relations of RST is shown for all three languages. Again, the percentages represent the number of occurrences of EDUs between which a discourse cue is used to express a certain rhetorical relation. In the case of Circumstance in the Danish texts, we can observe that 94.5 % of the EDUs in which a Circumstance relation is expressed have been marked by some kind of discourse cue. The table also shows the distribution of explicitly signalled discourse cues in relation to linkage. For practical reasons, only the percentages of explicitly signalled intersentential discourse cues have been indicated, that is, explicitly marked discourse cues that connect EDUs in different sentences. These numbers are shown in italics and should be read as follows: in the case of Elaboration in the English texts, we can see that 13.3 % (out of the total 4.5 %) of the EDUs in which an Elaboration relation is explicitly signalled are found between intersententially related EDUs. The remaining 86.7 % are, thus, found between intrasententially related EDUs.

		Danish		English		Italian	
		Total	*Inter-*	**Total**	*Inter-*	**Total**	*Inter-*
SEMANTIC RELATIONS	Circumstance	94.5 %	*0.0 %*	63.4 %	*0.0 %*	66.7 %	*0.0 %*
	Condition	83.3 %	*0.0 %*	96.0 %	*0.0 %*	100.0 %	*0.0 %*
	Elaboration	13.5 %	*47.7 %*	4.5 %	*13.3 %*	13.0 %	*28.6 %*
	Evaluation	0.0 %	*0.0 %*	0.0 %	*0.0 %*	0.0 %	*0.0 %*
	Interpretation	0.0 %	*0.0 %*	0.0 %	*0.0 %*	0.0 %	*0.0 %*
	Means	85.7 %	*0.0 %*	50.0 %	*0.0%*	60.7 %	*0.0 %*
	Non-Volitional Cause	56.3 %	*83.3 %*	22.0 %	*81.8 %*	44.2 %	*55.9 %*
	Non-Volitional Result	52.4 %	*54.5 %*	32.0 %	*75.0 %*	48.1 %	*38.5 %*
	Otherwise	-	-	-	-	100.0 %	*0.0 %*
	Purpose	75.0 %	*0.0 %*	80.5 %	*0.0 %*	85.0 %	*0.0 %*
	Solutionhood	0.0 %	*0.0 %*	0.0 %	*0.0 %*	0.0 %	*0.0 %*
	Unconditional	100.0 %	*0.0 %*	66.7 %	*0.0 %*	100.0 %	*0.0 %*
	Unless	100.0 %	*0.0 %*	42.9 %	*0.0 %*	100.0 %	*0.0 %*
	Volitional Cause	67.2 %	*73.2 %*	37.1 %	*53.8 %*	61.7 %	*54.0 %*
	Volitional Result	77.8 %	*66.7 %*	50.0 %	*66.7 %*	48.0 %	*41.7 %*
PRAGMATIC RELATIONS	Antithesis	29.4 %	*40.0 %*	28.6 %	*33.3 %*	20.0 %	*100.0 %*
	Background	7.1 %	*80.0 %*	6.5 %	*80.0 %*	7.0 %	*80.0 %*
	Concession	50.7 %	*72.2 %*	55.4 %	*77.4 %*	58.5 %	*41.7 %*
	Enablement	0.0 %	*0.0 %*	0.0 %	*0.0 %*	0.0 %	*0.0 %*
	Evidence	41.7 %	*44.0 %*	16.3 %	*42.9 %*	48.9 %	*45.4 %*
	Justify	52.1 %	*28.0 %*	44.4 %	*15.0 %*	57.1 %	*30.0 %*
	Motivation	50.0 %	*50.0 %*	0.0 %	*0.0 %*	25.0 %	*50.0 %*
	Preparation	33.3 %	*100.0 %*	0.0 %	*0.0 %*	33.3 %	*100.0 %*
	Restatement	36.4 %	*0.0 %*	50.0 %	*0.0 %*	80.0 %	*50.0 %*
	Summary	15.4 %	*100.0 %*	16.7 %	*100.0 %*	15.8 %	*100.0 %*
MULTINUCLEAR	Disjunction	100.0 %	*0.0 %*	100.0 %	*0.0 %*	100.0 %	*0.0 %*
	Conjunction	70.6 %	*23.2 %*	63.3 %	*15.9 %*	72.4 %	*16.0 %*
	Contrast	74.1 %	*10.0 %*	66.7 %	*14.3 %*	83.9 %	*7.7 %*
	Joint	0.0 %	*0.0 %*	0.0 %	*0.0 %*	0.0 %	*0.0 %*
	List	40.5 %	*40.0 %*	53.8 %	*28.6 %*	63.6 %	*7.1 %*
	Multin. Restatement	-	-	50.0 %	*100.0 %*	0.0 %	*0.0 %*
	Sequence	64.5 %	*35.0 %*	64.0 %	*50.0 %*	63.3 %	*57.9 %*

Table 6.9: Distribution of discourse cues in relation to rhetorical relations and linkage

Table 6.9 reveals a number of interesting aspects of how text structure, discourse structure and information structure interact with each other. The first aspect which surfaces from the numbers is that some relations are never signalled by discourse cues across all three languages. This holds for Evaluation, Interpretation, Solutionhood, Enablement and Joint. Secondly, it is noteworthy that some relations are rarely marked: Background, Elaboration and Summary. Both of these findings are very much in line with the findings of previous studies (Knott & Dale, 1994; Taboada, 2004b). Thirdly, the numbers in the table show remarkable similarities across the three languages regarding the distributions of explicitly marked discourse cues. Take, for instance, Antithesis with around 25 % explicitly signalled occurrences in all three languages, Concession with around 55 %, and Conjunction with around 70 %. These together with other relations which occur less frequently in the texts of the corpus such as Disjunction could indicate that there is a more universal pattern of signalling rhetorical relations across the three languages.

Table 6.9 also reveals a number of cross-linguistic differences. And this is where the three linguistic structures interact. As regards the relations of Circumstance and Means, we can see that English (63.4 % and 50 % respectively) and Italian (66.7 % and 60.7 %) explicate these less frequently than does Danish (94.5 % and 85.7 %). This is mainly because English and Italian use different textualisations of Circumstance relations as compared with Danish, in particular non-finite realisations, which do not necessarily entail an explicit discourse cue; cf. also Table 5.6, Table 5.7 and Table 5.8.

Some of the same results are found in the four causal relations of Non-Volitional Cause and Result together with Volitional Cause and Result. Here, the cross-linguistic differences are not confined to Danish versus English and Italian but also apply to English versus Danish and Italian in the case of Non-Volitional Cause, Non-Volitional Result and Volitional Cause. Here, Danish is more similar to Italian than to English. As such, the picture is rather complex and cannot only be explained by the differences in textualisations, but needs to be clarified through a more detailed survey of the rhetorical hierarchy of the nucleus-satellite relationship in Cause-Result relations; see the discussion in Section 5.5.2. However, a survey of this kind requires more data than available in the present corpus.

Moreover, the differences between the explication of the Evidence relations call for attention. In the Danish and Italian texts, the relation is signalled approximately half of the times (41.7 % and 48.9 % respectively), whereas the English texts contain only explicit discourse cues in Evidence relations in 16.3 % of the occurrences. In Table 6.4, we witnessed that the Danish and

Italian discourse particles *jo* and *infatti* were relatively frequent, and that they effectively do not have any linguistic counterpart in English displaying the same frequency. This could be seen as a plausible explanation for why Evidence is less frequently signalled in English than in Danish and Italian[10].

Finally, regarding the distribution of discourse cues in relation to rhetorical relations and linkage, it can be observed that Danish and English exhibit relatively similar patterns compared to Italian. In line with the previous observations in this chapter and in Chapters 4 and 5, the Italian texts display lower frequencies of explicitly marked discourse cues between sentences in many rhetorical relations as compared with the Danish and English texts.

6.3 Summing up

In this last chapter of the three analyses of linguistic structures in parliamentary discourse, the concept of information structure was addressed. The study of information structure, which was here approached from a somewhat different angle than the more frequently employed thematic approach (Lambrecht, 1994), included the examination of two parameters. Both parameters focus on the way that information is packaged in the Danish, English and Italian texts, again departing from the notion of Elementary Discourse Units (EDUs). The first parameter was linkage, where I investigated how the EDUs of the texts were linked intersententially and intrasententially. The second parameter was signal, where the implicit and explicit marking of rhetorical relations between EDUs was studied through the notion of discourse cues.

In the analysis of the first parameter, linkage, I started the analysis by surveying the basic numbers of the Europarl texts, discovering that the Italian mean sentence lengths were strikingly higher than the Danish and English counterparts. But, since cross-linguistic comparisons of sentence lengths, in general, are subject to a high degree of uncertainty concerning the nature of the observed differences, I carried out a number of tests to establish what might have caused these differences. The first test in which the texts analysed in this thesis were compared with their official translations in the two other languages revealed that the differences in sentence length could not entirely be caused by typological differences between Germanic and Romance languages. The second test which examined the distribution of the number of EDUs per sentence also revealed interesting differences between the three languages: in particular, the sentences of the Danish and English texts usually contained one or two EDUs per sentence, whereas the

---

[10] Another explanation, pointed out by Gisela Redeker (personal communication), could be that English uses *do*-constructions to signal Evidence relations, but as verbs are not regarded as discourse cues in this thesis, these constructions have not been annotated.

Italian sentences exhibited an almost equal distribution of sentences with one, two, three and four EDUs. Furthermore, the Italian sentences included more EDUs per sentence (up to 13) than Danish and English, where sentences with more than seven EDUs were extremely rare. Lastly, I discussed how these differences affected the text structure and discourse structure of the three languages, and the way in which emphasis is put on different parts of the argumentation. The results of the analysis of linkage reveal a far more varied linkage pattern in Italian than in Danish and English and reconfirm many of the findings of the analysis of text structure presented in Chapter 4.

In the analysis of the second parameter, signal, the results disclosed different aspects. First of all, English showed a less frequent use of discourse cues as compared with Danish and Italian. One of the reasons for this was the lower frequency of discourse cues in textualisations of levels (a) and (b) of the deverbalisation scale. Secondly, another analysis showed that the three languages tend to prefer different types of discourse cues: Danish prefers coordinators and English mainly subordinators, whereas Italian distributes the different discourse cue items more equally. In a more detailed analysis of the most frequent discourse cues, the three languages exhibited the same preferences for *and* and *but* together with their Danish and Italian counterparts. Lower on the list, we found that English had many occurrences of the textual organisational cue *firstly* and its cognate expressions, whereas both Danish and Italian had a marked number of occurrences of the pragmatic discourse particles *jo* and *infatti*. Lastly, we saw that across all three languages some rhetorical relations were never signalled, while others were rarely signalled by discourse cues.

In conclusion, this analysis of information structure revealed not only interesting cross-linguistic differences between Danish, English and Italian, but it also demonstrated how the three linguistic structures under investigation interact with each other. We have seen how the way EDUs are linked to each other affects the textualisation of these, which, again, affects the way in which the rhetorical relations between the EDUs are signalled by discourse cues. This chapter concludes the three chapters devoted to the analysis of linguistic structures in parliamentary discourse in Danish, English and Italian. These three chapters have established the basis for the following chapter on how these structural differences and similarities can be related to contextual factors viewed from an intercultural rhetoric perspective.

7   Linguistic structures in context

In the previous three chapters, I have examined a number of linguistic phenomena of parliamentary discourse in Danish, English and Italian from three different perspectives: a text structural perspective in Chapter 4, a discourse structural perspective in Chapter 5, and an information structural perspective in Chapter 6. In these three chapters, I have investigated how writers across three different languages employ various linguistic means when creating persuasive discourses, without taking into consideration contextual factors such as linguistic and rhetorical traditions.

This chapter is devoted to a discussion of the similarities and differences observed in linguistic structures between Danish, English and Italian by examining the various cross-linguistic patterns from an intercultural rhetoric perspective, and thus including contextual factors. The chapter comprises the last two steps in the 12-step analysis proposed by Moreno (Connor & Moreno, 2005; Moreno, 2008), as described in Section 2.3. Unlike the previous three chapters, this chapter will start with three examples of parliamentary discourse from each of the three languages in my corpus. The three samples have been chosen so as to represent three different types of communicative patterns across the three languages – patterns which I argue can be correlated with different linguistic constraints and rhetorical preferences. Next, the three communicative patterns are placed in a general discussion of linguistic and rhetorical typologies accompanied by selected references to the contrastive literature on linking language, cognition and rhetoric. Finally, each language is categorised within these typologies.

7.1   Introductory remarks on the relationship between language and cognition

Before presenting the three examples and the underlying typologies, however, a few comments shall be made on the relationship between language and cognition. The analyses and results presented in the previous three chapters indicate that the three languages exhibit a marked preference for the linguistic means that writers tend to favour for linking together the units of a text. As mentioned in Chapter 2, this study, and its underlying ideas and methodology, are closely associated with the interdisciplinary field of intercultural rhetoric, departing from the assumption that 'writing is culturally influenced in interesting and complex ways' (Connor, 2002, p. 495). As such, this thesis can be considered a part of the increasing bulk of research conducted within the cross-disciplinary fields of language, cognition and culture in their widest definitions.

Admittedly, this may not have been the easiest path to follow, as the existence of 'a correlation between language, cognition and mentality is still an open question, although it has been discussed for more than two hundred years' (Durst-Andersen, 2010, p. 29). In fact, Korzen (2005a, p. 55) traces the discussion even further back in history and discusses some interesting statements by the 18th- and 19th-century philosophers Vico and Humboldt. But correlating language and cognition is a essential path to follow in a study that sets out to answer to what extent discourse patterns can be associated with contextual factors.

In contrast to various studies of cultural aspects in international business communication (e.g. Beamer, 2000; Hofstede, 1984; Lewis, 2006), the approach in intercultural rhetoric is to move from examining linguistic features of particular texts that are subject to specific rhetorical situations to contextual factors. As illustrated in Figure 2.1 illustrating the various linguistic structures and cognitive dimensions, Chapters 4 to 6 moved from the lower levels of lexicon and morphology to the middle levels of syntax, text, discourse, and information structure, now arriving at the upper levels of context. By following this bottom-up approach, the point of departure is constituted by naturally occurring linguistic data found in the corpus. The converse approach is to adopt a top-down approach, which would take contextual factors as its point of departure. I have chosen the former approach because of the obvious biases related to the latter. The correlations I draw between linguistic structures and contextual aspects should be perceived as genre-specific rather than universal, and as descriptive rather than prescriptive. As in any corpus-based study, a different data set may result in a different set of findings. This being said, most of my findings are supported by a number of earlier findings and observations. I refer to these in the following where applicable.

7.1.1 Language, cognition and linguistic structures

The most salient point in a study that ties up language with cognition is whether language *influences* thought or whether language *controls* thought, or whether there is no correlation between language and thought at all. In intercultural rhetoric, the majority of studies adhere to the assumption of a correlation based on influence (Moder, 2004, p. 10), also referred to as the 'weak' version of the Sapir-Whorf hypothesis of linguistic relativity; the correlation based on control is referred to as the 'strong' version (Gumperz & Levinson, 1996; Lucy, 1992; Sapir, 1964; Whorf, 1954). In this thesis, the weak version is followed. Much inspiration has also been drawn from subsequent revisions of this version, in particular Slobin's (1996a) revision of correlating differences in linguistic structures with differences in rhetorical patterns. Slobin does

not correlate language with thought *per se*, but talks of fitting one's thoughts into linguistic frames. This idea is coined in the notion of 'thinking for speaking' defined as (p. 76):

> a special form of thought mobilised for communication. The activity of thinking takes on a particular quality when it is employed in the activity of speaking. In the evanescent time frame of constructing utterances in discourse one fits one's thoughts into available linguistic frames. 'Thinking for speaking' involves picking those characteristics of objects and events that (a) fit some conceptualisation of the event and (b) are readily encodable in language.

These 'linguistic frames' are also taken to comprise the notion of linguistic structures as applied in this thesis. They are, as discussed in Chapter 2, based on a combination of grammatical constraints (Sausurre's *langue*) and stylistic preferences (*parole*) which continuously influence and are influenced by contextual factors such as the communication situation, the writer's cultural background and linguistic traditions. Below, this idea is illustrated with the three cognitive dimensions from Figure 2.1 in Chapter 2, showing how language can be viewed as a window on human cognition.
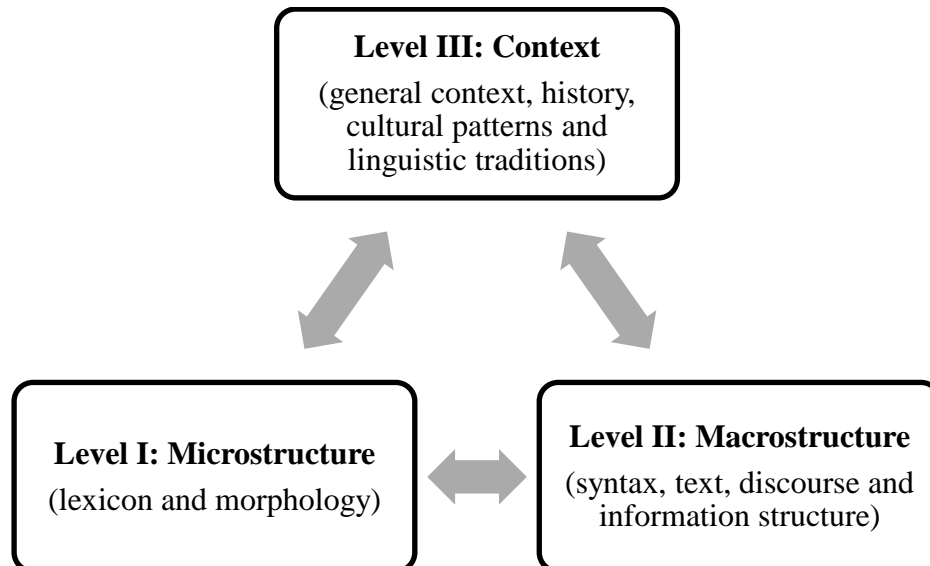
Figure 7.1: Relations between cognitive dimensions

In the context of this study, we have seen how Level I features influence Level II features. In Chapter 4, it was shown that participial constructions (Level I) are extremely rare in Danish, whereas they are relatively frequent in English and Italian. This results in a more hierarchical

text structure (Level II) in English and Italian and a higher frequency of non-finite realisations of main verbs. However, evidence that Levels I and II are capable of influencing or being influenced by Level III features has still not be discussed in this thesis.

7.2   Three examples

This section deals with three examples from my corpus in each of the languages under consideration. The examples capture the kinds of cross-linguistic differences that the previous three chapters have examined in detail. The first example in 146) is taken from the Italian part of the corpus. The topic of the speech is animal experimentation. The text is an excerpt of the original, which contains three subsequent paragraphs. Below the Italian L1, an English translation is found.

146)   Signor Presidente, ho apprezzato la coerenza con cui l'onorevole Roth-Behrendt ha portato avanti la sua relazione e soprattutto l'esigenza di porre fine a una discussione che ormai dura da molti anni. Condivido anche le preoccupazioni da parte dell'opinione pubblica sulla sperimentazione animale: quando questa è ripetitiva su ingredienti e principi attivi conosciuti da anni e già immessi in commercio, allora il sacrificio di nuovi animali è assolutamente inutile. Ma quando, signor Commissario, come io sostengo nell'emendamento sottoscritto da oltre cinquanta parlamentari, si tratta di nuovi cosmetici contenenti ingredienti nuovi, mai testati sperimentalmente prima al fine di caratterizzarne il profilo tossicologico negli animali da laboratorio, in tali condizioni io sono convinto, da scienziato, che sia obbligatorio procedere ad una prima fase di sperimentazione animale prima dell'uso nell'uomo e prima dell'immissione in commercio. La mancata sperimentazione animale di cosmetici nuovi potrebbe portare - lo dico con la massima chiarezza - all'insorgenza di potenziali effetti tossici, sia nei bambini sia nelle donne sia nell'uomo, potenziale epato- e nefrotossicità, potenziale neurotossicità ed effetti anche cancerogeni. Chi è responsabile, poi, caro Commissario? Per tutte queste ragioni mi auguro fortemente che l'Assemblea possa approvare un emendamento, da me presentato con cinquanta parlamentari, che va nella direzione anche di non ostacolare quello che è il progresso della scienza. <ep-01-04-02.txt:42>

Mr President, I appreciate the consistent way in which Mrs Roth-Behrendt has developed her report and, in particular, the need to bring to an end a debate which has lasted for many years now. I also share the public's concerns regarding animal experimentation; if this is a matter of repeat trials on active ingredients and principles which have been known for years and are already on the market, then the sacrifice of

more animals serves absolutely no purpose. However, Commissioner, when, as I point out in an amendment supported by over 50 Members of Parliament, it is a matter of new cosmetics containing new ingredients which have never been tested in the past in order to establish their toxicological profile in laboratory animals, then, as a scientist, I am convinced that it is essential to carry out an initial set of experiments on animals before cosmetics are used by human beings and before they are placed on the market. If new cosmetics are not tested on animals, it could lead – I would like to make this quite clear – to potentially toxic risks for men, women and children alike. There would be a risk of toxic effects on the liver and nervous system, and even cancer. And who would be responsible then, my dear Commissioner? For all these reasons, I strongly urge the House to adopt an amendment tabled by myself and 50 other Members, which calls for barriers not to be placed in way of the progress of science.

This is a typical example of the Italian texts in my corpus. The writer begins by expressing his attitude towards the topic of the session (a report on animal experimentation), which is a widely used opening in all the texts across the three languages. Then, the writer starts his argumentation, which could be represented in the following way using Toulmin's (1958) 'Model of Argument':
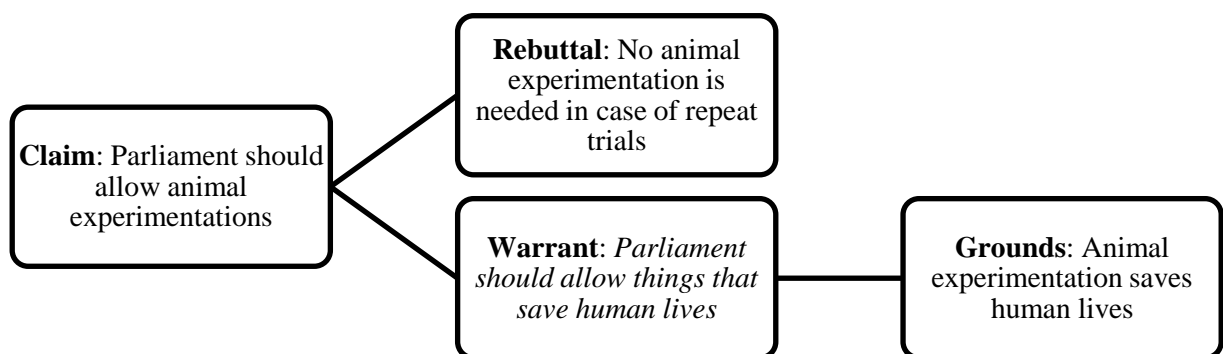
Figure 7.2: Argument model of 'Animal experimentation' text

As we can see, all elements which according to Toulmin are necessary to present a good and realistic argument are present in the text, some more implicit than other, e.g. the warrant (shown in italics). The argument does not pretend to be stronger than it actually is, which can be seen by the inclusion of a rebuttal (or Concession in RST terms) and qualifiers indicating the conditions under which the argument is true (e.g. use of modal verbs and adverbials such as *potenziale/potentially*). In terms of text structure, the text contains a high number of subordinate EDUs, some of which are realised syntactically with non-finite verb forms (mostly as participial modifiers: *contenenti/containing*; *sottoscritto/signed*; *presentato/presented*, but also as

nominalisations: *prima dell'uso/before use*; *primo dell'immissione in commercio/before market release*). Other EDUs are realised as insertions (indicated by means of hyphens in the text) which, although being syntactically coordinated with another EDU, give the text more complexity. In terms of information structure, it can be observed that the sentences are relatively long (35.33 words per sentence; 40.56 if the rest of the text is included) with many EDUs in each sentence. The rhetorical relations between the EDUs are in many cases signalled by a discourse cue (*quando/if*; *ma/but*; *in tali condizioni/under these circumstances*; *per tutte queste ragioni/for all of these reasons*) underlining the discourse structure of the argument presented. In several respects, then, this way of structuring arguments corresponds with the prototypical formal approach to logical argumentation (see Gross & Dearin, 2003, p. 43ff).

The second example in 147) comes from the English part of the corpus. The speech is taken from the same session as the Italian text above, thus dealing with the same topic of animal experimentation. Again, this text has been chosen because it features many of the aspects found in the English texts in the corpus. This text excerpt contains two paragraphs, and the shift between these has been indicated with <P>.

147)   Mr President, I would like to congratulate the rapporteur on a very good report. There was also a very good debate in committee.

<P>

The issue of animal testing has been of concern to this Parliament for a long time and is also reflected in the level of lobbying by constituents who want to keep the ban on marketing cosmetics. It is time for us to ensure that action is taken; that the concerns of our citizens and those who care about animal welfare are properly addressed by a complete ban on sales of cosmetics tested on animals and a testing ban. As other Members have said, there have been enough postponements: from 1998 to 2000, and then to 2002. We have to set a date and keep to it. There can be no justification for making animals suffer in tests when there are proven, effective and safe alternative testing methods and 8,000 cosmetic ingredients already in use. And we are talking about cosmetics, not medical products; we are talking about shampoos and anti-wrinkle creams, and as Mr Davies said earlier, there is certainly no shortage of those in the shops already. <ep-01-04-02.txt:45>

In some respects, this text resembles its Italian counterpart: it contains a clear introductory part and Toulmin's model can be used to represent the structure of the argument. In other respects, however, this text does not resemble the Italian one. A major difference concerns text structure: the EDUs not only consist of subordinate constructions but are also coordinated with each other as coordinated main and matrix clauses (as in the last sentence: *we are talking …*; *we are talking … and … there is certainly no shortage …*) or as coordinated constructions where the second EDU can be said to have some element elided (Ø) (*It is time for us to ensure that action is taken*; Ø *that the concerns of our citizens and those who care about animal welfare are properly addressed*). Moreover, the morphological realisations of the EDUs with finite verb forms entail more syntactic and rhetorical coordination than we saw in the Italian text above, causing a more linear text and discourse structure. Another difference concerns information structure: the sentences are remarkably shorter than the Italian counterpart (24 words per sentence; 26.33 in the whole text), and the signals of rhetorical relations are either semantically vague (several occurrences of *and*) or peripheral to the argumentation (*when*; *as*). Unlike the Italian text, which attempted to persuade the reader by means of logical argumentation, the English text can be said to achieve this by moving the audience: the writer creates a flow by varying the length of the sentences and by highlighting the importance of EDUs with short, audience-involving sentences (*We have to set a date and keep to it*) and repetitions of words and phrases (*we are talking*, repeated twice). These are classical rhetorical figures of speech, also referred to as bi- or tricolons etc. (cf. Jørgensen & Villadsen, 2009, p. 190). Attention is drawn to the words in the text, not to its structure.

The third introductory example in 148) has been taken from the Danish texts. Here, the topic is not animal experimentation but the outcome of an EU meeting in Berlin. The excerpt contains the entire first paragraph and one sentence from the second. Note that this text does not include the usual formal greeting of the president.

148)   Vejen til helvede er brolagt med gode forsætter. Jeg er ikke klar over, om dette udmærkede og lærerige ordsprog også findes på de øvrige ti officielle sprog. Det findes i hvert fald på dansk, og jeg bliver mindet om det næsten hver eneste dag i Den Europæiske Unions institutioner, for disse institutioner - Rådet, Kommissionen og dette Parlament - har aldrig fattet visdommen, den meget simple visdom i den norske håndbog for fjeldvandrere, der lyder således: I tilfælde af uoverensstemmelse mellem kortet og terrænet skal man følge terrænet. Her i dette Parlament og i alle EU's institutioner

inklusive Kommissionen og Rådet følger man konsekvent og med næsten religiøs nidkærhed det én gang vedtagne kort og ignorerer terrænet. Det hævner sig, og jeg skal nævne tre punkter uden at gå i detaljer. Det tillader min taletid jo ikke.

[...]

Det er en markant påmindelse om, at man ikke kan ignorere terrænet, uden at terrænet hævner sig på et tidspunkt. <ep-99-03-10.txt:31>

The road to hell is paved with good intentions. I am not sure whether this excellent and instructive proverb is also found in the ten other official languages. It exists in Danish at least, and I am reminded of it almost every day in the institutions of the European Union, for these institutions – the Council, the Commission and this Parliament – have never grasped the wisdom, the very simple wisdom contained in the Norwegian handbook for mountain walkers, which says: in the event of disagreement between the map and the terrain, you should follow the terrain. In this Parliament and in all the EU's institutions, including the Commission and the Council, people consistently and with an almost religious zeal follow the once adopted map and ignore the terrain. The terrain will have its revenge, and I shall mention three points without going into detail. My speaking time, you know, does not allow that.
[...]
This is a sharp reminder that you cannot ignore the terrain without the terrain taking revenge at some stage.

First of all, the introductory section of this text is quite different from the English and Italian text in 146) and 147). The writer does not address the topic but starts by referring to a Danish proverb *(Vejen til helvede er brolagt med gode forsætter/The road to hell is paved with good intentions)*, which he is not even sure exists in the languages spoken by the non-Danish part of the audience. He then relates the proverb to the way the EU institutions behave and adds yet another saying from a Norwegian mountain walkers' handbook to his claim (*I tilfælde af uoverensstemmelse mellem kortet og terrænet skal man følge terrænet/in the event of disagreement between the map and the terrain, you should follow the terrain*). The intended effect of associating these two sayings with the EU is to highlight that the way EU acts goes against any common sense: if the EU does not change its policies, it will eventually destroy itself. In contrast with the English and Italian way of arguing, the Danish text is more implicit: there is no explicit claim, grounds or warrant, and the writer has not included any rebuttal.

In terms of linguistic structures, we can see that most of the EDUs are coordinated main and matrix clauses, most of which are juxtaposed through coordinators (*Det findes i hvert fald på dansk, og jeg bliver mindet om det næsten hver eneste dag/It exists in Danish at least, and I am*

reminded of it almost every day; Det hævner sig, <u>og</u> jeg skal nævne tre punkter/The terrain will have its revenge, <u>and</u> I shall mention three points). Moreover, most of the main verbs have been realised as finite constructions. In this respect, the text structure resembles the English excerpt above. In terms of discourse structure, the analysis is slightly more complicated due to the information structure of the text. There are hardly any useful discourse cues, except for *for* (*for/because)* in the third sentence and the pragmatic discourse particle *jo* (translated as *you know*) in the last sentence of the first paragraph; see the discussion of this cue in Section 6.2.5. In addition, many of the sentences are very short containing one or two EDUs each; the only exception in this regard is the sentence including the Norwegian saying. The mean sentence length of the seven sentences is 22.57 words.

The Danish text has many of the characteristics of a narrative in that it is up to the reader to infer the different elements of the argumentation. Instead of trying to move or persuade the reader as directly as we saw in the English and Italian texts, the Danish text teaches the reader – most clearly observed in the last sentence, which repeats the moral of the narrative presented in the first paragraph (*man kan ikke ignorere terrænet, uden at terrænet hævner sig på et tidspunkt/you cannot ignore the terrain without that the terrain takes revenge at some stage*).

The three examples that I have just presented differ from each other in several respects. They differ in the way that language is used rhetorically. They differ in the way that EDUs are structured linguistically. And they differ in the way that they try to persuade stylistically. However, the three examples also have some features in common. The Italian and the English texts both have a quite clear discourse structure, and the English and Danish texts both use coordination as the primary means to connect EDUs. In the sections that follow, I shall discuss the interfaces of linguistics and rhetoric in relation to the three languages under consideration. I will do this by drawing on insights from linguistic and rhetorical typology that take into account how writers with different linguistic backgrounds are constrained by the languages they speak and by the rhetorical traditions upheld by the writers of the three languages.

7.3   Linguistic and rhetorical typology

In contrast to linguistic typology, the literature on rhetorical typology is very scarce. It has therefore been necessary to draw on insights from linguistic typology to describe a rhetorical typology for Danish, English and Italian. But as we shall see, the two approaches to distinguish between different language types can easily be combined. In fact, they complement each other in a number of ways, cf. Johnstone (1986, p. 182):

Discourse modelled on logic demands complex subordination, and the complex syntax of subordination characteristic of European languages calls forth logic.

Rhetorical typology has traditionally investigated the effects of lexicalisation typologies on the patterns of sentence and discourse content in different languages (e.g. Folman & Sarig, 1990; Nir & Berman, 2010). In an early paper on rhetorical typology, Johnstone (1989) distinguishes between 'persuasive strategies' and 'persuasive style', which in some respects resembles Saussure's (1959) distinction between *langue* (the language system) and *parole* (the use of this system). Persuasive strategies are defined as 'the various means of persuasion available to any speaker', and persuasive style is intended to account for 'a speaker's general tendency, resulting in part from cultural and historical factors, to adopt one particular persuasive strategy in any situation' (Johnstone, 1989, p. 142). It can therefore be argued that Johnstone's persuasive strategies overlap with Saussure's *langue* understood as the abstract language system, and persuasive style overlaps with *parole* meaning the concrete instances language use, cf. also Nir & Berman (2010, p. 758):

> Earlier cross-linguistic analyses of extended discourse suggest that variation between languages may depend on language typology not only in the sense of the structural means available to speaker-writers of a given language, but also in terms of the expressive options that they favour when constructing discourse. That is, contrastive rhetoric in the sense used here derives from a complex interplay between the repertoire of linguistic constructions in a given target language, on the one hand, and rhetorical preferences governing how speaker-writers of the language select to deploy and alternate these structures to meet particular discourse functions, on the other.

Various linguistic phenomena have been examined in a number of experiments, many providing evidence for the hypothesis that different languages exhibit different ways of structuring information. Hypothesising that narratives constitute a fundamental and universal reflection of human thought (cf. Berman & Slobin, 1994; Bruner, 1986; van Dijk & Kintsch, 1983), several studies have examined which linguistic means speakers of different languages employ when are asked to retell short movies sequences (e.g. Berman & Slobin, 1994; Chafe, 1980; Skytte et al., 1999). Among these, Berman & Slobin's (1994; see also Slobin, 1996a) study demonstrated how the different patterns used by the speakers that participated in the experiment were related

to Talmy's (1985, 1987) typology of verb- and satellite-framed languages. Talmy's distinction has also been the point of departure in recent studies by Korzen and colleagues (Cresti & Korzen, 2010; Korzen & D'Achille, 2005; Korzen & Lavinio, 2009; Korzen, 2012), who characterise Danish (along with other Scandinavian languages) as an 'endocentric' language, where verbs contain the most important information, and Italian (along with other Romance languages) as an 'exocentric' language, where nouns contain the most important information. In these studies, the primary evidence comes from motion verbs as well as nominal lexicalisation patterns (see also Herslund & Baron, 2003). In an survey of the typical linguistic patterns of exocentric and endocentric languages, Korzen (2005a, p. 63) accounts for how the lower level of lexicon/morphology and the middle level syntax/text/discourse/information structure influence the cognitive patterns in the various languages, see also Lundquist (2010) and Korzen (2005b, 2005c). The characteristics of exocentric and endocentric languages are summarised in Table 7.1.

Typical linguistic patterns	Exocentric languages (Romance)	Endocentric languages (Scandinavian)
Level I (lexicon and morphology)	abstract verbs specific verbs non-finite verb realisations	specific verbs abstract nouns finite verb realisations
Level II (syntax)	subordination	coordination
(text, discourse, and information structure)	hierarchical	linear
↓ Typical cognitive patterns	↓	↓
Cognitive focus	on the relation between the units	on each single unit
Thought patterns	abstract synthetic	concrete analytical
Typical speech acts	interpretation	narration

Table 7.1: Exocentric and endocentric languages

The alleged differences between exocentric and endocentric languages are very much in line with the results presented in Chapters 4, 5 and 6, at least for Danish and Italian Level II features. But whereas the text types investigated in the studies cited above were narratives (or, more specific, retellings of movie clips), the texts investigated in this thesis are argumentative political speeches. Even though narratives and argumentative texts may differ in some respects, they both have an important rhetorical component (Bruner, 1986; Chatman, 1990, p. 6ff), which also led Slobin (1996b) in his experiments with retellings of the 'Frog stories' to conclude that languages can be divided into different typologies. In particular, he discussed rhetorical typology versus grammatical/linguistic typology. Similar to the assumptions behind Figure 7.1 and Table 7.1, rhetorical typologies are said to be guided by linguistic typologies, which means that languages with different grammatical options prefer different rhetorical styles and strategies. Again, this confirms the various patterns observed across the Danish, English and Italian texts analysed in this thesis. It also echoes the original assumption in intercultural rhetoric found in Kaplan (1966) that different languages have different 'culturo-linguistic systems' that produce different types of discourse or different rhetorically structured texts. The rhetorical organisation of texts may thus reflect aspects of contextual factors as argued within the closely related discipline of ethnolinguistics/cultural linguistics (see Palmer, 1996; Sharifian & Palmer, 2007):

> Cultural conceptualisations also provide analytic tools for explorations of pragmatic aspects of language. First, the use of pragmatic devices, such as pragmatic markers, may be associated with culture specific conceptualisations [...]. When we say the use of a certain linguistic device has a given implied meaning, we are in fact referring to conceptualisations that the speaker/hearer associates with the use of the device in a particular context. It is of course well-known in the area of pragmatics that different cultures may have different pragmatic norms and devices and thus it may be stated that across different cultures, different devices might be associated with similar or overlapping cultural schemas and, in some cases, similar devices may give rise to contrasting cultural schemas. At the discourse level, both the content of discourse and its rhetorical organisation may reflect cultural conceptualisations of experience [...]. (Sharifian, 2011, p. 31)

That 'rhetorical organisation may reflect cultural conceptualisations of experience' was also the point of departure for Johnstone's (1989) rhetorical typology. Her persuasive strategies constitute the range of options the writer selects in deciding on an appropriate tactic or combination of tactics for persuasion in a given situation. These tactics vary according to the rhetorical situation (Bitzer, 1968) and are considered part of the writer's communicative competences. Among these tactics, Johnstone concentrates on three: one that persuades by formal argumentation, called 'quasilogic', a term borrowed from Perelman (Perelman & Olbrechts-Tyteca, 1969; Perelman, 1977, pp. 79–108); another that persuades by moving the audience, called 'presentation'; and a third strategy that persuades by teaching the audience, called 'analogy'. The three strategies and their characteristics are listed in Table 7.2, which is taken from Johnstone (1989, p. 145).

	Quasilogical	Presentational	Analogical
Conceptual correlates	ideas are persuasive	people are persuasive	culture is persuasive
	institutions make decisions	individuals make decisions	history, tradition make decisions
	structure is crucial	words are crucial	aptness is crucial
Culture/geography	canonically Western (though not so typically as we suppose)	Eastern (in older and more religious tradition)	

Table 7.2: Persuasive strategies

Along with persuasive strategies, there are persuasive styles. These styles are represented in Table 7.3, again taken from Johnstone (1989, p. 151). Persuasive styles are the writer's initial, reflexive choice of strategy or strategies that the writer assumes to be most effective or applicable in the given situation. Other things being equal, persuasive styles, Johnstone argues (1989, p. 143), 'are culturally predisposed' (cf. the last row in Table 7.2). Each style, then, relates to a strategy, as represented in Table 7.3.

As can be deduced from the two tables, 'Westerners' are claimed to be most likely to use quasilogical argumentation, whereas 'Easterners' are most likely to use presentational or analogical argumentation. In quasilogical argumentation, the goal is to convince the audience with logically irrefutable arguments, making it difficult not to accept the writer's claim or conclusion. In this style, ideas are persuasive and do not necessary depend on the person who puts them forward or on the way the arguments are presented linguistically. Instead, attention is paid to structure following Aristotle's epistemological principle that *inventio* and *dispositio* (invention and arrangement) are prior to *elocutio* (style). The linguistic features of quasilogical argumentation are high frequencies of subordinate constructions, frequent use of discourse cues and a highly integrated text structure (in fact, Johnstone (1989, p. 146) refers to Chafe's (1982) notion of 'integration', see Section 4.3.3., in this thesis).

	Quasilogic	**Presentation**	**Analogy**
Distinguishing model	model from formal logic; convincing	model from poetry; moving	model from narrative; teaching
Linguistic correlates	use of 'logical connectives': thus, hence, therefore …	'rhetorical deixis': here, now, this …	formulaic language: 'You know what they say'; 'That reminds me'
		visual metaphors: behold, look, see	'the words of the ancestors'; proverbs
	subordination; integration	coordination/parataxis/ parallelism; involvement	chronology; timeless past ('once upon a time'); involvement

Table 7.3: Persuasive styles

In presentational argumentation, the goal is to move the audience by repeating central passages (parallelism) and hereby emphasising the importance of these in the audience's consciousness. This gives the argumentation rhythm, as we know from poetry, and a more expository nature.

Instead of the ideas, it's the people that present them that are persuasive. The linguistic characteristics are high frequencies of coordinate structures, varied sentence lengths and deictic or visual metaphors instead of discourse cues.

The third strategy, analogical argumentation, aims to persuade by teaching the audience. This is done by making reference to historical precedents or analogies using proverbs or anecdotes. In order to understand the argumentation as argumentative, the audience usually has to make lateral, abductive leaps between past events and current issues. In some cases, these leaps are explicitly provided by the writer. In many ways, the linguistic features of analogical persuasion resemble those of presentational persuasion, with coordinate structures and short sentences as used in narratives.

It is essential to underline that each individual writer has access to all three persuasive strategies, and that I have found instances of all strategies in all three languages in the corpus. In Johnstone's paper, the point was illustrated by comparing three texts produced by the very same person. However, I also hope to show, bearing in mind the three examples from my corpus in Section 7.2 and the typological differences between exocentric and endocentric languages in Table 7.1, that certain linguistic, rhetorical and contextual settings entail certain persuasive strategies, turning these strategies into the dominant persuasive styles of writers in these settings. More specifically, I argue that the writers of the Italian texts in my corpus primarily resort to quasilogical argumentation, the writers of the English texts to presentational argumentation, and the writers of the Danish texts to analogical argumentation. [11] In the following three subsections, each concentrating on one of the three languages, I provide further evidence for these claims with additional exemplification and comparisons to previous studies.

7.3.1 Italian formality

The analyses of the three linguistic structures analysed in this thesis indicate that the Italian texts in the corpus have the following characteristics as compared with their Danish and English counterparts:

- Text structure (Chapter 4)

---

[11] There are, of course, other persuasive strategies, some of which Johnstone (1989, n. 4) also cites. But as these deal with nonverbal communication or strategies used by specific ethnic discourse communities, I have not included them here. A different but also interesting approach is found in Glenn, Witmeyer & Stevenson (1977) who propose three other typologies: the factual-inductive (which focuses on facts and inductively moves towards a conclusion), the axiomatic-deductive (which relies on general principles and deduces implications for specific situations), and the affective-intuitive (which is based on the use of emotional or affective messages); cf. also Gudykunst & Kim (2003, p. 299f). I have, however, not included these either, because they are not directly concerned with linguistic structures as employed in this thesis.

- Dependency: relatively high frequency of subordinate EDUs
- Realisation: relatively high frequency of non-finite verb forms
- Discourse structure (Chapter 5)
  - Source of coherence: same overall distributions of semantic and pragmatic rhetorical relations as in the Danish and English texts
  - Hierarchy: slightly more rhetorical coordination between syntactically subordinate EDUs
- Information structure (Chapter 6)
  - Linkage: longer sentences and more EDUs per sentence
  - Signal: relatively high frequencies of discourse cues

Linguistically speaking, these characteristics suggest that the Italian writers have created relatively complex argumentative texts in terms of linguistic structures. The text structure is hierarchical with many subordinate EDUs that are often realised with non-finite verb forms such as gerunds, participial or nominalised constructions. In many cases, subordinate EDUs relate to other subordinate EDUs resulting in a deeper text structure. The discourse structure in a number of respects resembles that of Danish and English, but due to the differences in text and information structure the textualisations of certain rhetorical relations differ. The information structure is more complex in that the Italian texts contain longer sentences with many EDUs, a feature that has consequences for text and discourse structure. As regards the linguistic typology proposed by Korzen and colleagues of exocentric and endocentric languages, see Table 7.1, the Italian texts considered in this study clearly fall within the exocentric typology. Additionally, my observations generally adhere to the previous studies cited in Section 4.3; cf. also Lo Cascio (1991, p. 24) who states that Italian sentences are more complex and contain more subordinate clauses than, for example, English and Dutch.

Rhetorically speaking, the above-mentioned linguistic characteristics suggest that the Italian writers tend to prefer a quasilogical style of persuasion; see Table 7.2 and Table 7.3. Examples of this persuasive style are found in examples 149)-155) from texts in the corpus. The first example in 149) constitutes an entire text.

149)   Signora Presidente, desidero informare il Parlamento di una grave iniziativa che è stata presa la settimana scorsa, in Italia, dal governo regionale della Regione Lazio. Questo ha infatti costituito una commissione per il controllo dei libri di testo di storia

nelle scuole, commissione che deve controllare fino a che punto - come dicono i neofascisti che sono al governo nel Lazio - questi testi sono inficiati di marxismo. Il fatto è che questi testi presentano soltanto la storia d'Italia, della Seconda guerra mondiale, del nazismo e del fascismo assumendo, come ovvio, che non si possono mettere sullo stesso piano i valori per cui si sono battuti i resistenti e gli antifascisti e le idee che hanno ispirato Hitler e Mussolini.

<P>

Questa è un'iniziativa molto grave che tende a restaurare in Italia una censura, per ora soltanto sui libri di testo ma presto forse anche su altro. Siccome i partiti cosiddetti liberali, anche europei, tendono sempre più ad allearsi con forze d'ispirazione xenofoba, razzista, autoritaria, credo che questo sia un grave pericolo non solo per l'Italia, ma segnali qualcosa che concerne tutta l'Europa. Prego i colleghi di firmare un appello che sto facendo circolare, firmato anche da altri intellettuali europei. <ep-00-11-13.txt:31>

Madam President, I would like to inform Parliament of a serious initiative undertaken last week in Italy by the regional government of the Lazio region. The regional government has established a committee to monitor the history textbooks used in schools, a committee which is to ascertain the extent to which these texts are influenced by Marxism - as the neo-fascists who are in power in Lazio say. The fact is that these texts merely present the history of Italy, of World War II, of nazism and fascism, on the assumption that it is clear that the values for which the resistance and the antifascists were fighting and the ideas motivating Hitler and Mussolini cannot be regarded as equal.

<P>

This is an extremely serious initiative towards reintroducing censorship in Italy, limited to textbooks for the moment but which will soon be extended to other areas. Since the allegedly liberal parties, including the European liberal parties, always tend to ally themselves with parties of xenophobic, racist, authoritarian inspiration, I feel that this is not only a serious danger for Italy, but also an indication of something which concerns the whole of Europe. I would ask the Members to sign a petition which I am circulating, which also bears the signatures of other European intellectuals.

The writer begins here by introducing the context in which the argumentation is to be understood. The main purpose is to persuade the other MEPs to sign a petition against the monitoring of textbooks in the Italian region Lazio. The writer builds up the argumentation by using a range of typical quasilogical properties. The prefix 'quasi-' refers to the fact that this type of argumentation is not logical in formal terms, but resembles logical reasoning in many respects (cf. Perelman & Olbrechts-Tyteca, 1969, p. 194) and takes its effectiveness from

formal, demonstrative logic creating logically incontrovertible arguments. In example 149), this can be seen in the writer's use of transitivity. The writer claims that because an equality exists between *a* and *b* and between *b* and *c*, it therefore also exists between *a* and *c*: *If the European Union (a) is against censorship (b), and censorship (b) is introduced in a Member State (c), then the European Union (a) ought to ban that Member State (c)*, cf. the last paragraph in 149). The transitivity, then, remains informal because the equality here between *a* and *c* is not objectively transitive. An example of formal transitivity is numerical superiority. In Italian, quasilogical argumentation is described at length in Mortara Garavelli's *Manuale di retorica* (1988, pp. 90–97) and in Lo Cascio's *La grammatica dell'argomentare* (1991, pp. 323–327).

Another instance of quasilogical style is found in the last sentence of example 149) in which the writer asks the other MEPs to sign the petition. Here, the writer again uses transitivity: *if the petition (a) has been signed by other European intellectuals (b), and European intellectuals (b) only sign documents that are worth supporting (c), then the petition (a) is worth supporting (c).* Or slightly more appealing in a different form: *if European intellectuals (a) have signed the petition (b), and MEPs (c) sign the petition (b), then MEPs (c) ought to be considered European intellectuals (a).* By using the notions transitivity informally, the rhetorical force of the writer's arguments does not arise from their formal validity but from the ways they resemble the structures of formal arguments.

A second example of quasilogical style of persuasion is found in 150). The purpose of this example is to show how Italian writers package various argumentative elements (cf. Toulmin's model in Figure 7.2) within the same sentence. Here, we can identify marked rhetorical relations such as Concession *(dobbiamo però avere il coraggio/we have to be prepared)*, Condition (*se non sono scelte strategiche che/if the strategies selected do not)* and Cause *(poiché abbiamo a disposizione/because we have at our disposal).* Note also the three coordinated relative clauses *(che io sottoscrivo, che hanno sollevato altri colleghi e che quindi dovrà essere anche al centro della conciliazione/which I consider important, which other Members have raised and which should therefore be at the centre of the conciliation process)*, which, although they are coordinated syntactically, are not coordinated rhetorically.

150)   Ora, io vengo da un paese che avrebbe tutto l'interesse, con i suoi 8.000 km di coste, ad avere quanti più porti possibile iscritti nella categoria "porto europeo"; dobbiamo però avere il coraggio di fare delle scelte, che possono essere anche difficili proprio perché, se non sono scelte strategiche che individuano davvero quelle località e

quei porti che hanno queste caratteristiche, rischiamo, dati fra l'altro i pochi finanziamenti - dovremmo discutere semmai su questo tema, poiché abbiamo a disposizione solo 400 milioni di euro - di vanificare gli interventi, rischiamo cioè di non essere efficaci. Certo, rimane aperto un tema - che io sottoscrivo, che hanno sollevato altri colleghi e che quindi dovrà essere anche al centro della conciliazione - cioè il tema delle isole, delle regioni ultraperiferiche: il tema, in altre parole, di quelle situazioni in cui anche dimensioni diverse da quelle che descriviamo non possono pero' penalizzare la realtà portuale di questi vari paesi e di queste aree. <ep-00-10-02.txt:102>

Now, I come from a country which, with its 8,000 km of coastline, should have great interest in having as many ports as possible included in the "European port" category; however, we must be prepared to take some decisions, which could also be difficult decisions for the very reason that, if the strategies selected do not genuinely identify areas and ports that have these characteristics, we will be in danger, given, amongst other things, the lack of funding – we shall have to talk about this in any case because we have only EUR 400 million at our disposal – we will be in danger of undoing any good work, in danger, that is, of being ineffective. One issue remains to be resolved – which I consider important, which other Members have raised and which should therefore be at the centre of the conciliation process – and that is the issue of the islands and the outermost regions: in other words, the issue of those situations in which, even where a port does not meet the volume criteria we have laid down, its activity cannot be allowed to suffer as a result.

The last examples in 151)-155) have been chosen to illustrate the frequent convention in Italian texts to signal conclusions. Whereas the English texts contain a smaller number of simple conclusive signals such as *finally*, *to sum up* etc., the Italian writers mark their conclusions more explicitly and elaborately with discourse cues that contain anaphoric references to the rest of the entire text or parts of the preceding text. Below, I have underlined some of these.

151)   <u>Per tutti questi motivi</u>, caro Commissario, io condivido la sua relazione, la relazione della Commissione, ... <ep-01-04-02.txt:42>

<u>For all these reasons</u>, dear Commissioner, I support your report, the Commission's report, ...

152)   <u>Per i motivi già esposti</u> diciamo "no" all'eliminazione progressiva delle sovvenzioni alle coltivazioni di tabacco in Europa operata al buio, ... <ep-02-03-14.txt:124>

For the reasons given above, we say 'no' to the under-cover gradual elimination of subsidies for tobacco growing in Europe, ...

153)   <u>Per l'insieme di queste riflessioni</u> a me pare corretto riconoscere alle regioni che hanno competenze legislative la possibilità di ricorrere direttamente alla Corte di giustizia ... <ep-03-01-13.txt:73>

For all these reasons, I feel it is right to give the regions with legislative powers the right to approach the Court of Justice directly ...

154)   <u>Per tali motivi</u> voteremo a favore dell'aumento della dotazione, ... <ep-96-10-21.txt:57>

For these reasons, we shall be voting in favour of increased funding, ...

155)   <u>Per queste ragioni</u> voglio ringraziare per il lavoro da loro svolto i parlamentari che hanno partecipato al comitato di conciliazione, ... <ep-96-12-09.txt:29>

For those reasons, I should like to thank those honourable Members who formed part of the Conciliation Committee for their work, ...

Again, these discourse cues 'borrow' the rhetorical strength from formal reasoning to which they have some kind of propositional similarity. Moreover, the linguistic similarity to formal reasoning in the argumentative metalexis of the discourse cues (*motivi, riflessioni, ragioni*; all translated with *reasons*) gives the reader the impression of an argument that has formal validity.

Formal – or formality – is, in other words, the keyword in the Italian texts not only rhetorically speaking but also linguistically. In Section 2.1.2, I discussed the correlations between formality and linguistic structures, arguing that texts written in a high, formal register express an intellectualisation or 'logification' of the content. I mentioned a number of characteristics of formal register which could be found in all of the Italian texts in the corpus. Along the same lines, Cerruti (2009, p. 276f) describes the linguistic features of formal register in Italian as high frequencies of subordinate clauses, non-finite realisations and explicit discourse cues. I argue, then, that the writers of the Italian texts in my corpus have partially employed linguistic means that from an Italian perspective have to be employed in a formalised setting, cf. also Hymes (1972) and Lo Cascio (1991, p. 24): 'Typically, formation of complex utterances reveals a more formal and sophisticated cultural, cognitive and social situation, but is

also related to the linguistic norms of a given linguistic community'. [12] Among these, the parliamentary habit in Italy of reading aloud from a prepared text, see Section 2.4, could also explain the formal, less oral, register in the texts.

Another interesting contextual factor in this regard is to be found in grammars. As noted by Ferrari (forthcoming, sec. 4), a vast majority of the Italian grammars contain a section on text linguistic matters: Ferrari's own text grammar (Ferrari & Zampese, 2000) is, in fact, entitled *Dalla frase al testo* (From the clause to the text), but even more syntactically oriented grammars (e.g. Dardano & Trifone, 1997; Renzi, Salvi, & Cardinaletti, 1995) include sections and chapters on linguistic structures above the level of the sentence; cf. also Kaplan (1988, pp. 291–297) for a discussion of the role of educational systems in language and rhetoric teaching. We know the concept of text grammars from English as well (e.g. Werlich, 1983), but textual features are usually not included in elementary grammars. This particular focus on text linguistic phenomena in Italian grammars, along with the other contextual factors mentioned above, may explain why the Italian writers prefer the quasilogical style of persuasion and its concomitant linguistic structure characteristics.

7.3.2 English involvement

The analyses of the three linguistic structures analysed in this thesis indicate that the English texts in the corpus have the following characteristics as compared with their Danish and Italian counterparts:

- Text structure
    - Dependency: relatively high frequency of subordinate EDUs
    - Realisation: relatively high frequency of non-finite verb forms
- Discourse structure
    - Source of coherence: same overall distributions of semantic and pragmatic rhetorical relations as in the Danish and Italian texts
    - Hierarchy: slightly more rhetorical subordination between syntactically coordinate EDUs
- Information structure
    - Linkage: short sentences and few EDUs per sentence
    - Signal: lower frequencies of discourse cues

---

[12] My translation of *'Un'articolazione complessa rivela in genere una situazione culturale, cognitiva e sociale più alta e più sviluppata, ma è anche legata ai codici di uso vigenti all'interno di una comunità linguistica'*.

In terms of linguistic features, these characteristics indicate that the English writers have created argumentative texts that in some respects resemble the complex Italian texts and in other respects do not. The text structure is hierarchical with many subordinate EDUs often realised as non-finite verb forms. But contrary to Italian, the English text structure is more linear, as the number of EDUs per sentence is relatively low. The discourse structure in some respects resembles the Danish and Italian ones, but due to the differences in text and information structure the textualisations of certain rhetorical relations differ. It has already been mentioned that in terms of information structure the English texts have short sentences, and compared to the Danish and Italian texts the English texts contain lower frequencies of explicitly marked discourse cues. In other words, the English texts exhibit both exocentric and endocentric features confirming the status as a typological 'hybrid', as proposed by Herslund & Baron (2003).

Rhetorically speaking, I argue that these characteristics suggest that the English writers tend to prefer a presentational style of persuasion. Given the number of subordinate EDUs in the English texts, I admit that it may seem odd to classify the persuasive strategy in the English texts as presentational. As we saw earlier, subordination is usually characteristic of texts where a quasilogical strategy has been applied. But if we examine Johnstone's papers on presentational strategies of persuasion in more detail (Johnstone, 1983, 1986, 1987, 1989), it becomes evident that the terms 'coordination', 'parataxis' and 'parallelism', see Table 7.2, are intended in a much broader manner than the way coordination has been used in this thesis so far. In fact, the terms cover not only syntactically and rhetorically coordinate EDUs as in my analysis, but also repetition of noun and verb phrases, restatements, modifications and a number of related rhetorical figures of speech such as alliteration, bicolons and epistrophes. Examples of this presentational persuasive style are found in the following excerpts in 156)-159) from texts in the corpus. In the first set of examples, the relevant linguistic material expressing repetitions has been underlined.

156)   On entrepreneurship, <u>is it not about time</u> that we heard some of the results from the many benchmarking exercises which have been undertaken? <u>Is it not time</u> that we named and shamed those Member States who are not taking the actions necessary to improve entrepreneurship and competitiveness? <ep-02-02-27.txt:83>

157)   What we need to know is whether the Commission is determined to get as good a result as it can. We now have to make the best of a bad job. The ruling has been made, whether we like it or not. What we now have to do is try to salvage something from it.
<P>
When Mrs Kinnock, Mr Thomas and I were in the Windward Islands in May it was quite obvious that the producers there were living in fear - and I mean fear - of the WTO ruling. Now that it has appeared, goodness knows what they feel like. They said to us that they accept that liberalization and globalization are inevitable in the future. But what they need more than anything is time - time to adjust to a very competitive market; in some cases it will take at least five years and in other cases it will take a far longer period until they can compete with dollar-banana producers.
[...]
Parliament wants to help. We want to play our part. We will do so and we will support you; but we need a commitment from the Commission that this is a real attempt to salvage something from this deal. <ep-97-09-16.txt:102>

158)   How can they argue that the Davos package is too generous to President Mandela's South Africa? How can such a ridiculous proposal as that on port and sherry be allowed to scupper the greatest opportunity that we have to offer real and practical solidarity to South Africa? <ep-99-02-24.txt:58>

159)   It is why the Essex Race Equality Council will this year start a new anti-racist project with young people in Thurrock, host a national exhibition on racial diversity and this very week convenes a first meeting of local women's groups on questions facing minority ethnic women. It is why the Essex Returners Unit with Essex Tech has launched an action plan following on from the Year which includes new recruitment procedures for local police, race awareness for local training organizations and new support for local minority ethnic associations. <ep-98-01-28.txt:80>

As can be see from the examples above, coordination, repetition and parallelism are not only used aesthetically as rhetorical figures of speech; they are also used to create a rhythmic flow in the text as in poetry. This flow can partially be substituted for or function as discourse cues as seen in 160). In this example, the writer uses short coordinated main clauses to express an Antithesis between two EDUs, and then repeats this Antithesis a few sentences later. Note also

how emphasis has been put on the last EDU in each Antithesis couplet mainly as a result of the negations (*not*) in the first EDUs. Reference in this regard is also made to Table 5.1, which suggests that English writers prefer Antithesis to Concession as compared with Danish and Italian writers. EDUs are shown between square brackets.

160)   [Linguistic and cultural diversity is not about defining or redefining boundaries;] [it underpins the cultural strength of our enlarging European Union.] [...] [A region is not about its boundaries,] [it is about its cultures and its languages,] [and, even more so, it is about its people.] <ep-03-01-13.txt:64>

Another linguistic correlate in presentational argumentation strategy is the use of visual metaphors and 'rhetorical deixis' (Johnstone (1989, p. 148) cites Lakoff (1974) for the latter term). Examples of visual metaphors are verbs such as *look*, *see*, *behold* and rhetorical deixis covers words such as *here*, *now* and *this,* which create spatial and temporal realms. In the English part of the corpus, these kinds of metaphors occur relatively frequently, and their presence may partly account for the 'lacking' discourse cues. In examples 161)-162), we can observe how visual metaphors and rhetorical deixis help to make the writers' claims present, as if the claims were actually visible to the reader. Again, the relevant linguistic material has been underlined.

161)   The European Union exists partly so that we can work together and give value and share our experiences. Therefore I am delighted that when we look at the fifth framework programme, as we looked at the fourth, we from all those fifteen countries say: 'That is enough. You may be curious but you are not going to work on that subject.' The consensus this morning amongst most groups shows that the Commissioner is quite right to take further the regulatory body that she has in mind and she has seen that there is a wish to have that strengthened. I can assure you that the Committee on Research, Technological Development and Energy will be looking very carefully at the ethical dimensions of the fifth framework programme. <ep-97-03-11.txt:36>

162)   There has already been reference here tonight to the need for revision. Even before we adopt the directive we are talking about revision and I think that is very wise. Some of the concerns that have been expressed here tonight, particularly about dealing with

illegal content and consumer acceptance, need to be being already pursued by the Commission before we adopt this fully as a European directive. <ep-00-05-03.txt:120>

The ways in which visual metaphors and rhetorical deixis are used in the examples above create some of the same effects as the previous examples of coordination, repetition and parallelism in that they make the writers' claims maximally present in the readers' consciousness. Similarly, the linguistic means employed here gives the reader the impression that the text is not only a piece of rhetorical literacy but in fact a personal speech due to its many features from orality.

Just as formality was argued to be the keyword in the Italian texts, orality and involvement are argued to be the key components of the English texts. Unlike the complex and formal Italian texts, the English texts share many features with the modes of oral discourse (Chafe, 1982; Ong, 1982). This goes hand in hand with the epistemological context of presentational persuasion in which *people* are responsible for persuasion, see Table 7.2, and it is their choices of repeated words, phrases and rhythms that move readers to belief.

The similarities with orality also conform to the so-called 'KISS principle' of the plain language movement in English, that is, 'keep it short and simple' (also found in other versions: 'keep it short and sweet', 'keep inadequacies sparse, scribe' or 'keep it simple, stupid') as described in a number of English style guides and grammars (e.g. *How to write clearly/Scrivere chiaro/Skriv klart*, 2011, p. 6). Here, the idea is to make written (and spoken) discourse as simple and well written as possible, which among other things includes using short sentences to achieve clarity, and simple words to achieve more credibility or *ethos*.

Even though presentational persuasion in Johnstone's model is said to be more eastern than western, it is interesting that she associates this style with communities and individuals that in one way or another feel that their values and beliefs are threatened, or with people that have to defend these; examples include 'traditional Islamic theocracies in the Arab world' and the world of the missionary (ibid, p. 151). In the context of this thesis, this is interesting because many of the British MEPs and politicians in general are known to be very sceptical of the EU (Anderson, 2005; Clifford, 2008; Forster, 2002), partially because they fear that the EU weakens the nation state. This could be one of the reasons that presentational persuasion style is the predominant strategy found in the English texts of my corpus. It could perhaps also explain why the English writers appear to be much more 'present' in their texts than the Italian writers, or what I in the title of this subsection have referred to as 'involvement'.

In a last example from the English part of the corpus, this involvement of the English writers is emphasised by the extensive use of *we*, which appears in all sentences and in almost every single EDU throughout the entire text in 163).

163) Mr President, <u>we</u> would like to congratulate the rapporteur for having brought together two sensitive issues which many of <u>us</u> would prefer not to see linked since the question of security often provides a questionable alibi, feeding the prejudices of those who prefer to restrict asylum and develop a policy of deterrence towards those needing to exercise their individual and fundamental rights. <u>We</u> are pleased that <u>we</u> can concur with the rapporteur on his statements on the need to uphold human rights and civil liberties.

<P>

<u>We</u> believe that civil liberties are not secondary to security; <u>we</u> are told that a key value of the European Union is the protection of freedom, justice and civil liberties, and <u>we</u> agree with the last speaker that <u>we</u> need to be vigilant on this. <u>We</u> also fully support the view that the governing principle must be that of international obligations, fairly and humanely applied.

<P>

So for example, <u>we</u> can agree with paragraph 12, that immediate and general detention should be avoided. <u>We</u> also fully support conclusion 5 of the report, which stresses the need for an holistic approach towards asylum and immigration looking at this as a horizontal policy objective and considering all policy areas, particularly those with an external dimension, such as trade, development, environment and agriculture.

<P>

<u>We</u> have long argued that <u>we</u> need to look at the effect of European Union policies in acting as push factors, and after Johannesburg <u>we</u> must be even more aware that what <u>we</u> are aiming for is sustainable development, reducing the causes of seeking asylum. <u>We</u> trust that will form part of the open coordination. <u>We</u> agree with the number of concerns raised about open coordination: in a number of other policy areas, it has almost become a democracy bypass, leaving out the European Parliament and often also national parliaments. But <u>we</u> hope that, through the open coordination method, solutions to difficult issues, such as finding best practice for legal entry for asylum seekers and other immigrants, will be found. <ep-02-09-23.txt:62>

7.3.3 Danish storytelling

The analyses of the three linguistic structures analysed in this thesis indicate that the Danish texts in the corpus have the following characteristics as compared with their English and Italian counterparts:

- Text structure
    - Dependency: relatively high frequency of coordinate EDUs
    - Realisation: relatively high frequency of finite verb forms
- Discourse structure
    - Source of coherence: same overall distributions of semantic and pragmatic rhetorical relations as in the English and Italian texts
    - Hierarchy: slightly more rhetorical subordination between syntactically coordinate EDUs
- Information structure
    - Linkage: short sentences and few EDUs per sentence
    - Signal: relatively high frequencies of discourse cues

In terms of linguistic features, these characteristics indicate that the Danish writers have created argumentative texts that in certain respects differ from the English and, in particular, from the Italian texts. The text structure is linear with a relatively high frequency of coordinate EDUs often realised with finite verb forms. Like the English texts, the sentences are shorter than the Italian, and the number of EDUs per sentence is also lower. The discourse structure in a number of respects resembles the English and Italian ones, but due to the differences in text and information structure the textualisations of certain rhetorical relations differ. As we saw with the Italian texts, the Danish texts contain more discourse cues than the English.

In relation to the linguistic typology proposed by Korzen and colleagues of exocentric and endocentric languages, see Table 7.1, the Danish texts clearly fall within the endocentric typology. Additionally, my observations generally conform to the previous studies cited in Section 4.3, indicating that the Danish texts in my corpus share a number of features with other text types in Danish.

Rhetorically speaking, I argue that these characteristics suggest that the Danish writers tend to prefer an analogical style of persuasion. I do this because I have found a number of typical features of analogical persuasion strategy in the texts. The first feature is the tendency of calling

to mind traditional wisdom as we saw in the Danish example in Section 7.2, or calling to mind more recently acquired wisdom through arguments by analogy. In the two following examples, the writers use arguments of this type, in 164), by drawing on insights from other laws (*I andre love er sammenhængen mellem indre og ydre miljø erkendt/In other legislation, the connection between the internal and external environment is recognised*) and, in 165), by referring to experiences with water supply services from Denmark and Great Britain.

164)   Hr. formand, det er glædeligt, at vi her har et forslag, der fokuserer på organiske opløsningsmidler. Disse flygtige stoffer har en særlig giftig virkning på mennesker og miljø. Derfor skal vi bestræbe os på at begrænse anvendelsen mest muligt. I mange år har opløsningsmidler været under mistanke for at føre til hjerneskader. Det er - som hr. Blokland nævnte - set i udstrakt grad hos malere, der er i daglig kontakt med stofferne. Herfra stammer begrebet »malersyndromet«. Disse erfaringer bør være med i dette direktiv, som vi behandler i dag, fordi folkesundheden ikke bare er et fritidsbegreb, men så sandelig også et emne for arbejdsmiljøet. I andre love er sammenhængen mellem indre og ydre miljø erkendt, og derfor virker det lidt mystisk, at Kommissionen i denne tekst ikke vil anerkende sammenhængen, men det vil kommissionsmedlemmet forhåbentlig redegøre nærmere for. Jeg ser også en klar forbindelse til direktivet om kemiske agenser, som Freddy Blak er ordfører for. <ep-98-01-13.txt:232>

Mr President, I am pleased that we have here a proposal which focuses on organic solvents. These liquid chemicals have a particularly toxic effect on people and the environment. We should therefore strive to limit their use as much as possible. For many years, solvents have been suspected of causing brain damage. As Mr Blokland said, this is seen to a large extent among painters who are in daily contact with these chemicals. That is where the term 'painter syndrome' comes from. This experience should be included in the directive which we are considering today, because public health is not just a leisure concept, but is indeed also a topic for the work environment. In other legislation, the connection between the internal and external environment is recognised, and it is therefore a little mystifying that the Commission does not want to recognise the connection in this text. But hopefully the Commissioner will explain that more fully. I also see a clear connection with the directive on chemical agents for which Freddy Blak is the rapporteur.

165)   I Danmark er vandforsyningen baseret på et solidarisk princip, hvad angår ressourcetilgængelighed og priser. I Danmark er vandforsyningen decentral og primært offentlig og i mindre omfang privat, men altid bruger-/ejerstyret. Det er en meget vigtig faktor for at sikre en fortsat kildebeskyttelse. Den offentlige del af den danske vandforsyning er meget aktiv med hensyn til at kræve så høje beskyttelsesniveauer som muligt og kræve eventuelle forureningskilder stoppet. Den private del - derimod - af den danske vandforsyning er meget tilfreds med de udvandede krav, som EU nu barsler med, og som indebærer, at grænseværdierne for pesticider i vores drikkevand svækkes.

<P>

Erfaringerne fra England og andre steder viser også, at når vandforsyningen bliver gjort til en vare og privatiseres, så forsvinder ressourcebeskyttelsen. Londons private vandforsyning lobbyedes således på lige fod med den europæiske kemiindustri for, at vi blot skulle tillade store mængder sprøjtemidler i drikkevandet, dengang EU forhandlede om drikkevandsdirektivet - så var det nemmere at tjene penge. <ep-03-09-03.txt:262>

In Denmark, the supply of water is based on a principle of solidarity when it comes to prices and the accessibility of resources. In Denmark, the supply of water services is decentralised, primarily public and to a lesser extent private, but it is always user/owner controlled. That is a very important factor for ensuring the continued protection of sources of water. The public dimension of the Danish water supply services is very active in demanding as high levels of protection as possible and in demanding an end to possible sources of pollution. The private dimension is, however, very satisfied with the diluted requirements that the EU is now coming up with and that means a weakening of the limit values for pesticides in our drinking water.

<P>

Experience in Great Britain and elsewhere also shows that, when the water supply is turned into a commodity and privatised, there is no more protection of resources. At the time when the EU was negotiating in connection with the Drinking Water Directive, London's private water suppliers were thus lobbied on the same basis as the European chemical industry with a view to our simply allowing large quantities of pesticides in the drinking water. It was easier to earn money that way.

Contrary to the English texts, where claims and truths were made available to the reader by being stated and restated, the analogical rhetoric of the Danish texts persuades by teaching. The argumentation style is rooted in the belief that analogies are persuasive, and that decisions are best made with reference to historical precedents (Johnstone, 1989, p. 152). These preferences for analogies and precedents in Danish have also been observed in the legal system by Baron (2003b, 2007), who observes a number of differences between the Danish variant and the

French variant. See also Korzen (2009b, p. 207) on the linguistic differences between Danish and Italian legal texts, and McCool (2009, p. 33f) on the linguistic and cultural differences between common law and civil law traditions in general. Interestingly, the Danish writers in my corpus often make reference to precedents from Denmark; we saw this in 165), and the excerpt in 166) constitutes yet another example of this, see the seven underlined references.

166) Det er vigtigere at fastholde pengepolitikken som en hellig ko end at give de arbejdsløse et job. Jeg finder det vigtigere at tage hensyn til familiernes økonomi og sikre, at farmand kommer hjem med en lønseddel i stedet for en fyreseddel, og sådan tænkte den danske regering, der kom til i 1993. Vi havde i 1992 9,2% arbejdsløse, både i Euroland, altså i de 11 lande, og i Danmark. Fra 1992 til 1998 er ledigheden i Danmark faldet til 5,1%. I Euroland er den steget til 10,9%. Vi har i Danmark haft en samlet vækst på 20% mod 10% i Euroland. Hvis den danske regering havde gjort det samme i dag, så ville det have været direkte ulovligt og grundlovsstridigt, fordi underskuddet på finansloven dengang var 3,9%. Nu har Danmark også overskud på finansloven, fordi de, der fik fyresedlen skiftet ud med en lønseddel, også begyndte at betale skat i stedet for at hæve understøttelse. Jeg synes, statsministrene skulle studere det danske eksempel, når de mødes til økonomisnak i Lissabon. <ep-00-03-13.txt:27>

It is more important to stick to the holy cow of monetary policy than to find jobs for the unemployed. I, on the other hand, think it is more important to take account of families' finances and to ensure that dad comes home with a wage packet instead of a redundancy notice, and so too did the Danish Government which came to power in 1993. In 1992 we had 9.2% unemployment, both in Denmark and in the eleven countries of the good old EU. Between 1992 and 1998, unemployment in Denmark fell to 5.1%. In the EU, it increased to 10.9%. In Denmark, we have had combined growth of 20%, compared with 10% in the EU. If the Danish Government had acted in the same way under the present dispensation, it would have been downright illegal and unconstitutional because the budget deficit was 3.9% at that time. Now, Denmark also has an actual budget surplus because those who exchanged their redundancy notices for wage packets also began to pay tax instead of drawing benefit. I believe that the EU's prime ministers ought to study the Danish example when they meet for their tête-à-tête on the economy in Lisbon.

The way the audience is reminded of time-tested values through analogies can be considered an indirect mode of storytelling. In contrast to the English and Italian texts, many Danish texts begin in the same manner as narratives. Instead of referring to previous speakers who have addressed the same topic, or briefly summarising the most important aspects of the given topic,

some Danish writers open their speeches with historical references, again avoiding the formal greeting of the president, see 167)-168).

167)   Vort kendskab til behovet for en konsekvent omlægning af energiforbruget er ikke af nyere dato. Nogenlunde samtidig med 1970'ernes olieprischok erkendte de vestlige industrilande behovet for øget forskning i anvendelse af vedvarende energi. Det skete først og fremmest ud fra et hensyn til forsyningssikkerheden og ud fra erkendelsen af, at vort energibehov ikke i det uendelige ville kunne dækkes via brugen af fossile brændstoffer. Siden da har miljøpolitiske hensyn forståeligt nok spillet en mindst lige så stor rolle i overvejelserne - ikke mindst i lyset af den kendsgerning, at verdens samlede energiforbrug forventes at stige eksplosivt inden for det kommende kvarte århundrede som følge af industrialiseringen af en lang række tredjeverdenslande.

<P>

På denne baggrund er det på høje tid, at Den Europæiske Union gennemfører en handlingsplan til fremme af vedvarende energikilder, hvori der tages særligt hensyn til de enkelte EU-landes særlige forudsætninger for anvendelse af bestemte former for vedvarende energi (f.eks. øget anvendelse af solenergi i Middelhavslandene, udvidet brug af vindenergi i de nordvestlige EU-lande etc.). <ep-96-07-04.txt:52>

We have long been aware of the need for a systematic reorganisation of energy consumption. Around the time of the oil price shock in the 1970s, the western industrialised countries recognised the need to step up research into the use of renewable energy. This was above all out of a concern for security of supplies, stemming from the realisation that our energy needs could not be covered indefinitely by the use of fossil fuels. Since then, environment policy considerations have understandably come to play an equally important part in the debate – not least in view of the fact that the world's overall energy consumption is expected to increase dramatically within the next quarter of a century, as a result of the industrialisation of many third world countries.
<P>
Against this background, it is high time that the European Union implemented an action plan to promote renewable energy sources, taking special account of the particular conditions for using certain forms of renewable energy in the individual Member States: increased use of solar energy in the Mediterranean countries, for example, further use of wind power in the north-western Member States, and so on.

168)   Det er nu fem år siden, Parlamentet påbegyndte forhandlingerne om statutten, og for nogle kolleger er der her tale om en betydningsfuld, politisk proces, der på én gang sikrer Parlamentets magt og legitimitet – Parlamentets ret til at vedtage sin egen statut er

udtryk for et føderalt, europæisk demokrati. For andre, til gengæld, er det en farce af næsten surrealistiske dimensioner. <ep-03-06-02.txt:67>

It is now five years since Parliament embarked upon the negotiations concerning the Statute and, for some of my fellow MEPs, this is a significant political process that, at one and the same time, secures Parliament's power and legitimacy – Parliament's right to adopt its own statute is/being an expression of a federal, European democracy. For others, however, it is a farce of almost surrealistic dimensions.

In correspondence with the characteristics of analogical persuasion style, see Table 7.2, the Danish writers, like the English, are more visible or involved more explicitly in the texts than the Italian writers, who by contrast build up their argumentation with systems and structures. Examples such as analogies create empathy (Gabrielsen & Juul Christiansen, 2010, p. 97), and sometimes the Danish writers even act as characters themselves in their own narratives. This is seen in example 169), where the writer uses himself in arguing that the Nordic Passport Union should continue to exist within the EU, hereby using a *logos* appeal.

169)   Jeg synes, at der på denne nordiske aften, fru formand, også skal lyde en lille hilsen fra Danmark i det kor af stemmer, der har været fremme. <u>Jeg bor selv i nærheden af den dansk-tyske grænse</u>. Det er jo ikke blot grænsen mellem Danmark og Tyskland. Det er også grænsen mellem de øvrige nordiske lande og Tyskland. Den nordiske pasunion, som ifølge min landsmand, Ole Krarup, ikke må hedde således, men bør kaldes et nordisk pasfrit område, er jo ikke blot noget praktisk, det er i virkeligheden også et politikum, som siger noget om den nordiske sammenhæng og også om den nordiske samfølelse. Det synes jeg som dansker, at det er vigtigt at pege på midt i alt det europæiske og midt i disse Schengen-tider. <ep-96-06-19.txt:226>

I think that on this Scandinavian evening, a welcoming voice should be heard from Denmark among the chorus of speeches which have been made. <u>I myself live near the Danish-German border</u>. This is not only the border between Denmark and Germany, but also the border between Germany and the other countries of Scandinavia. The Nordic Passport Union, which my fellow countryman Mr Krarup says must not be called thus, but referred to as a Nordic passport-free area, is not just a practical question, but in fact a political one as well, which says something about the Nordic context and also about Nordic solidarity. As a Dane, I think it is important to draw attention to this, in these days of Europe and Schengen.

As with presentational persuasion, Johnstone (1989) points out that analogical persuasion is more commonly found in 'eastern' cultures than in 'western' cultures. The fact that analogical

persuasion strategy seems to be the predominant style in the Danish texts under scrutiny may be difficult to explain with reference to contextual factors. However, by comparing my observations with previous contrastive observations, I believe that there are a number of plausible explanations. In a contrastive study of Danish and French, Lundquist (2010) concludes that Danish writers are more 'pragmatic' than French writers. Lundquist (2010, p. 16) states that Danes prefer reality over ideas, the particular over the general, concretisations over abstractions, and real life experiences over general principles, among other things. These preferences speak indeed in favour of an analogical persuasion strategy (and to a certain extent also a presentational). In the same vein, Korzen (2005a, p. 66) cites a number of scholars who define the Danish meeting culture as 'relaxed' and 'down to earth', characterised by a 'distrust of rhetoric' and a 'palaver democracy' (cf. also Hofstede, 1984; Lewis, 2006), suggesting that there could be a correlation between linguistic structures and social structures; cf. also Johnstone (1987, p. 91): 'the way we make discourse coherent reflects the way we make the world coherent'. In this respect, it is interesting to see that the 'founding father' of rhetoric in Denmark, Fafner, in his books, argued that rhetoric in a Danish context should be perceived as an 'oral intentionality' instead of being perceived as an 'art of persuasion' (Roer & Lund Klujeff, 2009, p. 13). Moreover, it is also interesting to see that comparisons of the writing traditions in Denmark and Italy also highlight the differences between analogical and quasilogical persuasive strategies:

> [The writing tradition] in Italy is strengthened through the high priority that is given to writing competence in the educational system. In accordance with rhetorical norms ("il bello scrivere"), emphasis is put on linguistic variation by choosing the appropriate level of style and expressing oneself in an abstract and synthetic manner throughout a written text. These are skills that are described in detail in Italian grammars and textbooks on written communication. By contrast, there is a much weaker writing tradition in the Danish educational system. (Skytte & Korzen, 2000, p. 49)[13]

---

[13] My translation of *'Den tidligere nævnte skriftsprogstradition i Italien forstærkes gennem den høje prioritering, man i undervisningen tillægger den skriftlige sprogfærdighed. Der lægges i overensstemmelse med retorikkens forskrifter ("il bello scrivere") vægt på sproglig variation med valg af korrekt stilniveau og evne til at udtrykke sig abstrakt og syntetisk i en velkomponeret fremstilling. Det er færdigheder, der er nøje beskrevet i italienske grammatikker og lærebøger vedr. skriftlig fremstilling. Over for dette står den skriftlige fremstilling svagt i det danske undervisningssystem'.*

In a last example from the Danish part of the corpus, the analogical style of the Danish writers is emphasised by the extensive use of analogies. Firstly, the writer compares the introduction of the euro with waving goodbye to the Titanic. Secondly, he sarcastically questions the legitimacy of the monetary union by asking which textbook the euro nations have used. Thirdly, he compares the introduction of the euro with the Chinese cultural revolution. And finally, he talks about EMU going against common sense.

170)   Fru formand, jeg har det lige nu som dem, der stod på kajen og vinkede farvel til det flotte skib Titanic, som med garanti ikke kunne synke. Jeg håber ikke, at de skændes så meget på kommandobroen, at skibet synker. Men skal det synke, så er det bedre, at det sker nu, hvor passagererne endnu kan svømme i land og bruge de nationale valutaer som redningsbælte. Den rigtige farefulde færd starter i år 2002, hvor redningsbælterne smides ud. Så kan der blive farligt mytteri med skænderier og det, der er værre, mellem f.eks. Frankrig og Tyskland. Euroen er et farligt eksperiment, som ender med at kunne lægge det nødvendige internationale samarbejde for had, og min gruppe har stemt nej i dag.

<P>

Jeg ville ellers gerne lykønske de 11 euro-lande med evnen til at se bort fra al økonomisk teori. I hvilken lærebog står det, at en møntunion fungerer bedst, når der kun er penge og valutapolitik? Hvor i alverden står det, at valutakursen skal være konstant, når konkurrenceevnen forringes eller forbedres? Jeg håber for de 11 lande, at eksperimentet lykkes. Det har trods alt større chance for at lykkes end den kinesiske kulturrevolution, hvor man også satte sig ud over økonomiens love. Men al fornuft siger, at ØMU'en i sin nuværende udformning enten lider skibbrud ved mødet med markedskræfterne eller må følges op af en politisk union med en fælles regering til at føre indkomstpolitik, finanspolitik og alle andre økonomiske politikker i præcise doseringer til gavn for menneskene i Europa. Husk: En fælles valuta er ikke livets mål, men højst et redskab til at gøre mennesker lykkelige. Da møntunionen blev vedtaget, havde vi 12 millioner ledige. Vi har nu over 18 millioner uden job, så hvis møntunionen skal kaldes en succes, så er det målestokken, der er gal. <ep-98-05-02.txt:64>

Madam President, I feel like someone standing on the quayside waving farewell to the beautiful, ship Titanic, which was guaranteed to be unsinkable. I sincerely hope that those in command on the bridge do not keep on squabbling until the ship sinks. But if sink it must, it is better for it to sink now, while the passengers can still swim ashore,

clutching their national currencies like life belts. The truly perilous voyage begins in the year 2002, when the life belts are to be cast away. Then there could be a dangerous mutiny, with squabbling and worse between France and Germany. The euro is a dangerous experiment leading to the necessary international collaboration for hatred, and my group has today voted against it.

<P>

I must congratulate the eleven euro nations on their ability to disregard all economic theory. Which textbook says that a monetary union performs best in terms of monetary and foreign exchange policy alone? Where on earth does it say that exchange rates should be fixed regardless of whether competitivity decreases or improves? For the sake of the eleven nations, I do hope the experiment succeeds. At the end of the day, its chances of success are greater than those of the Chinese cultural revolution, which also put itself beyond the laws of economics. But common sense tells us that the EMU in its present guise will either founder when it meets market forces, or it will have to be backed up by a political union with a common government determining income policy, fiscal policy and all other economic policies, meting out precise doses to benefit the people of Europe. Remember: a common currency is not the purpose of life, but at best, a tool to make people happy. When currency union was adopted, there were 12 million unemployed. We now have more than 18 million out of work, so if currency union is to be called a success, there must be something wrong with the yardstick.

7.4 Summing up

This chapter has presented a description of the linguistic structures of the Danish, English and Italian texts in the corpus from an intercultural rhetoric perspective. I have suggested that certain linguistic constraints and rhetorical preferences in the three languages under investigation prompt certain persuasive strategies, turning these strategies into the dominant persuasive styles of writers in these languages.

In the first part of the chapter, it was stated that my suggestions should be interpreted in the context of this study and not as universal features for all texts in the given languages. It was emphasised that any writer has access to a range of persuasive strategies, and that the chosen strategy is made on the basis of the rhetorical situation in question. I hope to have shown that language and contextual factors do not determine choices, but that they may lead to favouring certain choices over others. In line with many of the studies cited in this chapter from linguistic and rhetorical typology, I too found instances of all strategies across the texts in all three languages, but the overall impression was that each language showed tendencies to use one strategy to a larger extent than the other two strategies.

The second part of the chapter was devoted to correlating the linguistic structures of the texts studied in Chapters 4 to 6 with rhetorical strategies and styles. With reference to previous proposals for linguistic and rhetorical typologies, it was shown that the Italian writers make use of typical features of a quasilogical persuasion style in the texts analysed, and that the texts in general adhere to the characteristics of exocentric languages. I argued that the Italian writers employed more formal language and a more formal way of arguing than the Danish and English writers and correlated this with a tendency in Italian to concentrate on structure, both in grammars and communicative situations in general. In contrast, I found that the English writers typically make use of a presentational persuasion style, with features from both the exocentric and endocentric typologies. The writers involve themselves in a more personal way than the Italian writers and there is a stronger emphasis on words than structures. I argued that one of reasons could be that this strategy mainly has been correlated with people who in some way or another feel that their beliefs and values are threatened, which corresponds with the widely observed British euroscepticism. Lastly, I demonstrated how the Danish writers mainly applied an analogical persuasion style through an endocentric language type. The extensive use of different kinds of narratives in the Danish texts was correlated with the previously observed Danish pragmatism and tendency to avoid direct confrontations.

I have summarised these tendencies in Figure 7.3, which shows the continuum of persuasive styles considered in this chapter. The arrows between the three languages indicate that a smaller number of instances of presentational persuasion style were found in the Italian texts, a smaller number of instances of quasilogical and analogical persuasion style in the English texts, and a smaller number of instances of presentational persuasion style in the Danish texts, too.
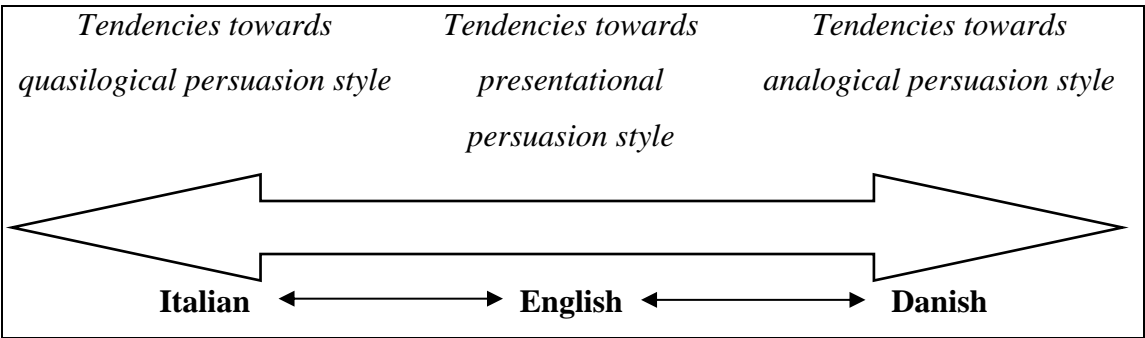
Figure 7.3: Persuasive styles in Danish, English and Italian

This chapter concludes the analysis and closes the circle around the cross-linguistic study of linguistic structures in parliamentary discourse across Danish, English and Italian. I have described linguistic structures, and the role of these in relation to the cognitive dimensions of

language in Chapter 2, then moved on to an in-depth analysis of the texts in Chapters 4 to 6, with particular attention paid to their relation to the lower level features of lexicon and morphology (see Figure 2.1). In this chapter, I have expanded the view of linguistic structures by correlating these with contextual features from the higher levels of intercultural rhetorical patterns and linguistic traditions.

8   Conclusions and perspectives

This thesis has presented a study of the structure of discourse. In particular, the thesis has addressed the similarities and differences in linguistic structures of argumentative texts in Danish, English and Italian. The linguistic structures were interpreted from a cognitive-functional perspective, as such comprising a number of textual phenomena that were analysed in a corpus of parliamentary discourse from the European Union. The corpus contained speeches held by native speakers of Danish, English and Italian in their mother tongue, covering a range of political topics, from global warming to in-house smoking policies. The speeches were held as parts of a larger debate, entailing relatively short texts with carefully presented and selected arguments.

The aim of this thesis was to study some of the linguistic means that writers employ when creating persuasive discourses. The linguistic means considered were conceptualised as belonging to three different structural domains which account for different ways of linking discourse units in a text: a syntactically organised text structure, a rhetorically organised discourse structure and an information packaging organised information structure. The structural domains were juxtaposed into a single analytical framework by drawing on insights from the cross-field of linguistics and rhetoric; in particular, from the research disciplines of text linguistics, discourse analysis and intercultural rhetoric.

A second aim of this thesis was to find out whether, and how, linguistic structures vary across different languages. In this study, the analyses were performed on Danish, English and Italian texts, following the same methodological steps. This comparison allowed us to examine how the three different languages structure and relate discourse units, and to discuss whether linguistic differences could be influenced by contextual factors such as linguistic and rhetorical traditions.

The main contributions of the thesis are as follows: (1) a unified framework for analysing text structure, discourse structure and information structure; (2) the application of English-based theories to Danish and Italian and comparisons between Danish, English and Italian text-building strategies; (3) a corpus-based characterisation of parliamentary discourse; (4) a study of the characteristics of text structure, discourse structure and information structure in parliamentary discourse for all three languages; and (5) a proposal for correlating linguistic structures with rhetorical strategies.

Below, I would like to concentrate on what I consider to be the perspectives of this thesis. First of all, an overview shall be given of the answers to the research questions posited in

Chapter 1, the introductory chapter. Secondly, reference shall be made to a number of possible applications of this analysis of linguistic structures from a more general point of view.

The first research question asked how various linguistic structures within and above the level of the sentence are related to each other. This question was addressed in Chapter 2, arguing that lower level features in lexicon and morphology influence higher level features in syntax, text, discourse and information structure, see Figure 2.1. I also showed that the lower levels are influenced by higher levels. Throughout the thesis, the view has been defended that we can interpret these interactions between the various linguistic structures by analysing three types of relations between Elementary Discourse Units (EDUs): syntactic relations (text structure), rhetorical relations (discourse structure) and information packaging relations (information structure). As part of these three structures, six objects of analysis were defined: syntactic dependency and morphological realisation accounting for text structure; rhetorical hierarchy and source of coherence accounting for discourse structure; and linkage and signal accounting for information structure. Figure 8.1 represents the various interactions between linguistic structures through the six objects of analysis. The idea behind the figure is that whenever the textualisation of one element is changed, the remaining elements are potentially influenced by this change. For example, if the linkage is changed between two EDUs, from intrasentential to intersentential, it may affect the signalling of a rhetorical relation, the realisation of the main verb and the dependency between the EDUs. Additionally, the change may also affect the source of coherence and the rhetorical hierarchy between the EDUs. I hope to have shown that the three linguistic structures of text, discourse and information should not be treated as separate structures with nothing in common, but that it can be highly beneficial to juxtapose them into one analytical framework, especially in a cross-linguistic context.

The second research question related to the manifestation of linguistic structures in argumentative discourse. By investigating the linguistic structures by object of analysis, we have observed that, in terms of text structure, syntactic subordination occurs more frequently than syntactic coordination, and finite constructions occur more frequently than non-finite and verbless constructions, in all three languages. In order to perform this analysis, a deverbalisation scale was introduced encompassing the various syntactic constructions, from independent sentences with all grammatico-semantic characteristics at their disposal to highly integrated non-finite and verbless constructions with hardly any syntactic independence left. This was shown in Chapter 4.
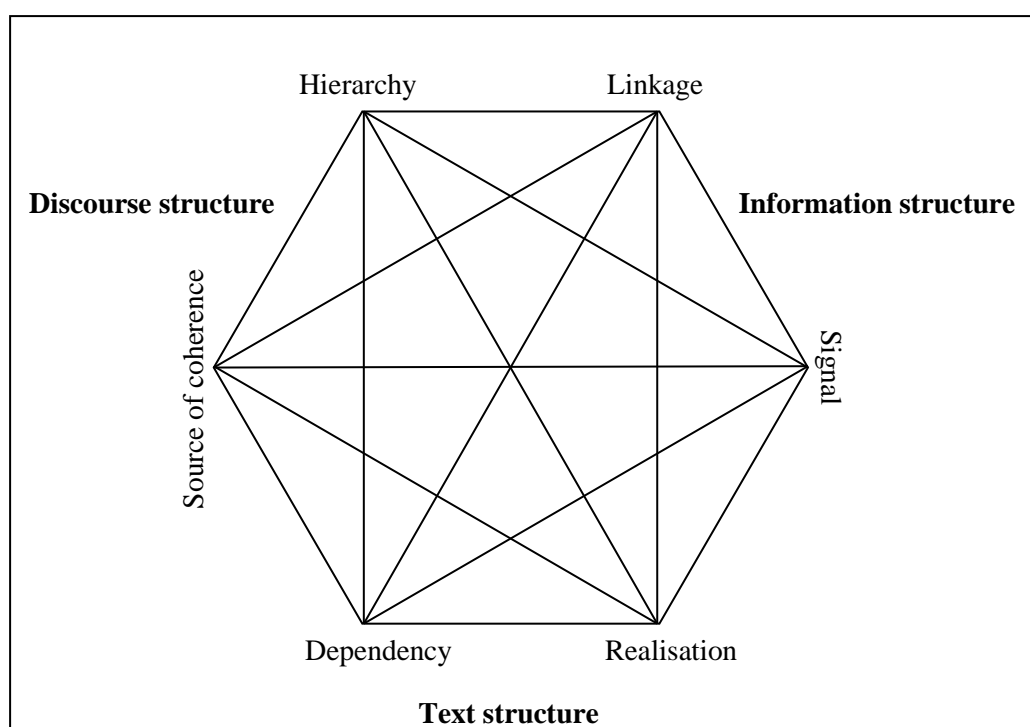
Figure 8.1: Interactions between linguistic structures

In terms of discourse structure, the rhetorical analysis using Rhetorical Structure Theory (RST) showed that a vast majority of the rhetorical relations between EDUs can be ascribed a semantic source of coherence relating events or situations, and that most of these relations occur in mononuclear text spans. The two most frequent rhetorical relations were Elaboration and Conjunction. The analysis testing the linguistic textualisation of the relations revealed that some relations were found mainly between EDUs inside the same sentence (Circumstance, Condition, Means, Otherwise, Purpose, Unconditional, Unless and Disjunction), whereas other relations typically were found between EDUs in different sentences (Solutionhood, Background, Enablement, Preparation, Summary, Joint and Multinuclear Restatement). I suggested that there could be a tendency for semantic relations to be textualised inside the same sentence, while pragmatic relations tend to be textualised across sentences. The analysis also showed that rhetorical coordination and subordination normally corresponded with syntactic coordination and subordination, although examples of all possible pairings between the four types were observed. This was shown in Chapter 5.

In the analysis of information structure, it was shown that most EDUs are linked intrasententially, and that the sentences in the texts of the corpus generally contain one to four EDUs per sentence. It was also observed that the rhetorical relations between EDUs often

remain implicit, that is, not signalled by any discourse cue. Discourse cues were found to link EDUs inside the same sentence more frequently than they were found to link EDUs in different sentences. The two most frequent discourse cues were *and* and *but* together with their cross-linguistic counterparts. This was shown in Chapter 6.

The third research question examined whether linguistic structures vary in argumentative discourse across Danish, English and Italian. A related question is whether English-based analysis theories and methods are adequate for the analysis of linguistic structures in Danish and Italian. The results reported in the different chapters of the thesis point to a relatively large number of differences in textualisation. The results can be summarised as follows. In the analysis of text structure, Danish diverged from English and Italian by using coordination between EDUs more frequently. Even though English and Italian showed similar distributions of subordination, the Italian texts had a more complex use of subordination, in that many EDUs were subordinated to other subordinate EDUs, resulting in a more hierarchical text structure. Danish also diverged in using almost exclusively finite verb forms, with the exception of some infinitives and other non-finite constructions. Given that many of these observations have been made in earlier studies, the results can be seen as a confirmation of basic characteristics of the typological differences between the three languages: Danish puts more weight on the verbs in EDUs, whereas English and Italian put more emphasis on the nouns, for example, by modifying these with participial constructions and postmodifiers.

Turning now to discourse structure, the analyses reveal no significant differences between Danish, English and Italian in their distribution of rhetorical relations. The analysis testing for types of relations found in the corpus showed the same distribution of semantic and pragmatic relations. However, the analysis testing the linguistic realisation of relations showed that Danish and English tend to distribute the sources of coherence almost equally in relation to linkage, whereas Italian in all cases displays a higher tendency to prefer intrasentential linkage of EDUs, in particular when it comes to pragmatic relations. Moreover, it could be observed that four relations in particular tend to be textualised differently in Danish and English than in Italian: Evaluation, Interpretation, Antithesis and Evidence. Other cross-linguistic differences were observed between the tendency in Danish to restrict non-finite textualisations to express Elaboration, Means or Purpose and the tendencies in English and Italian to allow non-finite textualisations to express a wide variety of rhetorical relations: Circumstance, Elaboration, Means, Purpose, Cause, Concession and Conjunction. Thus, the similarities in text structure patterns between English and Italian can in certain cases be refound in the discourse structure of

the two languages. In terms of rhetorical hierarchy, the differences observed were that English tended to have more rhetorically subordinated EDUs that were syntactically coordinated with other EDUS as compared with Danish and Italian. It was also observed that Italian tended to have more rhetorically coordinated EDUs that were syntactically subordinated to other EDUs. The most salient contribution of the analysis of discourse structure was that the writers in all three languages seemed to avail themselves of roughly the same relations in the same situations, but that the textualisation of the relations differed. The most likely conclusion here is that – regardless of linguistic constraints and rhetorical preferences – writers across the three languages generally use the same rhetorical relations to build up persuasive discourse; they merely textualise them differently. Text types and genres appear to determine the distribution of rhetorical relations more than languages do, which can be seen as confirmation that RST can also be applied to Danish and Italian texts. However, this calls for further investigation.

The analysis of information structures investigated linkage and signalling. The Italian texts tended to have many more EDUs inside the same sentence as compared with Danish and English. This was also one of the reasons that the Italian sentences were significantly longer than the Danish and English counterparts. In the English texts, rhetorical relations between EDUs were signalled less frequently than in the Danish and Italian texts. In addition, the three languages preferred different types of discourse cues: Danish favours coordinators and English subordinators, whereas Italian tends to have a more equal distribution of the different discourse cues. It was concluded that the Italian concept of a sentence does not necessarily correspond to that of Danish and English, and that considering EDUs as the object of analysis has proven to be useful in this cross-linguistic study. This was shown in Chapters 4 to 6.

The fourth and last research question looked for correlations between linguistic structures and contextual factors such as linguistic and rhetorical traditions. Chapter 7 provided a discussion of the possible correlations between language, cognition and contextual factors by mapping the findings of the previous three chapters, Chapters 4 to 6, with linguistic and rhetorical typologies. It was suggested that certain linguistic traditions and rhetorical preferences in the three languages under investigation evoke certain persuasive strategies. These strategies in turn become the dominant persuasive styles of the writers in these languages. With reference to previous proposals for linguistic and rhetorical typologies, it was shown how the Italian writers make use of typical features of quasilogical persuasion style in the texts analysed, and that the texts in general adhere to the characteristics of exocentric language types. It was argued that the Italian writers applied a more formal register and way of arguing than the Danish and English

writers. This finding was correlated with a tendency in Italian to put much focus on structure, both in grammars and communicative situations in general. By contrast, English writers were found to make use of the presentational persuasion style, with features from both the exocentric and endocentric language types. The writers involve themselves in a more personal way than the Italian writers and place more emphasis on words than structures. It was argued that one of reasons could be that this strategy has mainly been associated with people who in some way or other feel that their beliefs and values are threatened, which corresponds with the widely observed British euroscepticism. Lastly, I illustrated that the Danish writers mainly employed an analogical persuasion style through an endocentric language structure. The extensive use of different kinds of narratives in the Danish texts was correlated with Danish tendency to pragmatism and avoidance of direct confrontations.

A number of applications can be derived from this thesis. From a general point of view, it is a contribution to the study of how people from different cultures argue in favour and against a given topic. I also believe that it is a contribution to theoretical discourse analysis, since it provides insights into the analysis of argumentative discourse from a contrastive point of view. But more specifically, I refer to three major areas that can benefit from my results: second language teaching, translation studies and intercultural communication.

In second language teaching, the complete analysis of the characteristics of any text type provides useful information for learners who need to be proficient in that text type. In intercultural rhetoric, which takes its point of departure in second language teaching and English as a Foreign Language, the importance of studying variation within and above the level of the sentence has long been emphasised. Understanding intralinguistic variations between different text types is important, but understanding cross-linguistic variations in comparable text types is perhaps even more important in a second language teaching context. It is my hope that this thesis may assist in the process of designing materials and activities that will help students understand and produce texts and registers appropriately in a foreign language. I also hope that that a pairing of linguistics and rhetoric, such as presented in Chapter 7, may prompt more interaction between these two very parallel disciplines.

The field of translation studies concerns itself with many of the same aspects as second language teaching. Apart from possible applications similar to the ones mentioned above, I believe that my observations of cross-linguistic differences between the use of dependencies, realisation, linkage and discourse cues highlight important stylistic preferences which any translator working with languages that do not have the same typological configuration must be

aware of. Moreover, the subfield of statistical machine translation is still not completely capable of producing whole well-translated texts. I hope that some of the findings in this thesis may be used to generate better scripts and algorithms by scholars in this area.

Intercultural communication is another field that may benefit from the results of my thesis. Effective communication requires texts be organised into a coherent discourse structure. But languages vary considerably in how they do this, thereby posing a challenge for effective intercultural communication. Instead of relying on one's own preferred persuasion style to be the most appropriate in any context, we need to take into consideration that people with different linguistic and cultural backgrounds do not necessarily employ or expect the same linguistic means when attempting to persuade others in similar communication situations. This is of particular importance in a professional business context, and a more profound understanding of cross-linguistic differences in the organisation of texts across types and genres is needed. With this thesis, I hope to have demonstrated how this can be done systematically.

References

Abelen, E., Redeker, G., & Thompson, S. A. (1993). The rhetorical structure of US-American and Dutch fund-raising letters. *Text*, *13*(3), 323–350.

Abélès, M. (1992). *La vie quotidienne au parlement européen*. Paris: Hachette.

Abraham, W., Givón, T., & Thompson, S. A. (Eds.). (1995). *Discourse grammar and typology: Papers in honor of John W.M. Verhaar*. Amsterdam: John Benjamins.

Aijmer, K. (2002). *English discourse particles. Evidence from a corpus*. Amsterdam: John Benjamins.

Andersen, G. (2001). *Pragmatic markers and sociolinguistic variation: A relevance-theoretic approach to the language of adolescents*. Amsterdam: John Benjamins.

Anderson, P. J. (2005). A flag of convenience? Discourse and motivations of the London-based eurosceptic press. In R. Harmsen & M. Spiering (Eds.), *Euroscepticism: Party politics, national identity and European integration* (pp. 151–170). Amsterdam: Rodopi.

Artstein, R., & Poesio, M. (2008). Inter-coder agreement for computational linguistics. *Computational Linguistics*, *34*(4), 555–596.

Asher, N. (1993). *Reference to abstract objects in discourse*. Dordrecht & Boston: Kluwer Academic.

Asher, N., & Lascarides, A. (2003). *Logics of conversation*. Cambridge: Cambridge University Press.

Asher, N., & Vieu, L. (2005). Subordinating and coordinating discourse relations. *Lingua*, *115*(4), 591–610.

Austin, J. L. (1962). *How to do things with words*. Boston: Harvard University Press.

Azar, M. (1999). Argumentative text as rhetorical structure: An application of Rhetorical Structure Theory. *Argumentation*, *13*, 97–114.

Bache, C., & Kvistgaard Jakobsen, L. (1980). On the distinction between restrictive and non-restrictive relative clauses in modern English. *Lingua*, *52*, 243–267.

Ballard, D. L., Conrad, R., & Longacre, R. E. (1971). The deep and surface grammar of interclausal relations. *Foundations of language*, *4*, 70–118.

Ballard, M. (1995). La traduction de la conjonction "and" en francais. In M. Ballard (Ed.), *Relations discursives et traduction* (pp. 221–293). Lille: Presses Universitaires de Lille.

Baron, I. (Ed.). (2003a). *Language and culture*. Copenhagen: Samfundslitteratur.

Baron, I. (2003b). Diversité linguistique et cultures juridiques: les langues constituent-elles un obstacle à l'intégration européenne? In I. Baron (Ed.), *Language and culture* (pp. 153–166). Copenhagen: Samfundslitteratur.

Baron, I. (2007). Attention! Surveillance vidéo: La linguistique des langues et du droit. *Scolia*, *22*, 105–123.

Baron, I. (2009). Les notions de liberté et de démocratie au Danemark et en France: Valeurs universelles ou culturelles? In I. Korzen & C. Lavinio (Eds.), *Lingue, culture e testi istituzionali* (pp. 231–252). Florence: Franco Cesati.

Baroni, M., & Bernardini, S. (2005). A new approach to the study of translationese: Machine-learning the difference between original and translated text. *Literary and Linguistic Computing*, *21*(3), 259–274.

Bateman, J. A., & Rondhuis, K. J. (1997). Coherence relations: Towards a general specification. *Discourse Processes*, *24*(1), 3–49.

Bayley, P. (Ed.). (2004). *Cross-cultural perspectives on parliamentary discourse*. Amsterdam: John Benjamins.

Bazzanella, C. (1990). Phatic connectives as interactional cues in contemporary spoken Italian. *Journal of Pragmatics*, *14*, 629–647.

Bazzanella, C. (2008). *Linguistica e pragmatica del linguaggio*. Rome & Bari: Laterza.

Beamer, L. (2000). Finding a way to teach cultural dimensions. *Business Communication Quarterly*, *63*(3), 111–118.

Becher, V. (2011). *The use of parallel and comparable data for analysis of abstract anaphora in German and English*. PhD Thesis. Universität Hamburg.

Bellier, I. (2002). European identity, institutions and languages in the context of the enlargement. *Journal of Language and Politics*, *1*, 85–114.

Berman, R. A., & Slobin, D. I. (Eds.). (1994). *Events in narrative: A cross-linguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Biber, D. (1991). *Variation across speech and writing*. Cambridge: Cambridge University Press.

Biber, D., Connor, U., & Upton, T. A. (Eds.). (2007). *Discourse on the move: Using corpus analysis to describe discourse structure*. Amsterdam: John Benjamins.

Biber, D., Conrad, S., & Reppen, R. (Eds.). (1998). *Corpus Linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.

Bitzer, L. (1968). The rhetorical situation. *Philosophy & Rhetoric*, *1*(1), 1–14.

Blakemore, D. (1987). *Semantic constraints on Relevance*. Oxford: Blackwell Publishing.

Blakemore, D. (2002). *Relevance and linguistic meaning. The semantics and pragmatics of discourse markers*. Cambridge: Cambridge University Press.

Blakemore, D., & Carston, R. (2005). The pragmatics of sentential coordination with "and." *Lingua*, *115*, 569–589.

Blatt, F. (1946). *Latinsk syntaks i hovedtræk*. Copenhagen: Nordisk Forlag.

Blühdorn, H. (2008). Subordination and coordination in syntax, semantics and discourse: Evidence from the study of connectives. In C. Fabricius-Hansen & W. Ramm (Eds.), *"Subordination" versus "coordination" in sentence and text – from a cross-linguistic perspective* (pp. 59–85). Amsterdam: John Benjamins.

Born, J., & Schütte, W. (1995). *Eurotexte: Textarbeit in einer Institution der EG*. Tübingen: Narr.

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

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.

Bruner, J. (1986). *Actual minds, possible worlds*. Cambridge: Harvard University Press.

Buch-Kromann, M. (2010). Open challenges in treebanking: some thoughts based on the Copenhagen Dependency Treebanks. In *Nealt Proceedings Series* (Vol. 10).

Buch-Kromann, M., Gylling, M., Jelsbech Knudsen, L., Korzen, I., & Müller, H. H. (2010). *The inventory of linguistic relations used in the Copenhagen Dependency Treebanks*. Working paper. Copenhagen Business School.

Buch-Kromann, M., Hardt, D., & Korzen, I. (2011). Syntax-centered and semantics-centered views of discourse. Can they be reconciled? In S. Dipper & H. Zinsmeister (Eds.), *Beyond semantics: Corpus-based investigations of pragmatic and discourse phenomena* (pp. 17–30). Bochum: Bochumer Linguistische Arbeitsberichte 3.

Buch-Kromann, M., & Korzen, I. (2010). The unified annotation of syntax and discourse in the Copenhagen Dependency Treebanks. In *Proceedings of the Fourth Linguistic Annotation Workshop (LAW-IV)* (pp. 127–131). Uppsala, Sweden: ACL.

Buch-Kromann, M., Korzen, I., & Müller, H. H. (2009). Uncovering the "lost" structure of translations with parallel treebanks. *Copenhagen Studies in Language*, *38*, 199–224.

Byron, D., & Heeman, P. (1998). *Discourse marker use in task-oriented spoken dialogue*. Computer Science Technical Report 664. University of Rochester.

Carletta, J. C. (1996). Assessing agreement on classification tasks: The kappa statistic. *Computational Linguistics*, *22*(2), 249–254.

Carlson, L., & Marcu, D. (2001). *Discourse tagging reference manual*. ISI Technical Report ISI-TR-545.

Carlson, L., Marcu, D., & Okurowski, M. E. (2003). Building a discourse-tagged corpus in the framework of Rhetorical Structure Theory. In J. van Kuppevelt & R. Smith (Eds.), *Current directions in discourse and dialogue* (pp. 85–112). Dordrecht & Boston: Kluwer Academic Publisher.

Cartoni, B., & Meyer, T. (2012). Extracting directional and comparable corpora from a multilingual corpus for translation studies. In *Proceedings of the 8th international conference on Language Resources and Evaluation (LREC)* (Vol. 2, pp. 2132–2137). Istanbul.

Cerruti, M. (2009). Premesse per uno studio della variazione di registro in italiano. *Rivista italiana di dialettologia*, *33*, 267–282.

Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. Li (Ed.), *Subject and topic* (pp. 25–55). New York: Academic Press.

Chafe, W. (Ed.). (1980). *The Pear Stories: Cognitive, cultural and linguistic aspects of narrative production*. Norwood, NJ: Ablex.

Chafe, W. (1982). Integration and involvement in speaking, writing and oral literature. In D. Tannen (Ed.), *Spoken and written language: Exploring orality and literacy* (pp. 35–53). Norwood, NJ: Ablex.

Chafe, W. (2003). Discourse: overview. In *International Encyclopedia of Linguistics*. Oxford: Oxford University Press.

Chatman, S. (1990). *Coming to terms: The rhetoric of narrative in fiction and film*. Cornell: Cornell University Press.

Chesterman, A. (1998). *Contrastive functional analysis*. Amsterdam: John Benjamins.

Chilton, P. A. (2003). *Analysing political discourse: Theory and practice*. London: Taylor & Francis Group.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Boston: MIT Press.

Christiansen, T. (2011). *Cohesion: A discourse perspective*. Bern: Peter Lang.

Chuquet, H., & Paillard, M. (1987). *Approche linguistique des problèmes de traduction anglais-français*. Gap/Paris: Editions Ophrys.

Clark, H., & Haviland, S. (1977). Comprehension and the given-new contract. In R. Freedle (Ed.), *Discourse production and comprehension* (pp. 1–40). Hillsdale. NJ: Lawrence Erlbaum Associates.

Clark, H., & Schaefer, E. (1989). Contributing to discourse. *Cognitive Science*, *13*, 259–294.

Clark, H., & Wilkes-Gibbes, D. (1986). Referring as a collaborative process. *Cognition*, *22*, 1–39.

Clear writing throughout Europe. (2011). In *Proceedings of the conference "Clear writing throughout Europe."* Brussels: European Commission Directorate-General for Translation.

Clifford, C. (2008). *The making of eurosceptic Britain: Identity and economy ina post-imperial state*. Hampshire, UK: Ashgate Publishing.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46.

Collins, B., & Mees, I. (2003). *Practical phonetics and phonology: a resource book for students*. Oxford: Routledge.

Connor, U. (1996). *Contrastive rhetoric: Cross-cultural aspects of second-language writing*. Cambridge: Cambridge University Press.

Connor, U. (2002). New directions in contrastive rhetoric. *TESOL Quarterly*, *36*(4), 493–510.

Connor, U. (2004). Intercultural rhetoric research: beyond texts. *Journal of English for Academic Purposes*, *3*(4), 291–304.

Connor, U. (2011). *Intercultural rhetoric in the writing classroom*. Ann Arbor: University of Michigan Press.

Connor, U., & Moreno, A. I. (2005). Tertium Comparationis: A vital component in contrastive rhetoric research. In P. Bruthiaux & R. B. Kaplan (Eds.), *Directions in applied linguistics: essays in honour of Robert B. Kaplan* (pp. 153–164). Clevedon: Multilingual Matters Ltd.

Connor, U., Nagelhout, E., & Rozycki, W. (Eds.). (2008). *Contrastive rhetoric: Reaching to intercultural rhetoric*. Amsterdam: John Benjamins.

Conte, M.-E. (1999). *Condizioni di coerenza. Ricerche di linguistica testuale*. (B. Mortara Garavelli, Ed.) (2nd ed.). Alessandria: Edizioni dell'Orso.

Cook, P., & Bildhauer, F. (2011). Annotating information structure: The case of topic. In S. Dipper & H. Zinsmeister (Eds.), *Beyond semantics: Corpus-based investigations of pragmatic and discourse phenomena* (pp. 45–56). Bochum: Bochumer Linguistische Arbeitsberichte 3.

Cornish, F. (2009). Discourse anaphora and the FDG Contextual Component. In E. Keizer & G. Wanders (Eds.), *Web Papers in Functional Discourse Grammar 82* (pp. 97–115). Amsterdam: Universiteit van Amsterdam.

Coseriu, E. (1956). *La geografía lingüística*. Montevideo: Universidad de la Republica Faculdad de humanidas y ciencias.

Cosmai, D. (2000). Il linguaggio delle istituzioni comunitarie tra creazione terminologica e resa traduttiva. *Rivista internazionale di tecnica della traduzione*, *5*, 1–16.

Cosme, C. (2008). A corpus-based perspective on clause linking patterns in English, French and Dutch. In C. Fabricius-Hansen & W. Ramm (Eds.), *"Subordination" versus "Coordination" in Sentence and Text. A cross-linguistic perspective* (pp. 89–114). Amsterdam: John Benjamins.

Cresti, E., & Korzen, I. (Eds.). (2010). *Language, cognition and identity: Extensions of the endocentric/exocentric language typology*. Florence: Firenze University Press.

Cristofaro, S. (2003). *Subordination*. Oxford: Oxford University Press.

Cruse, A., & Croft, W. (2004). *Cognitive Linguistics*. West Nyack, NY: Cambridge University Press.

Da Cunha, I., & Iruskieta, M. (2010). Comparing rhetorical structures in different languages: The influence of translation strategies. *Discourse Studies*, *12*(5), 563–598.

Dardano, M., & Trifone, P. (1997). *La nuova grammatica della lingua italiana*. Bologna: Zanichelli.

De Beaugrande, R., & Dressler, W. (1981). *Introduction to text linguistics*. London & New York: Longman.

De Saussure, F. (1959). *Course in general linguistics*. (P. Meisel & H. Saussy, Eds.). New York: Columbia University Press.

De Swaan, A. (1999). Language and culture in transnational society. *European Review*, *7*(04), 507–518.

Degand, L., Lefèvre, N., & Bestgen, Y. (1999). The impact of connectives and anaphoric expressions on expository discourse comprehension. *Document Design*, *1*, 39–51.

Degand, L., & Sanders, T. (2002). The impact of relational markers on expository text comprehension in L1 and L2. *Reading and Writing*, *15*(7-8), 739–757.

Dinesh, N., Lee, A., Miltsakaki, E., Prasad, R., Joshi, A., & Webber, B. (2005). Attribution and the (non-)alignment of syntactic and discourse arguments of connectives. In *Proceedings of the Workshop on Frontiers in Corpus Annotation II: Pie in the Sky* (pp. 29–36).

Dirven, R. (2004). *Cognitive exploration of language and linguistics*. Amsterdam: John Benjamins.

Duranti, A. (1997). *Linguistic anthropology*. Cambridge: Cambridge University Press.

Durst-Andersen, P. (2010). Language, cognition and mentality. In P. Durst-Andersen & E. Lange (Eds.), *Mentality and thought: North, South, East and West* (pp. 29–50). Copenhagen: Copenhagen Business School Press.

Eggins, S., & Martin, J. R. (1997). Genres and registers of discourse. In T. A. van Dijk (Ed.), *Discourse as structure and process* (pp. 230–256). London: SAGE Publications.

Eisenhart, C., & Johnstone, B. (2008). Discourse analysis and rhetorical studies. In C. Eisenhart & B. Johnstone (Eds.), *Rhetoric in detail. Discourse analysis of rhetorical talk and text* (pp. 3–21). Amsterdam: John Benjamins.

English Style Guide. A handbook for authors and translators in the European Commission. (2012).

Enkvist, N. E. (1985). Introduction: Coherence, composition, and text linguistics. In N. E. Enkvist (Ed.), *Coherence and composition: A symposium* (pp. 11–26). Åbo, Finland: Tidningsbokhandeln.

Erman, B. (1992). Female and male usage of pragmatic exspressions in same-sex and mixed-sex interaction. *Language Variation and Change*, *4*, 217–234.

Ernout, A., & Thomas, F. (2002). *Syntaxe latine*. Paris: Librairie Klincksieck.

Fabricius-Hansen, C. (1996). Informational density: a problem for translation and translation theory. *Linguistics*, *34*(3), 521–565.

Fabricius-Hansen, C. (1998). Information density and translation, with special reference to German - Norwegian - English. In S. Johansson & S. Oksefjell (Eds.), *Corpora and cross-linguistic research: Theory, method, and case studies* (pp. 197–234). Amsterdam: Rodopi.

Fabricius-Hansen, C. (1999). Information packaging and translation: Aspects of translational sentence splitting (German–English/Norwegian). In M. M. Doherty (Ed.), *Sprachspezifische Aspekte der Informationsverteilung* (pp. 175–214). Berlin: Akademie-Verlag.

Fabricius-Hansen, C., & Ramm, W. (2008). Editors' introduction: Subordination and coordination from different perspectives. In C. Fabricius-Hansen & W. Ramm (Eds.), *"Subordination" versus "Coordination" in Sentence and Text. A cross-linguistic perspective* (pp. 1–30). Amsterdam: John Benjamins.

Fairclough, N. (1989). *Language and power*. London & New York: Longman.

Ferrari, A. (forthcoming). La linguistica del testo. In G. Iannaccaro (Ed.), *La linguistica italiana all'alba del terzo millennio (1997-2010)*. Rome: Bulzoni.

Ferrari, A., & Zampese, L. (2000). *Dalla frase al testo: una grammatica per l'italiano*. Bologna: Zanichelli.

Fetzer, A., & Speyer, A. (2012). Discourse relations in English and German discourse: Local and not-so-local constraints. *Intercultural Pragmatics*, *9*(4), 413–452.

Firbas, J. (1974). Some aspects of Czechoslovak approach to problems of functional sentence perspective. In F. Danes (Ed.), *Papers on Functional Sentence Perspective* (pp. 11–37). Prague: Academia.

Flowerdew, J. (2008). Critical discourse analysis and strategies of resistance. In V. K. Bhatia, J. Flowerdew, & R. H. Jones (Eds.), *Advances in discourse studies* (pp. 195–210). London: Routledge.

Fløttum, K. (Ed.). (2013). *Speaking of Europe : Approaches to complexity in European political discourse*. Amsterdam: John Benjamins.

Foley, W. A., & van Valin, R. D. (1984). *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.

Folman, S., & Sarig, G. (1990). Intercultural rhetorical differences in meaning construction. *Communication and Cognition*, *23*(1), 45–92.

Forbes, K., Miltsakaki, E., Prasad, R., Sarkar, A., Joshi, A., & Webber, B. (2002). D-LTAG system: Discourse parsing with a lexicalized tree adjoining grammar. *Journal of Logic, Language and Information*, *12*(3), 261–279.

Forster, A. (2002). *Euroscepticism in contemporary British politics: Opposition to Europe in the British conversative and labour parties since 1945*. New York: Routledge.

Foucault, M. (1969). *L'Archéologie du savoir*. Paris: Éditions Gallimard.

Fraser, B. (1988). Types of English discourse markers. *Acta Linguistica Hungarica*, *38*, 19–33.

Fraser, B. (1999). What are discourse markers? *Journal of Pragmatics*, *31*, 931–952.

Gabrielsen, J., & Juul Christiansen, T. (2010). *The power of speech*. Copenhagen: Hans Reitzels Forlag.

Galberg Jacobsen, H., & Skyum-Nielsen, P. (2003). *Dansk sprog. En grundbog*. Copenhagen: Schønberg.

Gale, W., & Church, K. (1993). A program for aligning sentences in bilingual corpora. *Computational Linguistics*, *19*(1), 75–102.

Gast, V., & Diessel, H. (Eds.). (2012). *Clause linkage in cross-linguistic perspective*. Berlin & Boston: De Gruyter.

Gast, V., & Schäfer, M. (2012). Relative clauses with adverbial meaning: A quantitative investigation of hybrid adjunct clauses in Latin. In V. Gast & H. Diessel (Eds.), *Clause Linkage in Cross-Linguistic Perspective* (pp. 363–392). Berlin & Boston: De Gruyter.

Gellerstam, M. (1986). Translationese in Swedish novels translated from English. In L. Wollin & H. Lindquist (Eds.), *Translation Studies in Scandinavia* (pp. 88–95). Lund: CWK Gleerup.

Girnth, H. (1996). Texte im politischen Diskurs. Ein Vorschlag zur diskursorientierten Beschreibung von Textsorten. *Muttersprache*, *106*(1), 66–80.

Givón, T. (1983). *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam: John Benjamins.

Givón, T. (2001). *Syntax: an introduction.* Amsterdam: John Benjamins.

Glenn, E. S., Witmeyer, D., & Stevenson, K. A. (1977). Cultural styles of persuasion. *International Journal of Intercultural Relations*, *1*(3), 52–66.

Grabski, M., & Stede, M. (2006). Bei: Intraclausal coherence relations illustrated with a German preposition. *Discourse Processes*, *41*(2), 195–219.

Granger, S. (2003). The corpus approach: a common way forward for Contrastive Linguistics and Translation Studies. In S. Granger, J. Lerot, & S. Petch-Tyson (Eds.), *Corpus-based Approaches to Contrastive Linguistics and Translation Studies* (pp. 17–29). Amsterdam: Rodopi.

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

Grimes, J. E. (1975). *The thread of discourse*. The Hague: Mouton.

Gross, A. G., & Dearin, R. D. (2003). *Chaim Perelman*. New York: SUNY Press.

Grosz, B. J., & Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational linguistics*, *12*(3), 175–204.

Gudykunst, W. B., & Kim, Y. Y. (2003). *Communicating with strangers* (4th ed.). New York: McGraw-Hill.

Gumperz, J. J., & Levinson, S. C. (Eds.). (1996). *Rethinking Linguistic Relativity*. Cambridge: Cambridge University Press.

Gylling, M. (forthcoming). English text structure in an endo-/exocentric perspective. In *Tra Romanista e Germanistica Lingua, Testo, Cognizione e Cultura*.

Gylling, M. (2012a). Relativsætningens retoriske funktioner. *Ny Forskning i Grammatik*, *19*, 65–81.

Gylling, M. (2012b). Causalità e conclusività nei discorsi politici: Un'ottica comparativa. In *Atti del XII Congresso SILFI - Dal manoscritto al web: canali e modalità di trasmissione dell'italiano*. Helsinki.

Gylling, M. (2012c). Comparing the structure of political discourse in three languages. In *II International Conference on Communication, Cognition and Media: Political and Economic Discourse* (pp. 96–97). Braga, Portugal: CATÓLICA: Universidade Catolica Portuguesa /Porto.

Gylling, M. (2013). Building a comparable corpus for annotating text and discourse structure. In *International Conference on Genre- and Register-related Text and Discourse Features in Multilingual Corpora* (pp. 39–41). Brussels.

Gylling, M., & Korzen, I. (2011). On discourse structure in Italian and Danish. In *Proceedings from Constraints of Discourse CID-IV* (pp. 5–11). Agay, France.

Haberlandt, K. (1982). Reader expectations in text comprehension. In J. F. Le Ny & W. Kintsch (Eds.), *Language and language comprehension* (pp. 239–249). Amsterdam: North-Holland.

Halliday, M. A. K. (1967). Notes on transitivity and theme in English, Parts 1-3. *Journal of Linguistics*, *3.1, 3.2,*.

Halliday, M. A. K. (1985). *An introduction to functional grammar*. London & Baltimore: E. Arnold.

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

Hansen-Schirra, S. (2007). Cohesive explicitness and explicitation in an English-German translation corpus. *Languages in Contrast*, *7*(2), 241–265.

Harder, P. (2005). Dansk Funktionel Lingvistik - en helhedsforståelse af forholdet mellem sprogstruktur, sprogbrug og kognition. In E. Engberg-Pedersen, M. Fortescue, P. Harder, L. Heltoft, M. Herslund, & L. F. Jakobsen (Eds.), *Dansk Funktionel Lingvistik - en helhedsforståelse af forholdet mellem sprogstruktur, sprogbrug og kognition* (pp. 1–38). Copenhagen: Københavns Universitet, CBS & RUC.

Hardt, D. (2013). A uniform syntax and discourse structure: The Copenhagen Dependency Treebanks. *Dialogue & Discourse*, *4*(2), 53–64.

Harris, Z. (1952). Discourse analysis. *Language*, *28*, 1–30.

Haspelmath, M. (Ed.). (2004). *Coordinating constructions*. Amsterdam: John Benjamins.

Hasund, K. (2002). Congratulation, like! - Gratulerer, liksom! Pragmatic particles in English and Norwegian. In L. E. Breivik & A. Hasselgren (Eds.), *From the COLT's mouth... and others. Language Corpora Studies in Honour of Anna-Brita Stenström* (pp. 125–139). Amsterdam: Rodopi.

Hatim, B. (1997). *Communication across cultures: translation theory and contrastive text linguistics*. Exeter, England: University of Exeter Press.

Heltoft, L., & Hansen, E. (2011). *Grammatik over det danske sprog*. Odense: Det Danske Sprog- og Litteraturselskab.

Herslund, M., & Baron, I. (2003). Language as world view. Endocentric and exocentric representations. In I. Baron (Ed.), *Language and culture* (pp. 29–42). Copenhagen: Samfundslitteratur.

Hobbs, J. R. (1979). Coherence and coreference. *Cognitive Science*, *6*, 67–90.

Hobbs, J. R. (1985). *On the coherence and structure of discourse*. Report No. CSLI-85-37. Stanford University.

Hobbs, J. R. (1990). *Literature and Cognition*. CSLI Report. Stanford University.

Hockett, C. F. (1958). *A course in modern linguistics*. MacMillan Education.

Hoey, M. (1983). *On the surface of discourse*. London & Boston: Allen & Unwin.

Hoey, M. (2001). *Textual interaction: an introduction to written discourse analysis*. London & New York: Routledge.

Hofstede, G. H. (1984). *Culture's consequences: International differences in work-related values*. Beverly Hills, CA: SAGE Publications.

Holliday, A. (1994). *Appropriate methodology and social context*. Cambridge: Cambridge University Press.

Holliday, A. (1999). Small cultures. *Applied Linguistics*, *20*, 237–264.

Hopper, P. J., & Thompson, S. A. (1984). The discourse basis for lexical categories in Universal Grammar. *Language*, *60*(4), 703–752.

Hovy, E., & Maier, E. (1992). *Parsimonious or profligate: How many and which discourse structure relations*. ISI Technical Report.

How to write clearly/Scrivere chiaro/Skriv klart. (2011). European Commission: Publications Office of the European Union.

Huddleston, R., & Pullum, G. (2002). *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Hymes, D. (1972). Models of interaction of language and social life. In J. J. Gumperz & D. Hymes (Eds.), *Directions in Sociolinguistics* (pp. 35–71). New York: Holt, Rinehart and Winston.

Ilie, C. (Ed.). (2010). *European parliaments under scrutiny: Discourse strategies and interaction practices*. Amsterdam: John Benjamins.

Irmer, M. (2011). *Bridging inferences: Constraining and resolving underspecification in discourse interpretation*. Berlin: Mouton de Gruyter.

Jansen, H. (2003). *Densità informativa. Tre parametri linguistico-testuali. Uno studio contrastivo inter- ed intralinguistico.* Copenhagen: Museum Tusculanum Press.

Johannessen, J. B. (1998). *Coordination*. Oxford: Oxford University Press.

Johansson, S. (2007). *Seeing through multilingual corpora: On the use of corpora in contrastive studies*. Amsterdam: John Benjamins.

Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science*, *4*, 71–115.

Johnson-Laird, P. N. (1983). *Mental models. Towards a Cognitive Science of Language, Inference and Consciousness*. Cambridge: Cambridge University Press.

Johnstone, B. (1983). Presentation as proof: The language of Arabic rhetoric. *Anthropological linguistics*, *25*, 47–60.

Johnstone, B. (1986). Arguments with Khomeini: Rhetorical situation and persuasive style in cross-cultural perspective. *Text*, *6*, 171–187.

Johnstone, B. (1987). Parataxis in Arabic: Modification as a model for persuasion. *Studies in language*, *11*, 85–98.

Johnstone, B. (1989). Linguistic strategies and cultural styles for persuasive discourse. In S. Ting-Toomey & F. Korzenny (Eds.), *Language, Communication, and Culture: Current Directions* (pp. 139–156). Beverly Hills, CA: SAGE Publications.

Johnstone, B. (Ed.). (2008). *Discourse Analysis* (2nd ed.). Malden: Blackwell Publishing.

Jucker, A. H., & Smith, S. W. (1998). And people just you know like "wow": Discourse markers as negotiating strategies. In A. W. Jucker & Y. Ziv (Eds.), *Discourse markers: Description and Theory* (pp. 171–202). Amsterdam: John Benjamins.

Just, S. N. (2009). No place like home? The role of the media in the framing of EUrope. *Journal of Language and Politics*, *8*(2), 244–268.

Jørgensen, C., & Villadsen, L. (Eds.). (2009). *Retorik. Teori og praksis*. Copenhagen: Samfundslitteratur.

Kaplan, R. B. (1966). Cultural thought patterns in intercultural education. *Language Learning*, *16*(1), 1–20.

Kaplan, R. B. (1988). Contrastive rhetoric and second language learning: Notes toward a theory of contrastive rhetoric. In A. C. Purves (Ed.), *Writing across languages and cultures. Issues in contrastive rhetoric* (pp. 275–304). Beverly Hills, CA: SAGE Publications.

Kapogianni, E. (2011). Graded salience effects on irony production and interpretation. In K. M. Jaszczolt & K. Allan (Eds.), *Salience and defaults in utterance processing* (pp. 53–81). Berlin: Mouton de Gruyter.

Kehler, A. (2002). *Coherence, reference, and the theory of grammar*. Stanford, CA.: CSLI Publications.

Klein, W., & Stutterheim, C. von. (1987). Quaestio und referentielle Bewegung in Erzählungen. *Linguistische Berichte*, *109*, 163–183.

Knott, A. (1996). *A data-driven methodology for motivating a set of coherence relations*. PhD Thesis. University of Edinburgh.

Knott, A., & Dale, R. (1994). Using linguistic phenomena to motivate a set of coherence relations. *Discourse Processes*, *18*(1), 35–62.

Knott, A., & Sanders, T. (1998). The classification of coherence relations and their linguistic markers: An exploration of two languages. *Journal of Pragmatics*, *30*, 135–175.

Koch, P., & Oesterreicher, W. (2001). Gesprochene Sprache und geschriebene Sprache. Langage parlé et langage écrit. In G. Holtus, M. Metzeltin, & C. Schmitt (Eds.), *Lexicon der romanistischen Linguistik* (Vol. I, 2., pp. 584–627). Tübingen: Max Niemeyer Verlag.

Koehn, P. (2005). Europarl: A parallel corpus for statistical machine translation. In *Proceedings of the 10th Machine Translation Summit* (Vol. 11, pp. 79–86). Phuket, Thailand.

Kortmann, B. (1996). *Adverbial subordination: A typology and history of adverbial subordinators based on European languages*. Berlin: Mouton de Gruyter.

Korzen, I. (1998). On the grammaticalisation of rhetorical satellites. A comparative study on Italian and Danish. *Copenhagen Studies in Language*, *22*, 65–86.

Korzen, I. (2003). Determinazione nominale e incorporazione in italiano: Un approccio pragmatico-testuale. *Cahiers Ferdinand de Saussure*, *56*, 35–65.

Korzen, I. (2005a). Sprogstruktur og tænkning: Endocentri og exocentri. In S. Bach, L. Cecchini, & A. Kratschmer (Eds.), *Ancora un'ultima occhiata* (pp. 47–71). Aarhus.

Korzen, I. (2005b). Struttura linguistica e schema cognitivo: Tipologie a confronto. In I. Korzen (Ed.), *Lingua, cultura e intercultura : L'italiano e le altre lingue. Atti del VIII Convegno*

Silfi. Societa Internazionale di Linguistica e Filologia Italiana - Copenhagen 22-26 giugno 2004 (pp. 123–134). Copenhagen: Samfundslitteratur.

Korzen, I. (2005c). Lingue endocentriche e lingue esocentriche: Lessico, testo e pensiero. In I. Korzen & P. D'Achille (Eds.), *Tipologia Linguistica e Società* (pp. 31–54). Florence: Franco Cesati.

Korzen, I. (2007). Linguistic typology, text structure and appositions. In I. Korzen, M. Lambert, & H. Vassiliadou (Eds.), *Langues d'Europe, l'Europe des langues. Croisements linguistiques, Scolia 22* (pp. 21–42).

Korzen, I. (2009a). Struttura testuale e anafora evolutiva: tipologia romanza e tipologia germanica. In I. Korzen & C. Lavinio (Eds.), *Lingue, culture e testi istituzionali* (pp. 33–60). Florence: Franco Cesati.

Korzen, I. (2009b). La costituzione italiana e danese: Valori impliciti ed espliciti. In I. Korzen & C. Lavinio (Eds.), *Lingue, culture e testi istituzionali* (pp. 201–230). Florence: Franco Cesati.

Korzen, I. (2010). L'italiano in una prospettiva de treebank: Il Copenhagen Dependency Treebank Project: Aspetti sintattici e testuali. In E. Cresti & I. Korzen (Eds.), *Language, Cognition and Identity: Extension of the Endocentric/Exocentric Language Typology* (pp. 77–99). Florence: Firenze University Press.

Korzen, I. (2011). La linguistica testuale comparativa rivisitata in chiave formale: L'italiano nelle Copenhagen Dependency Treebanks. *Journal of Romance Studies*, *8*, 255–269.

Korzen, I. (2012). Endo- og exocentrisk verbaltypologi: En genlæsning af Talmy - nu med (god) grund. *Ny Forskning i Grammatik*, *19*, 129–152.

Korzen, I., & Buch-Kromann, M. (2011). Anaphoric relations in the Copenhagen Dependency Treebanks. In S. Dipper & H. Zinsmeister (Eds.), *Beyond semantics: Corpus-based investigations of pragmatic and discourse phenomena* (pp. 83–98). Bochum: Bochumer Linguistische Arbeitsberichte 3.

Korzen, I., & D'Achille, P. (Eds.). (2005). *Tipologia linguistica e società. Considerazioni inter- e intralinguistiche.* Florence: Franco Cesati.

Korzen, I., & Gylling, M. (2011). What can contrastive linguistics tell us about translating discourse structure? In *Proceedings from Contrastive Linguistics - Translation Studies - Machine Translation - What can we learn from each other? (GSCL2011)* (pp. 5–11). Hamburg, Germany.

Korzen, I., & Gylling, M. (2012a). Text structure in a contrastive and translational perspective. *Translation: Computation, Corpora, Cognition*, *2*(1), 23–46.

Korzen, I., & Gylling, M. (2012b). Da mr. Bean alla Costituzione: Registri e stili di sei tipi di testo italiani. In *COLLOQUE ⊿IA II*. Copenhagen: University of Copenhagen.

Korzen, I., & Herslund, M. (Eds.). (1998). *Clause combining and text structure*. Copenhagen: Samfundslitteratur.

Korzen, I., & Lavinio, C. (Eds.). (2009). *Lingue, culture e testi istituzionali*. Florence: Franco Cesati.

Korzen, I., & Lundquist, L. (Eds.). (2003). *Sprogtypologi og oversættelse: Endocentriske og exocentriske sprog*. København: Samfundslitteratur.

Korzen, I., & Müller, H. H. (2011). The Copenhagen Dependency Treebanks: Forskellige niveauer - samme relationer. *Ny Forskning i Grammatik*, *18*, 173–196.

Krifka, M. (1993). Focus and presupposition in dynamic interpretation. *Journal of Semantics*, *10*, 269–300.

Kuppevelt, J. Van. (1995). Discourse structure, topicality and questioning. *Journal of Linguistics*, *31*(1), 109–147.

Källgren, G. (1979). Some types of textual cohesion and their effects on texts. In T. Pettersson (Ed.), *Paper from the Fifth Scandinavian Conference on Linguistics* (pp. 133–145). Stockholm: Almqvist & Wiksell.

Lakoff, R. (1974). Remarks on this and that. In *Proceedings of the tenth annual meeting of the Chicago Linguistics Society*. Chicago: University of Chicago.

Lambrecht, K. (1994). *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1).

Langacker, R. W. (1991). *Foundations of Cognitive Grammar: Descriptive application* (Vol. 2). Stanford, CA.: Stanford University Press.

Lehmann, C. (1988). Towards a typology of clause linkage. In J. Haiman & S. A. Thompson (Eds.), *Clause combining in grammar and discourse* (pp. 181–225). Amsterdam: John Benjamins.

Lenk, U. (1998). *Marking discourse coherence. Functions of discourse markers in spoken English*. Tübingen: Gunter Narr.

Lewis, R. D. (2006). *When cultures collide: Leading across cultures* (3rd ed.). Boston: Nicholas Brealey Publishing.

Lo Cascio, V. (1991). *Grammatica dell'argomentare. Strategie e strutture*. Florence: La Nuova Italia.

Longacre, R. E. (1983). *The grammar of discourse*. New York: Plenum Press.

Longacre, R. E., & Woods, F. (1976). *Discourse Grammar: Studies in indigenous languages of Colombia, Panama and Ecuador*. New York: Plenum Press.

Loock, R. (2010). *Appositive relative clauses in English: Discourse functions and competing structures*. Amsterdam: John Benjamins.

Loos, E. (2000). Language choice, linguistic capital and symbolic domination in the European Union. *Language Problems & Language Planning*, *24*(1), 37–53.

Loporcaro, M. (2013). *Profilo linguistico dei dialetti italiani*. Rome & Bari: Editori Laterza.

Lucy, J. A. (1992). *Language, diversity and thought*. Cambridge: Cambridge University Press.

Lundquist, L. (2003). Tekstlingvistik. In M. Herslund & B. Lihn Jensen (Eds.), *Sprog og sprogbeskrivelse* (pp. 227–253). Copenhagen: Samfundslitteratur.

Lundquist, L. (2005). Be orderly! Language types and interpretation of texts in the European Union. In I. Korzen & P. D'Achille (Eds.), *Tipologia Linguistica e Società* (pp. 110–133). Florence: Franco Cesati.

Lundquist, L. (2010). Cartesian Frenchmen and pragmatic Danes. A question of nouns and verbs? Working Paper, 74. Copenhagen: International Center for Business and Politics, Copenhagen Business School.

Lüngen, H., Puskàs, C., Bärenfänger, M., Hilbert, M., & Lobin, H. (2006). Discourse segmentation of German written text. In *Proceedings of the 5th International Conference on Natural Language Processing (FinTAL 2006)* (pp. 245–256). Berlin & Heidelberg: Springer-Verlag.

Lyons, J. (1977). *Semantics*. Cambridge: Cambridge University Press.

Maiden, M., & Parry, M. (1997). *The dialects of Italy*. London: Routledge.

Maiden, M., & Robustelli, C. (2007). *A reference grammar of modern Italian* (2nd ed.). McGraw-Hill.

Mamadouh, V. (1995). *De talen in het Europees parlement*. Amsterdam: Instituut voor Sociale Geografie, Universiteit van Amsterdam.

Mamadouh, V. (1999). A political-cultural map of Europe. Family structures and the origins of differences between national political cultures in the European Union. *GeoJournal*, *47*(3), 477–486.

Mann, W. C. (2005). RST website. Retrieved from http://www.sfu.ca/rst

Mann, W. C., & Matthiessen, C. M. I. M. (1991). Functions of language in two frameworks. *Word*, *42*, 231–250.

Mann, W. C., Matthiessen, C. M. I. M., & Thompson, S. A. (1992). Rhetorical Structure Theory and text analysis. In W. C. Mann & S. A. Thompson (Eds.), *Discourse Description: Diverse linguistic analyses of a fund-raising text* (pp. 39–78). Amsterdam: John Benjamins.

Mann, William C., & Thompson, S. A. (1987). *Rhetorical Structure Theory: A theory of text organization*. Technical report ISI/RS-97-190.

Mann, William C., & Thompson, S. A. (1988). Rhetorical Structure Theory: Toward a functional theory of text organization. *Text*, *8*(3), 243–281.

Marcato, C. (2007). *Dialetto, dialetti e italiano*. Bologna: Il Mulino.

Marcu, D., Romera, M., & Amorrortu, E. (1999). Experiments in constructing a corpus of discourse trees. In *Proceedings of the ACL Workshop on Standards and Tools for Discourse Tagging* (pp. 48–57).

Martin, J. R. (1992). *English text: System and structure*. Amsterdam: John Benjamins.

Martin, J. R., & Rose, D. (2003). *Working with discourse: Meaning beyond the clause*. London: Continuum.

Matthiessen, C. M. I. M., & Thompson, S. A. (1988). The structure of discourse and "subordination." In J. Haiman & S. A. Thompson (Eds.), *Clause Combining in Discourse and Grammar* (pp. 275–329). Amsterdam: John Benjamins.

McCool, M. (2009). *Writing around the world: A guide to writing across cultures*. London: Continuum International Publishing Group.

McEnery, T., Tono, Y., & Xiao, R. (2006). *Corpus based language studies: An advanced resource book*. London: Taylor & Francis Group.

McEnery, T., & Xiao, R. (2007). Parallel and comparable corpora: What is happening? In A. Gunilla & M. Rogers (Eds.), *Incorporating Corpora: Translation and the Linguist. Translating Europe* (pp. 18–31). Clevedon: Multilingual Matters Ltd.

Meyer, B. J. F., Brandt, D. M., & Bluth, G. J. (1980). Use of top-level structure in text: Key for reading comprehension of ninth-grade students. *Reading Research Quarterly*, *16*, 72–103.

Minsky, M. (1975). A framework for representing knowledge. In P. H. Winston (Ed.), *The Psychology of Computer Vision* (pp. 211–277). New York: McGraw-Hill.

Moder, C. L. (2004). Discourse across cultures, across disciplines: An overview. In C. L. Moder & A. Martinovic-Zic (Eds.), *Discourse across languages and cultures* (pp. 1–11). Amsterdam: John Benjamins.

Moreno, A. I. (2008). The importance of comparable corpora in cross-cultural studies. In U. Connor, E. Nagelhout, & W. Rozycki (Eds.), *Contrastive Rhetoric: Reaching to Intercultural Rhetoric* (pp. 25–41). Amsterdam: John Benjamins.

Mori, L. (2003). L'euroletto: genesi e sviluppo dell'italiano comunitario. In A. Valentini, P. Molinelli, P. Cuzzolin, & G. Bernini (Eds.), *Ecologia linguistica. Atti del XXXVI Convegno internazionale di Studi della Società di Linguistica italiana* (pp. 473–492). Rome: Bulzoni.

Mortara Garavelli, B. (1988). *Manuale di retorica*. Milan: Bompiani.

Mortensen, S. S. (2013). *Klassifikation af ledsætninger i dansk*. PhD Thesis. University of Copenhagen.

Mosegaard Hansen, M.-B. (1998). *The function of discourse particles: A study with special reference to spoken standard French*. Amsterdam: John Benjamins.

Muller, S. (2005). *Discourse markers in native and non-native English discourse*. Amsterdam: John Benjamins.

Muntigl, P. (2000). The European Union: Policy-making through organizational discursive practices. In P. Muntigl, G. Weiss, & R. Wodak (Eds.), *European Union discourses on un/employment. An interdisciplinary approach to employment policymaking and organizational change* (pp. 1–25). Amsterdam: John Benjamins.

Neff, J., Dafouz, E., Díez, M., Prieto, R., & Chaudron, C. (2004). Contrastive discourse analysis: Argumentative text in English and Spanish. In C. L. Moder & A. Martinovic-Zic (Eds.), *Discourse across languages and cultures* (pp. 267–283). Amsterdam: John Benjamins.

Nencioni, G. (1976). Parlato-parlato, parlato-scritto, parlato-recitato. In *Strumenti critici* (Vol. LX, pp. 1–56). Turin: Einaudi.

Nir, B., & Berman, R. A. (2010). Complex syntax as a window on contrastive rhetoric. *Journal of Pragmatics*, *42*(3), 744–765.

O'Donnell, M. (2004). RSTTool. Version 3.45.

Ochs, E. (1996). Linguistic resources for socializing humanity. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking Linguistic Relativity* (pp. 407–437). Cambridge: Cambridge University Press.

Ogina, R. (2007). *Il latino. Breve introduzione linguistica* (2nd ed.). Milan: Franco Angeli.

Okulska, U., & Cap, P. (2010). Analysis of political discourse: Landmarks, challenges and prospects. In U. Okulska & P. Cap (Eds.), *Perspectives in Politics and Discourse* (pp. 3–20). Amsterdam: John Benjamins.

Ong, W. (1982). *Orality and literacy*. London: Methuen.

Palmer, G. B. (1996). *Toward a theory of cultural linguistics*. Texas: Texas University Press.

Perelman, C. (1977). *L'empire rhétorique: rhétorique et argumentation*. Paris: J. Vrin.

Perelman, C., & Olbrechts-Tyteca, L. (1969). *The new rhetoric: A treatise on argumentation*. (J. Wilkenson & P. Weaver, Eds.). Notre Dame, IN: University of Notre Dame Press.

Perfetti, C. A., & Goldman, S. R. (1974). Thematization and sentence retrieval. *Journal of Verbal Learning and Verbal Behavior*, *13*, 70–79.

Peters, P. (2004). *The Cambridge guide to English usage*. Cambridge: Cambridge University Press.

Piemontese, M. E. (1998). Il linguaggio della pubblica amministrazione nell'Italia di oggi. Aspetti problematici della semplificazione linguistica. In G. Alfieri & A. Cassola (Eds.), *Atti del XXIX Congresso della Società di Linguistica Italiana* (pp. 269–292). Rome: Bulzoni.

Pierini, P. (2004). *Comparing Italian and English: An introdution*. Rome: Aracne.

Poesio, M., & Traum, D. R. (1997). Conversational actions and discourse situations. *Computational Intelligence*, *13*, 309–347.

Polanyi, L. (1988). A formal model of the structure of discourse. *Journal of Pragmatics*, *12*(5-6), 601–638.

Polanyi, L., Culy, C., Berg, M. van den, Thione, G. L., & Ahn, D. (2004). A rule based approach to discourse parsing. In *Proceedings of the SIGDIAL '04 Workshop* (pp. 108–117). Cambridge: ACL.

Pontoppidan, C. (2013). *Gør teksten klar. Principper for klart skriftsprog*. Copenhagen: Samfundslitteratur.

Prandi, M. (2010). Congiuntivo. In *Enciclopedia dell'Italiano*. Treccani.

Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joshi, A., & Webber, B. (2008). The Penn Discourse TreeBank 2.0. In *Proceedings of LREC*. Marrakech, Morocco.

Prince, E. F. (1986). On the syntactic marking of presupposed open propositions. *CLS (Parasession)*, *22*, 208–222.

Pym, A. (2005). Explaining Explicitation. In K. Karoly & Á. Fóris (Eds.), *New Trends in Translation Studies. In Honour of Kinga Klaudy* (pp. 29–34). Budapest: Akadémia Kiadó.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1972). *A comprehensive grammar of the English language*. London: Longman.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A grammar of contemporary English*. London: Longman.

Ramm, W., & Fabricius-Hansen, C. (2005). Coordination and discourse-structural salience from a cross-linguistic perspective. In M. Stede, C. Chiarcos, M. Grabski, & L. Lagerwerf

(Eds.), *Salience in Discourse: Multidisciplinary Approaches to Discourse 2005* (pp. 119–128). Münster: Nodus Publikationen.

Reidsma, D., & Carletta, J. C. (2008). Reliability measurement without limits. *Computational Linguistics*, *34*(3), 319–326.

Reisigl, M. (2007). *Nationale Rhetorik in Fest- und Gedenkreden. Eine diskursanalytische Studie zum „österreichischen Millennium" in den Jahren 1946 und 1996.* Tübingen: Stauffenburg.

Reisigl, M., & Wodak, R. (2009). The discourse-historical approach (DHA). In R. Wodak & M. Meyer (Eds.), *Methods for Critical Discourse Analysis* (2nd ed., pp. 87–121). London: SAGE Publications.

Renkema, J. (2004). *Introduction to Discourse Studies*. Amsterdam: John Benjamins.

Renkema, J. (Ed.). (2009a). *Discourse, of course: An overview of research in discourse studies*. Amsterdam: John Benjamins.

Renkema, J. (2009b). *The texture of discourse: Towards an outline of Connectivity Theory*. Amsterdam: John Benjamins.

Renzi, L., Salvi, G., & Cardinaletti, A. (Eds.). (1995). *Grande grammatica italiana di consultazione*. Bologna: Il Mulino.

Riazi, A. (2003). The invisible in translation: The role of text structure. *The Translation Journal*, *7*(2).

Rigotti, E. (2009). Gli agganci *forici nella struttura dei connettivi testuali. In F. Venier (Ed.), *Tra pragmatica e linguistica testuale* (pp. 421–440). Alessandria: Edizioni dell'Orso.

Rijkhoff, J. (2008). Layers, levels and contexts in Functional Discourse Grammar. In D. García Velasco & J. Rijkhoff (Eds.), *The Noun Phrase in Functional Discourse Grammar* (Vol. 195, pp. 63–115). Berlin/New York: Mouton de Gruyter.

Roer, H., & Lund Klujeff, M. (Eds.). (2009). *Retorikkens aktualitet*. Copenhagen: Hans Reitzels Forlag.

Ruiz Ruiz, J. (2009). Sociological discourse analysis: Methods and logic. *Forum: Qualitative Social Research*, *10*(2).

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organisation of turn-taking for conversation. *Language*, *50*, 696–735.

Sanders, T. (1997). Semantic and pragmatic sources of coherence: On the categorization of coherence relations in context. *Discourse Processes*, *24*(1), 119–147.

Sanders, T., & Noordman, L. G. M. (2000). The role of coherence relations and their linguistic markers in text processing. *Discourse Processes*, *15*, 1–35.

Sanders, T., Spooren, W. P. M., & Noordman, L. G. M. (1992). Toward a taxonomy of coherence relations. *Discourse Processes*, *15*, 1–35.

Sapir, E. (1964). *Culture, language and personality. Selected essays edited by Daniel G. Mandelbaum*. Berkeley-Los Angeles: University of California Press.

Sausy, L. (1977). *Grammaire latine complète*. Paris: Librairie Fernand Lanore.

Scarpa, F. (2001). *La traduzione specializzata. Lingue speciali e mediazione linguistica*. Milan: Hoepli.

Schank, R. C. (1972). Conceptual dependency: A theory of natural language understanding. *Cognitive Linguistics*, *3*, 552–631.

Schauer, H., & Hahn, U. (2001). Anaphoric cues for coherence relations. In *Proc. of 'Recent Advances in Natural Language Processing'- RANLP 2001*. Tzigov Chark, Bulgaria.

Schiffrin, D. (1987). *Discourse markers*. Cambridge: Cambridge University Press.

Searle, J. (1969). *Speech acts*. Cambridge: Cambridge University Press.

Sensini, M., & Roncoroni, F. (1997). *La grammatica della lingua italiana*. Milan: A. Mondadori.

Sharifian, F. (2011). *Cultural conceptualisations and language*. Amsterdam: John Benjamins.

Sharifian, F., & Palmer, G. B. (Eds.). (2007). *Applied cultural linguistics: Implications for second language learning and intercultural communication.* Amsterdam: John Benjamins.

Siegel, M. E. A. (2002). Like: The Discourse Particle and Semantics. *Journal of Semantics*, *19*(1), 35–71.

Skytte, G., & Korzen, I. (2000). *Italiensk-dansk sprogbrug i komparativt perspektiv*. Copenhagen: Samfundslitteratur.

Skytte, G., Korzen, I., Polito, P., & Strudsholm, E. (1999). *Tekststrukturering på italiensk og dansk / Strutturazione testuale in italiano e in danese.* (G. Skytte, I. Korzen, P. Polito, & E. Strudsholm, Eds.). Copenhagen: Museum Tusculanum Press.

Slembrouck, S. (2009). The parliamentary Hansard "verbatim" report: The written construction of spoken discourse. *Language and Literature*, *1*(2), 101–119.

Slobin, D. I. (1996a). From "thought and language" to "thinking for speaking." In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking Linguistic Relativity* (pp. 70–96). Cambridge: Cambridge University Press.

Slobin, D. I. (1996b). Two ways to travel: Verbs of motion in English and Spanish. In M. Shibatani & S. A. Thompson (Eds.), *Grammatical constructions: Their form and meaning* (pp. 195–219). Oxford: Clarendon Press.

Soria, C., & Ferrari, G. (1998). Lexical marking of discourse relations – some experimental findings. In *Proceedings of the ACL-98 Workshop on Discourse Relations and Discourse Markers* (pp. 36–42). Montreal, Canada.

Spooren, W. P. M. (1997). The processing of underspecified coherence relations. *Discourse Processes*, *24*, 149–168.

Stede, M. (2004). The Potsdam Commentary Corpus. In *Proceedings of the ACL 2004 Workshop on Discourse Annotation* (pp. 96–102).

Stede, M. (2008a). RST revisited: Disentangling nuclearity. In C. Fabricius-Hansen & W. Ramm (Eds.), *"Subordination" versus "coordination" in sentence and text – from a cross-linguistic perspective* (pp. 33–58). Amsterdam:: John Benjamins.

Stede, M. (2008b). Disambiguating rhetorical structure. *Research on Language and Computation*, *6*(3-4), 311–332.

Stede, M. (2011). *Discourse processing. Synthesis Lectures on Human Language Technologies* (Vol. 4). Morgan & Claypool Publishers.

Stubbs, M. (1983). *Discourse analysis: The sociolinguistic analysis of natural language*. Chicago: University of Chicago Press.

Stubbs, M. (1996). *Text and corpus analysis: Computer assisted studies of language and culture*. Oxford: Wiley-Blackwell.

Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Taboada, M. (2004a). Rhetorical relations in dialogue: A contrastive study. In C. L. Moder & A. Martinovic-Zic (Eds.), *Discourse across languages and cultures* (pp. 75–97). Amsterdam: John Benjamins.

Taboada, M. (2004b). *Building coherence and cohesion: Task-oriented dialogue in English and Spanish*. Amsterdam: John Benjamins.

Taboada, M. (2006). Discourse markers as signals (or not) of rhetorical relations. *Journal of Pragmatics*, *38*(4), 567–592.

Taboada, M. (2009). Implicit and explicit coherence relations. In J. Renkema (Ed.), *Discourse, of course: An overview of research in discourse studies* (pp. 127–140). Amsterdam: John Benjamins.

Taboada, M., & Mann, W. C. (2006a). Rhetorical structure theory: Looking back and moving ahead. *Discourse studies*, *8*(3), 423–460.

Taboada, M., & Mann, W. C. (2006b). Applications of rhetorical structure theory. *Discourse studies*, *8*(4), 567–588.

Taboada, M., Suárez, S. D., & Álvarez, E. G. (Eds.). (2013). *Contrastive discourse analysis. Functional and corpus perspectives*. Sheffield: Equinox.

Taboada, M., & Zabala, L. H. (2008). Deciding on units of analysis within Centering Theory. *Corpus Linguistics and Linguistic Theory*, *4*(1), 63–108.

Taft, M., Kacanas, D., Huen, W., & Chan, R. (2011). An empirical demonstration of contrastive rhetoric: Preference for rhetorical structure depends on one's first language. *Intercultural Pragmatics*, *8*(4), 503–516.

Talmy, L. (1985). Lexicalization patterns: semantic structure in lexcial forms. In T. Shopen (Ed.), *Language Typology and Syntactic Description, vol. III: Grammatical Categories and the Lexicon* (pp. 56–149). Cambridge: Cambridge University Press.

Talmy, L. (1987). The relation of grammar to cognition. In B. Rudzka-Ostyn (Ed.), *Topics on Cognitive Linguistics* (pp. 165–206). Amsterdam: John Benjamins.

Tannen, D. (1980). A comparative analysis of oral narrative strategies: Athenian Greek and American English. In W. Chafe (Ed.), *The Pear Stories: Cognitive, cultural and linguistic aspects of narrative production* (pp. 51–87). Norwood, NJ: Ablex.

Tannen, D. (1982). Analyzing discourse: Text and talk. In D. Tannen (Ed.), *Georgetown University Round Table on Languages and Linguistics* (pp. ix–xii). Washington D.C.: Georgetown University Press.

Teich, E. (2003). *Cross-linguistic variation in system and text: A methodology for the investigation of translations and comparable texts*. Berlin: Walter de Gruyter.

Thornbury, S. (2005). *Beyond the sentence: Introducing discourse analysis*. London: MacMillan Education.

Tofiloski, M., Brooke, J., & Taboada, M. (2009). A syntactic and lexical-based discourse segmenter. In *ACL Short 09 Proceedings of the ACL-IJCNLP 2009 Conference Short Papers* (pp. 77–80).

Tomlin, R. S. (1985). Foreground-background information and the syntax of subordination. *Text*, *5*, 85–122.

Toulmin, S. (1958). *The uses of argument*. Cambridge: Cambridge University Press.

Toury, G. (1995). *Descriptive Translation Studies and beyond*. Amsterdam: John Benjamins.

Traugott, E. C., & König, E. (1991). The semantics-pragmatics of grammaticalization revisited. In E. C. Traugott & B. Heine (Eds.), *Approaches to grammaticalization, vol. I* (pp. 189–218). Amsterdam: John Benjamins.

Traum, D. R. (1993). Rhetorical relations, action and intentionality in conversation. In *Proceedings of ACL SIG Workshop on Intentionality and Structure in Discourse Relations* (pp. 132–135).

Txurruka, I. G. (2000). *The semantics of "and" in discourse*. Technical Report ILCLI-00-LIC-9.

Vachek, J. (1966). *The Linguistic School of Prague. An introduction to its theory and practice*. London: Bloomington & London.

Vallduví, E. (1992). *The informational component*. New York & London: Garland Publishers.

Vallduví, E., & Engdahl, E. (1996). The linguistic realization of information packaging. *Linguistics*, *34*(3), 459–519.

Van der Vliet, N., Berzlánovich, I., Bouma, G., Egg, M., & Redeker, G. (2011). Building a discourse-annotated Dutch text corpus. In S. Dipper & H. Zinsmeister (Eds.), *Beyond semantics: Corpus-based investigations of pragmatic and discourse phenomena* (pp. 157–171). Bochumer Linguistische Arbeitsberichte 3.

Van der Vliet, N., & Redeker, G. (forthcoming-a). Explicit and implicit coherence relations in Dutch texts. In H. Gruber & G. Redeker (Eds.), *The Pragmatics of Discourse Coherence*.

Van der Vliet, N., & Redeker, G. (forthcoming-b). Exploring and evaluating RST annotations. In *MS2013*.

Van Dijk, T. A. (1979). Pragmatic connectives. *Journal of Pragmatics*, *3*, 447–456.

Van Dijk, T. A. (1980). *Macrostructures. An interdisciplinary study of global structures in discourse, interaction, and cognition*. Hillsdale, NJ: Lawrence Erlbaum.

Van Dijk, T. A. (2000). Parliamentary debates. In R. Wodak & T. A. van Dijk (Eds.), *Racism at the top: Parliamentary discourses on ethnic issues in six European states* (pp. 45–78). Klagenfurt: Drava Verlag.

Van Dijk, T. A. (2008). *Discourse and context: A socio-cognitive approach*. Cambridge: Cambridge University Press.

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

Van Halteren, H. (2008). Source language markers in EUROPARL translations. In *Proceedings of the 22nd International Conference on Computational Linguistics* (pp. 937–944). Manchester, UK: ACL.

Verspoor, M. H., & Sauter, K. (2000). *English sentence analysis: An introductory course*. Philadelphia: John Benjamins.

Vinay, J.-P., & Darbelnet, J. (1995). *Comparative stylistics of French and English: A methodology for translation*. (J. C. Sager & M.-J. Hamel, Eds.). Amsterdam: John Benjamins.

Völker, H. (2009). La linguistique variationnelle et la perspective intralinguistique. *Revue de Linguistique Romane*, *73*, 27–76.

Webber, B., Egg, M., & Kordoni, V. (2011). Discourse structure and language technology. *Natural Language Engineering*, *18*(4), 437–490.

Webber, B., & Prasad, R. (2009). Discourse structure: Swings and roundabouts. In B. Behrens & C. Fabricius-Hansen (Eds.), *Structuring information in discourse: The explicit/implicit dimension* (Vol. 1, pp. 171–190). Oslo Studies in Language.

Weinreich, U. (1954). Is a structural dialectology possible? *Word*, *10*, 388–400.

Werlich, E. (1983). *A text grammar of English*. Heidelberg: Quelle & Meyer.

Whorf, B. L. (1954). *Language, thought and reality. Selected writings of Benjamin Lee Whorf*. Cambridge: MIT Press.

Widdowson, H. G. (1979). *Explorations in applied linguistics*. Oxford: Oxford University Press.

Widdowson, H. G. (2004). *Text, context, pretext. Critical issues in Discourse Analysis*. Oxford: Blackwell Publishing.

Wodak, R. (2000). From conflict to consensus? The co-construction of a policy paper. In P. Muntigl, G. Weiss, & R. Wodak (Eds.), *European Union discourses on un/employment. An interdisciplinary approach to employment policy-making and organizational change* (pp. 73–114). Amsterdam: John Benjamins.

Wodak, R., & Chilton, P. A. (Eds.). (2005). *A new agenda in (Critical) Discourse Analysis*. Amsterdam: John Benjamins.

Wodak, R., & van Dijk, T. A. (Eds.). (2000). *Racism at the top: Parliamentary discourses on ethic issues in six European states*. Klagenfurt: Drava Verlag.

Wolf, F., & Gibson, E. (2005). Representing discourse coherence: A corpus-based study. *Computational Linguistics*, *31*(2), 249–287.

Wright, S. (2007). English in the European Parliament: MEPs and their language repertoires. *Sociolinguistica*, *21*, 151–165.

Writing for translation. (2010). Brussels: Translation Centre for the Bodies of the European Union.

Appendix A: Corpus overview with selected statistics

Language	File Name	Speaker	Sentences	Words	Words/ sentence
Danish	ep-00-01-18.txt:255	Camre	14	316	22.57
Danish	ep-00-02-02.txt:76	Krarup	8	320	40.00
Danish	ep-00-02-14.txt:70	Dybkjær	11	222	20.18
Danish	ep-00-02-15.txt:141	Busk	9	199	22.11
Danish	ep-00-03-13.txt:27	Bonde	18	332	18.44
Danish	ep-00-04-11.txt:271	Busk	13	264	20.31
Danish	ep-00-05-16.txt:52	Blak	28	407	14.54
Danish	ep-01-05-17.txt:43	Bonde	12	232	19.33
Danish	ep-01-07-05.txt:33	Blak	16	292	18.25
Danish	ep-01-09-05.txt:351	Rovsing	8	282	35.25
Danish	ep-01-10-24.txt:27	Krarup	14	243	17.36
Danish	ep-01-11-13.txt:318	Lund	16	324	20.25
Danish	ep-01-12-11.txt:98	Jensen	10	288	28.80
Danish	ep-02-09-04.txt:160	Haarder	12	263	21.92
Danish	ep-02-09-26.txt:25	Krarup	15	287	19.13
Danish	ep-02-10-23.txt:238	Haarder	8	264	33.00
Danish	ep-02-11-07.txt:11	Rovsing	8	238	29.75
Danish	ep-02-11-21.txt:34	Riis-Jørgensen	20	386	19.30
Danish	ep-02-12-18.txt:195	Haarder	10	299	29.90
Danish	ep-03-01-13.txt:129	Krarup	19	254	13.37
Danish	ep-03-02-11.txt:38	Sørensen	15	259	17.27
Danish	ep-03-04-08.txt:158	Sørensen	11	265	24.09
Danish	ep-03-05-13.txt:28	Andersen	12	202	16.83
Danish	ep-03-06-02.txt:67	Krarup	13	300	23.08
Danish	ep-03-09-03.txt:262	Andersen	13	269	20.69
Danish	ep-96-04-15.txt:56	Dybkjær	8	287	35.88
Danish	ep-96-06-19.txt:226	Kristoffersen	13	269	20.69
Danish	ep-96-07-04.txt:52	Rovsing	6	205	34.17

Danish	ep-96-09-18.txt:157	Jensen, Kirsten	10	277	27.70
Danish	ep-96-10-23.txt:28	Bonde	14	258	18.43
Danish	ep-96-12-11.txt:268	Dybkjær	10	338	33.80
Danish	ep-97-01-15.txt:207	Blak	7	221	31.57
Danish	ep-97-03-11.txt:126	Sindal	21	317	15.10
Danish	ep-97-06-09.txt:51	Dybkjær	10	296	29.60
Danish	ep-97-07-15.txt:277	Sindal	29	464	16.00
Danish	ep-97-10-22.txt:259	Dybkjær	7	210	30.00
Danish	ep-97-11-19.txt:161	Iversen	18	361	20.06
Danish	ep-98-01-13.txt:232	Jensen, Kirsten	11	229	20.82
Danish	ep-98-02-16.txt:20	Blak	19	331	17.42
Danish	ep-98-05-02.txt:64	Bonde	15	298	19.87
Danish	ep-98-06-16.txt:34	Iversen	11	227	20.64
Danish	ep-98-09-15.txt:97	Dybkjær	14	353	25.21
Danish	ep-98-11-16.txt:45	Kjer Hansen	10	355	35.50
Danish	ep-99-01-13.txt:44	Kofoed	17	331	19.47
Danish	ep-99-03-10.txt:31	Bonde & Sandbæk	10	332	33.20
Danish	ep-99-04-13.txt:31	Iversen	19	431	22.68
Danish	ep-99-05-06.txt:42	Sandbæk	13	260	20.00
Danish	ep-99-09-14.txt:25	Frahm	25	500	20.00
Danish	ep-99-10-05.txt:160	Nielson	12	275	22.92
Danish	ep-99-11-18.txt:143	Blak	16	305	19.06
English	ep-00-01-17.txt:40	Simpson	7	242	34.57
English	ep-00-02-02.txt:32	McMillan-Scott	12	279	23.25
English	ep-00-05-03.txt:120	Ahern	9	228	25.33
English	ep-00-09-04.txt:46	Patten	15	309	20.60
English	ep-00-10-02.txt:84	Ford	16	368	23.00
English	ep-00-11-13.txt:41	Andrews	14	228	16.29
English	ep-00-12-11.txt:42	Davies	12	291	24.25
English	ep-01-01-15.txt:54	Bushill-Matthews	9	268	29.78
English	ep-01-02-01.txt:12	Hyland	11	297	27.00
English	ep-01-04-02.txt:45	Evans, Jillian	12	316	26.33

English	ep-01-06-11.txt:59	Ludford	14	309	22.07
English	ep-01-10-01.txt:83	Farage	10	221	22.10
English	ep-01-11-12.txt:56	Corbett	13	334	25.69
English	ep-02-02-27.txt:83	McNally	9	201	22.33
English	ep-02-03-11.txt:79	Wallis	11	271	24.64
English	ep-02-05-29.txt:37	Farage	18	285	15.83
English	ep-02-07-01.txt:103	Hudghton	11	255	23.18
English	ep-02-09-23.txt:62	Lambert	10	334	33.40
English	ep-02-12-19.txt:83	MacCormick	8	201	25.13
English	ep-03-01-13.txt:64	Ó Neachtain	17	328	19.29
English	ep-03-01-29.txt:28	Evans, Jillian	13	356	27.38
English	ep-03-02-10.txt:116	Wyn	10	258	25.80
English	ep-03-03-10.txt:60	McNally	11	288	26.18
English	ep-03-04-09.txt:59	Wynn	12	305	25.42
English	ep-03-05-12.txt:29	Elles	6	224	37.33
English	ep-96-04-15.txt:28	Fitzsimons	11	240	21.82
English	ep-96-05-20.txt:33	Howitt	10	330	33.00
English	ep-96-07-03.txt:75	Watson	14	236	16.86
English	ep-96-09-04.txt:83	Tongue	17	335	19.71
English	ep-96-10-21.txt:42	Banotti	11	366	33.27
English	ep-96-12-09.txt:33	Crowley	12	260	21.67
English	ep-97-02-17.txt:30	Fitzsimons	7	251	35.86
English	ep-97-03-11.txt:36	McNally	11	300	27.27
English	ep-97-06-09.txt:99	Pollack	10	271	27.10
English	ep-97-07-14.txt:12	Titley	9	235	26.11
English	ep-97-09-16.txt:102	Wynn	22	361	16.41
English	ep-97-11-05.txt:183	Watts	18	464	25.78
English	ep-98-01-28.txt:80	Howitt	10	287	28.70
English	ep-98-03-09.txt:21	Green	10	254	25.40
English	ep-98-05-13.txt:201	McKenna	14	327	23.36
English	ep-98-06-15.txt:42	Ford	11	400	36.36
English	ep-98-09-15.txt:49	Collins, Gerard	13	354	27.23

English	ep-98-12-02.txt:150	Gallagher	12	266	22.17
English	ep-99-01-12.txt:178	Kinnock, Neil	9	279	31.00
English	ep-99-02-24.txt:58	Kinnock, Glenys	13	300	23.08
English	ep-99-04-12.txt:75	Gallagher	12	312	26.00
English	ep-99-05-04.txt:159	Spencer	12	271	22.58
English	ep-99-07-22.txt:25	Farage	13	255	19.62
English	ep-99-09-13.txt:59	Fitzsimons	18	446	24.78
English	ep-99-11-03.txt:27	Byrne	14	270	19.29
Italian	ep-00-02-02.txt:27	Bertinotti	12	307	25.58
Italian	ep-00-04-12.txt:41	Frassoni	17	575	33.82
Italian	ep-00-06-14.txt:176	Procacci	8	291	36.38
Italian	ep-00-09-04.txt:38	Pannella	9	240	26.67
Italian	ep-00-10-02.txt:102	Lisi	5	291	58.20
Italian	ep-00-11-13.txt:31	Vattimo	6	199	33.17
Italian	ep-00-12-13.txt:20	Dell'Alba	11	290	26.36
Italian	ep-01-01-15.txt:80	Caveri	10	325	32.50
Italian	ep-01-02-12.txt:64	Pittella	9	236	26.22
Italian	ep-01-04-02.txt:42	Nisticò	9	365	40.56
Italian	ep-01-09-04.txt:150	Dell'Alba	10	303	30.30
Italian	ep-01-10-01.txt:26	Ghilardotti	5	261	52.20
Italian	ep-01-12-11.txt:105	Pittella	10	215	21.50
Italian	ep-02-03-14.txt:124	Sbarbati	9	304	33.78
Italian	ep-02-04-11.txt:39	Podestà	9	327	36.33
Italian	ep-02-04-11.txt:43	Sartoni	7	277	39.57
Italian	ep-02-05-14.txt:71	Bigliardo	10	411	41.10
Italian	ep-02-10-21.txt:49	Trentin	6	287	47.83
Italian	ep-02-12-05.txt:14	Borghezio	5	264	52.80
Italian	ep-03-01-13.txt:73	Bodrato	6	209	34.83
Italian	ep-03-02-11.txt:273	Costa, Paolo	7	315	45.00
Italian	ep-03-03-11.txt:149	Della Vedova	9	288	32.00
Italian	ep-03-04-09.txt:235	Muscardini	5	319	63.80
Italian	ep-03-05-13.txt:150	Turchi	11	330	30.00

Italian	ep-03-07-01.txt:51	Fiori	9	273	30.33
Italian	ep-96-04-15.txt:59	Belleré	7	241	34.43
Italian	ep-96-05-08.txt:115	Tamino	8	317	39.63
Italian	ep-96-06-05.txt:85	La Malfa	4	277	69.25
Italian	ep-96-09-04.txt:57	Casini C.	10	367	36.70
Italian	ep-96-10-21.txt:57	Garosci	9	246	27.33
Italian	ep-96-12-09.txt:29	Tamino	8	222	27.75
Italian	ep-97-02-17.txt:42	Pettinari	7	249	35.57
Italian	ep-97-04-07.txt:81	Fantuzzi	9	307	34.11
Italian	ep-97-06-09.txt:52	Tamino	6	245	40.83
Italian	ep-97-09-15.txt:87	Belleré	5	225	45.00
Italian	ep-97-11-17.txt:24	Scapagnini	5	152	30.40
Italian	ep-97-12-03.txt:85	Moretti	5	209	41.80
Italian	ep-98-01-12.txt:43	Baldarelli	6	282	47.00
Italian	ep-98-03-11.txt:14	Caccavale	7	257	36.71
Italian	ep-98-05-02.txt:62	Garosci	13	293	22.54
Italian	ep-98-07-02.txt:29	Garosci	6	188	31.33
Italian	ep-98-10-05.txt:17	Dell'Alba	7	220	31.43
Italian	ep-98-11-18.txt:273	Filippi	9	215	23.89
Italian	ep-99-01-14.txt:178	Vecchi	8	296	37.00
Italian	ep-99-02-09.txt:276	Santini	14	396	28.29
Italian	ep-99-04-14.txt:157	Santini	18	359	19.94
Italian	ep-99-09-15.txt:33	Di Lello Finuoli	7	302	43.14
Italian	ep-99-10-05.txt:135	Prodi	12	291	24.25
Italian	ep-99-11-03.txt:59	Della Vedova	12	455	37.92
Italian	ep-99-12-14.txt:40	Prodi	24	668	27.83

Appendix B: Definitions of rhetorical relations

The following tables have been taken from Mann (2005).

N = Nucleus. S = Satellite. W = Writer. R = Reader.

Definitions of Presentational Relations (Pragmatic)			
Relation Name	**Constraints on either S or N individually**	**Constraints on N + S**	**Intention of W**
Antithesis	on N: W has positive regard for N	N and S are in contrast (see the Contrast relation); because of the incompatibility that arises from the contrast, one cannot have positive regard for both of those situations; comprehending S and the incompatibility between the situations increases R's positive regard for N	R's positive regard for N is increased
Background	on N: R won't comprehend N sufficiently before reading text of S	S increases the ability of R to comprehend an element in N	R's ability to comprehend N increases
Concession	on N: W has positive regard for N on S: W is not claiming that S does not hold;	W acknowledges a potential or apparent incompatibility between N and S; recognising the compatibility between N and S increases R's positive regard for N	R's positive regard for N is increased
Enablement	on N: presents an action by R (including accepting an	R comprehending S increases R's potential ability to perform the action in N	R's potential ability to perform the action in N increases

	offer), unrealised with respect to the context of N		
Evidence	on N: R might not believe N to a degree satisfactory to W on S: R believes S or will find it credible	R's comprehending S increases R's belief of N	R's belief of N is increased
Justify	none	R's comprehending S increases R's readiness to accept W's right to present N	R's readiness to accept W's right to present N is increased
Motivation	on N: N is an action in which R is the actor (including accepting an offer), unrealised with respect to the context of N	Comprehending S increases R's desire to perform action in N	R's desire to perform action in N is increased
Preparation	none	S precedes N in the text; S tends to make R more ready, interested or oriented for reading N	R is more ready, interested or oriented for reading N
Restatement	none	on N + S: S restates N, where S and N are of comparable bulk; N is more central to W's purposes than S is.	R recognises S as a restatement of N
Summary	on N: N must be more than one unit	S presents a restatement of the content of N, that is shorter in bulk	R recognises S as a shorter restatement of N

Definitions of Subject Matter Relations (Semantic)			
Relation Name	**Constraints on either S or N individually**	**Constraints on N + S**	**Intention of W**
Circumstance	on S: S is not unrealised	S sets a framework in the subject matter within which R is intended to interpret N	R recognises that S provides the framework for interpreting N
Condition	on S: S presents a hypothetical, future, or otherwise unrealised situation (relative to the situational context of S)	Realisation of N depends on realisation of S	R recognises how the realisation of N depends on the realisation of S
Elaboration	none	S presents additional detail about the situation or some element of subject matter which is presented in N or inferentially accessible in N in one or more of the ways listed below. In the list, if N presents the first member of any pair, then S includes the second: • set :: member • abstraction :: instance • whole :: part • process :: step • object :: attribute • generalisation :: specific	R recognises S as providing additional detail for N. R identifies the element of subject matter for which detail is provided.
Evaluation	none	on N + S: S relates N to degree of W's positive regard toward N.	R recognises that S assesses N and

			recognises the value it assigns
Interpretation	none	on N + S: S relates N to a framework of ideas not involved in N itself and not concerned with W's positive regard	R recognises that S relates N to a framework of ideas not involved in the knowledge presented in N itself
Means	on N: an activity	S presents a method or instrument which tends to make realisation of N more likely	R recognises that the method or instrument in S tends to make realisation of N more likely
Non-volitional Cause	on N: N is not a volitional action	S, by means other than motivating a volitional action, caused N; without the presentation of S, R might not know the particular cause of the situation; a presentation of N is more central than S to W's purposes in putting forth the N-S combination.	R recognises S as a cause of N
Non-volitional Result	on S: S is not a volitional action	N caused S; presentation of N is more central to W's purposes in putting forth the N-S combination than is the presentation of S.	R recognises that N could have caused the situation in S
Otherwise	on N: N is an unrealised situation on S: S is an unrealised situation	realisation of N prevents realisation of S	R recognises the dependency relation of prevention between the realisation of N and the realisation of S
Purpose	on N: N is an activity;	S is to be realised through the activity in N	R recognises that the activity in N is

	on S: S is a situation that is unrealised		initiated in order to realise S
Solutionhood	on S: S presents a problem	N is a solution to the problem presented in S;	R recognises N as a solution to the problem presented in S
Unconditional	on S: S conceivably could affect the realisation of N	N does not depend on S	R recognises that N does not depend on S
Unless	none	S affects the realisation of N; N is realised provided that S is not realised	R recognises that N is realised provided that S is not realised
Volitional Cause	on N: N is a volitional action or else a situation that could have arisen from a volitional action	S could have caused the agent of the volitional action in N to perform that action; without the presentation of S, R might not regard the action as motivated or know the particular motivation; N is more central to W's purposes in putting forth the N-S combination than S is.	R recognises S as a cause for the volitional action in N
Volitional Result	on S: S is a volitional action or a situation that could have arisen from a volitional action	N could have caused S; presentation of N is more central to W's purposes than is presentation of S;	R recognises that N could be a cause for the action or situation in S

Definitions of Multinuclear Relations		
Relation Name	**Constraints on each pair of N**	**Intention of W**
Conjunction	The items are conjoined to form a unit in which each item plays a comparable role	R recognises that the linked items are conjoined
Contrast	No more than two nuclei; the situations in these two nuclei are (a) comprehended as the same in many respects (b) comprehended as differing in a few respects and (c) compared with respect to one or more of these differences	R recognises the comparability and the difference(s) yielded by the comparison is being made
Disjunction	An item presents a (not necessarily exclusive) alternative for the other(s)	R recognises that the linked items are alternatives
Joint	None	None
List	An item comparable to others linked to it by the List relation	R recognises the comparability of linked items
Multinuclear Restatement	An item is primarily a reexpression of one linked to it; the items are of comparable importance to the purposes of W	R recognises the reexpression by the linked items
Sequence	There is a succession relationship between the situations in the nuclei	R recognises the succession relationships among the nuclei.

30. Ole Hinz
Den effektive forandringsleder: pilot, pædagog eller politiker?
Et studie i arbejdslederes meningstil-skrivninger i forbindelse med vellykket gennemførelse af ledelsesinitierede forandringsprojekter

31. Kjell-Åge Gotvassli
Et praksisbasert perspektiv på dynami-ske
læringsnettverk i toppidretten
Norsk ph.d., ej til salg gennem Samfundslitteratur

32. Henriette Langstrup Nielsen
Linking Healthcare
An inquiry into the changing perfor-mances of web-based technology for asthma monitoring

33. Karin Tweddell Levinsen
Virtuel Uddannelsespraksis
Master i IKT og Læring – et casestudie i hvordan proaktiv proceshåndtering kan forbedre praksis i virtuelle lærings-miljøer

34. Anika Liversage
Finding a Path
Labour Market Life Stories of Immigrant Professionals

35. Kasper Elmquist Jørgensen
Studier i samspillet mellem stat og erhvervsliv i Danmark under
1. verdenskrig

36. Finn Janning
A DIFFERENT STORY
Seduction, Conquest and Discovery

37. Patricia Ann Plackett
Strategic Management of the Radical Innovation Process
Leveraging Social Capital for Market Uncertainty Management

2006
1. Christian Vintergaard
Early Phases of Corporate Venturing

2. Niels Rom-Poulsen
Essays in Computational Finance

3. Tina Brandt Husman
Organisational Capabilities, Competitive Advantage & Project-Based Organisations
The Case of Advertising and Creative Good Production

4. Mette Rosenkrands Johansen
Practice at the top
– how top managers mobilise and use non-financial performance measures

5. Eva Parum
Corporate governance som strategisk kommunikations- og ledelsesværktøj

6. Susan Aagaard Petersen
Culture's Influence on Performance Management: The Case of a Danish Company in China

7. Thomas Nicolai Pedersen
The Discursive Constitution of Organi-zational Governance – Between unity and differentiation
The Case of the governance of environmental risks by World Bank environmental staff

8. Cynthia Selin
Volatile Visions: Transactons in Anticipatory Knowledge

9. Jesper Banghøj
Financial Accounting Information and Compensation in Danish Companies

10. Mikkel Lucas Overby
Strategic Alliances in Emerging High-Tech Markets: What's the Difference and does it Matter?

11. Tine Aage
External Information Acquisition of Industrial Districts and the Impact of Different Knowledge Creation Dimen-sions

A case study of the Fashion and Design Branch of the Industrial District of Montebelluna, NE Italy

12. Mikkel Flyverbom
Making the Global Information Society Governable
On the Governmentality of Multi-Stakeholder Networks

13. Anette Grønning
Personen bag
Tilstedevær i e-mail som inter-aktionsform mellem kunde og med-arbejder i dansk forsikringskontekst

14. Jørn Helder
One Company – One Language?
The NN-case

15. Lars Bjerregaard Mikkelsen
Differing perceptions of customer value
Development and application of a tool for mapping perceptions of customer value at both ends of customer-suppli-er dyads in industrial markets

16. Lise Granerud
Exploring Learning
Technological learning within small manufacturers in South Africa

17. Esben Rahbek Pedersen
Between Hopes and Realities: Reflections on the Promises and Practices of Corporate Social Responsibility (CSR)

18. Ramona Samson
The Cultural Integration Model and European Transformation.
The Case of Romania

2007
1. Jakob Vestergaard
Discipline in The Global Economy Panopticism and the Post-Washington Consensus

2. Heidi Lund Hansen
Spaces for learning and working
A qualitative study of change of work, management, vehicles of power and social practices in open offices

3. Sudhanshu Rai
Exploring the internal dynamics of software development teams during user analysis
A tension enabled Institutionalization Model; "Where process becomes the objective"

4. Norsk ph.d.
Ej til salg gennem Samfundslitteratur

5. Serden Ozcan
EXPLORING HETEROGENEITY IN ORGANIZATIONAL ACTIONS AND OUTCOMES
A Behavioural Perspective

6. Kim Sundtoft Hald
Inter-organizational Performance Measurement and Management in Action
– An Ethnography on the Construction of Management, Identity and Relationships

7. Tobias Lindeberg
Evaluative Technologies
Quality and the Multiplicity of Performance

8. Merete Wedell-Wedellsborg
Den globale soldat
Identitetsdannelse og identitetsledelse i multinationale militære organisatio-ner

9. Lars Frederiksen
Open Innovation Business Models
Innovation in firm-hosted online user communities and inter-firm project ventures in the music industry
– A collection of essays

10. Jonas Gabrielsen
Retorisk toposlære – fra statisk 'sted' til persuasiv aktivitet

12. Claus Bajlum
Essays on Credit Risk and Credit Derivatives

13. Anders Bojesen
The Performative Power of Competence – an Inquiry into Subjectivity and Social Technologies at Work

14. Satu Reijonen
Green and Fragile A Study on Markets and the Natural Environment

15. Ilduara Busta
Corporate Governance in Banking A European Study

16. Kristian Anders Hvass
A Boolean Analysis Predicting Industry Change: Innovation, Imitation & Business Models The Winning Hybrid: A case study of isomorphism in the airline industry

17. Trine Paludan
De uvidende og de udviklingsparate Identitet som mulighed og restriktion blandt fabriksarbejdere på det aftayloriserede fabriksgulv

18. Kristian Jakobsen
Foreign market entry in transition economies: Entry timing and mode choice

19. Jakob Elming
Syntactic reordering in statistical machine translation

20. Lars Brømsøe Termansen
Regional Computable General Equilibrium Models for Denmark Three papers laying the foundation for regional CGE models with agglomeration characteristics

21. Mia Reinholt
The Motivational Foundations of Knowledge Sharing

22. Frederikke Krogh-Meibom
The Co-Evolution of Institutions and Technology – A Neo-Institutional Understanding of Change Processes within the Business Press – the Case Study of Financial Times

23. Peter D. Ørberg Jensen
OFFSHORING OF ADVANCED AND HIGH-VALUE TECHNICAL SERVICES: ANTECEDENTS, PROCESS DYNAMICS AND FIRMLEVEL IMPACTS

24. Pham Thi Song Hanh
Functional Upgrading, Relational Capability and Export Performance of Vietnamese Wood Furniture Producers

25. Mads Vangkilde
Why wait? An Exploration of first-mover advantages among Danish e-grocers through a resource perspective

26. Hubert Buch-Hansen
Rethinking the History of European Level Merger Control A Critical Political Economy Perspective

2009

1. Vivian Lindhardsen
From Independent Ratings to Communal Ratings: A Study of CWA Raters' Decision-Making Behaviours

2. Guðrið Weihe
Public-Private Partnerships: Meaning and Practice

3. Chris Nøkkentved
Enabling Supply Networks with Collaborative Information Infrastructures An Empirical Investigation of Business Model Innovation in Supplier Relationship Management

4. Sara Louise Muhr
Wound, Interrupted – On the Vulnerability of Diversity Management

36. Annegrete Juul Nielsen
Traveling technologies and transformations in health care

37. Athur Mühlen-Schulte
Organising Development Power and Organisational Reform in the United Nations Development Programme

38. Louise Rygaard Jonas
Branding på butiksgulvet Et case-studie af kultur- og identitetsarbejdet i Kvickly

2011
1. Stefan Fraenkel
Key Success Factors for Sales Force Readiness during New Product Launch A Study of Product Launches in the Swedish Pharmaceutical Industry

2. Christian Plesner Rossing
International Transfer Pricing in Theory and Practice

3. Tobias Dam Hede
Samtalekunst og ledelsesdisciplin – en analyse af coachingsdiskursens genealogi og governmentality

4. Kim Pettersson
Essays on Audit Quality, Auditor Choice, and Equity Valuation

5. Henrik Merkelsen
The expert-lay controversy in risk research and management. Effects of institutional distances. Studies of risk definitions, perceptions, management and communication

6. Simon S. Torp
Employee Stock Ownership: Effect on Strategic Management and Performance

7. Mie Harder
Internal Antecedents of Management Innovation

8. Ole Helby Petersen
Public-Private Partnerships: Policy and Regulation – With Comparative and Multi-level Case Studies from Denmark and Ireland

9. Morten Krogh Petersen
'Good' Outcomes. Handling Multiplicity in Government Communication

10. Kristian Tangsgaard Hvelplund
Allocation of cognitive resources in translation - an eye-tracking and keylogging study

11. Moshe Yonatany
The Internationalization Process of Digital Service Providers

12. Anne Vestergaard
Distance and Suffering Humanitarian Discourse in the age of Mediatization

13. Thorsten Mikkelsen
Personligheds indflydelse på forretningsrelationer

14. Jane Thostrup Jagd
Hvorfor fortsætter fusionsbølgen udover "the tipping point"? – en empirisk analyse af information og kognitioner om fusioner

15. Gregory Gimpel
Value-driven Adoption and Consumption of Technology: Understanding Technology Decision Making

16. Thomas Stengade Sønderskov
Den nye mulighed Social innovation i en forretningsmæssig kontekst

17. Jeppe Christoffersen
Donor supported strategic alliances in developing countries

18. Vibeke Vad Baunsgaard
Dominant Ideological Modes of Rationality: Cross functional

integration in the process of product innovation

19. Throstur Olaf Sigurjonsson
Governance Failure and Icelands's Financial Collapse

20. Allan Sall Tang Andersen
Essays on the modeling of risks in interest-rate and inflation markets

21. Heidi Tscherning
Mobile Devices in Social Contexts

22. Birgitte Gorm Hansen
Adapting in the Knowledge Economy Lateral Strategies for Scientists and Those Who Study Them

23. Kristina Vaarst Andersen
Optimal Levels of Embeddedness The Contingent Value of Networked Collaboration

24. Justine Grønbæk Pors
Noisy Management A History of Danish School Governing from 1970-2010

25. Stefan Linder
Micro-foundations of Strategic Entrepreneurship Essays on Autonomous Strategic Action

26. Xin Li
Toward an Integrative Framework of National Competitiveness An application to China

27. Rune Thorbjørn Clausen
Værdifuld arkitektur Et eksplorativt studie af bygningers rolle i virksomheders værdiskabelse

28. Monica Viken
Markedsundersøkelser som bevis i varemerke- og markedsføringsrett

29. Christian Wymann
Tattooing The Economic and Artistic Constitution of a Social Phenomenon

30. Sanne Frandsen
Productive Incoherence A Case Study of Branding and Identity Struggles in a Low-Prestige Organization

31. Mads Stenbo Nielsen
Essays on Correlation Modelling

32. Ivan Häuser
Følelse og sprog Etablering af en ekspressiv kategori, eksemplificeret på russisk

33. Sebastian Schwenen
Security of Supply in Electricity Markets

2012

1. Peter Holm Andreasen
The Dynamics of Procurement Management - A Complexity Approach

2. Martin Haulrich
Data-Driven Bitext Dependency Parsing and Alignment

3. Line Kirkegaard
Konsulenten i den anden nat En undersøgelse af det intense arbejdsliv

4. Tonny Stenheim
Decision usefulness of goodwill under IFRS

5. Morten Lind Larsen
Produktivitet, vækst og velfærd Industrirådet og efterkrigstidens Danmark 1945 - 1958

6. Petter Berg
Cartel Damages and Cost Asymmetries

7. Lynn Kahle
Experiential Discourse in Marketing A methodical inquiry into practice and theory

8. Anne Roelsgaard Obling
Management of Emotions in Accelerated Medical Relationships

31. Fumiko Kano Glückstad
Bridging Remote Cultures: Cross-lingual concept mapping based on the information receiver's prior-knowledge

32. Henrik Barslund Fosse
Empirical Essays in International Trade

33. Peter Alexander Albrecht
Foundational hybridity and its reproduction
Security sector reform in Sierra Leone

34. Maja Rosenstock
CSR - hvor svært kan det være?
Kulturanalytisk casestudie om udfordringer og dilemmaer med at forankre Coops CSR-strategi

35. Jeanette Rasmussen
Tweens, medier og forbrug
Et studie af 10-12 årige danske børns brug af internettet, opfattelse og for-ståelse af markedsføring og forbrug

36. Ib Tunby Gulbrandsen
'This page is not intended for a US Audience'
A five-act spectacle on online communication, collaboration & organization.

37. Kasper Aalling Teilmann
Interactive Approaches to Rural Development

38. Mette Mogensen
The Organization(s) of Well-being and Productivity
(Re)assembling work in the Danish Post

39. Søren Friis Møller
From Disinterestedness to Engagement Towards Relational Leadership In the Cultural Sector

40. Nico Peter Berhausen
Management Control, Innovation and Strategic Objectives – Interactions and Convergence in Product Development Networks

41. Balder Onarheim
Creativity under Constraints
Creativity as Balancing 'Constrainedness'

42. Haoyong Zhou
Essays on Family Firms

43. Elisabeth Naima Mikkelsen
Making sense of organisational conflict
An empirical study of enacted sense-making in everyday conflict at work

2013

1. Jacob Lyngsie
Entrepreneurship in an Organizational Context

2. Signe Groth-Brodersen
Fra ledelse til selvet
En socialpsykologisk analyse af forholdet imellem selvledelse, ledelse og stress i det moderne arbejdsliv

3. Nis Høyrup Christensen
Shaping Markets: A Neoinstitutional Analysis of the Emerging Organizational Field of Renewable Energy in China

4. Christian Edelvold Berg
As a matter of size
THE IMPORTANCE OF CRITICAL MASS AND THE CONSEQUENCES OF SCARCITY FOR TELEVISION MARKETS

5. Christine D. Isakson
Coworker Influence and Labor Mobility Essays on Turnover, Entrepreneurship and Location Choice in the Danish Maritime Industry

6. Niels Joseph Jerne Lennon
Accounting Qualities in Practice Rhizomatic stories of representational faithfulness, decision making and control

7. Shannon O'Donnell
Making Ensemble Possible
How special groups organize for collaborative creativity in conditions of spatial variability and distance

TITLER I ATV PH.D.-SERIEN

1992

1. Niels Kornum
    *Servicesamkørsel – organisation, øko-
    nomi og planlægningsmetode*

1995

2. Verner Worm
    *Nordiske virksomheder i Kina
    Kulturspecifikke interaktionsrelationer
    ved nordiske virksomhedsetableringer i
    Kina*

1999

3. Mogens Bjerre
    *Key Account Management of Complex
    Strategic Relationships
    An Empirical Study of the Fast Moving
    Consumer Goods Industry*

2000

4. Lotte Darsø
    *Innovation in the Making
    Interaction Research with heteroge-
    neous Groups of Knowledge Workers
    creating new Knowledge and new
    Leads*

2001

5. Peter Hobolt Jensen
    *Managing Strategic Design Identities
    The case of the Lego Developer Net-
    work*

2002

6. Peter Lohmann
    *The Deleuzian Other of Organizational
    Change – Moving Perspectives of the
    Human*

7. Anne Marie Jess Hansen
    *To lead from a distance: The dynamic
    interplay between strategy and strate-
    gizing – A case study of the strategic
    management process*

2003

8. Lotte Henriksen
    *Videndeling
    – om organisatoriske og ledelsesmæs-
    sige udfordringer ved videndeling i
    praksis*

9. Niels Christian Nickelsen
    *Arrangements of Knowing: Coordi-
    nating Procedures Tools and Bodies in
    Industrial Production – a case study of
    the collective making of new products*

2005

10. Carsten Ørts Hansen
    *Konstruktion af ledelsesteknologier og
    effektivitet*

TITLER I DBA PH.D.-SERIEN

2007

1. Peter Kastrup-Misir
    *Endeavoring to Understand Market
    Orientation – and the concomitant
    co-mutation of the researched, the
    re searcher, the research itself and the
    truth*

2009

1. Torkild Leo Thellefsen
    *Fundamental Signs and Significance
    effects
    A Semeiotic outline of Fundamental
    Signs, Significance-effects, Knowledge
    Profiling and their use in Knowledge
    Organization and Branding*

2. Daniel Ronzani
    *When Bits Learn to Walk Don't Make
    Them Trip. Technological Innovation
    and the Role of Regulation by Law
    in Information Systems Research: the
    Case of Radio Frequency Identification
    (RFID)*

2010

1. Alexander Carnera
    *Magten over livet og livet som magt
    Studier i den biopolitiske ambivalens*