Bridging Inferences in Discourse Interpretation

Dissertation

zur Erlangung des akademischen Grades
doctor philosophiae
(Dr. phil.)

Eingereicht an der
Philologischen Fakultät der
Universität Leipzig

von

Matthias Irmer

Leipzig, im November 2009

Gutachter:

PD Dr. Johannes Dölling,
Institut für Linguistik, Universität Leipzig

Prof. Dr. Anita Steube,
Institut für Linguistik, Universität Leipzig

Prof. Dr. Nicholas Asher,
CNRS – Institut de Recherche en Informatique,
Université Paul Sabatier, Toulouse

Acknowledgements

Writing this thesis was only possible with the support of a number of people, to whom I want to give my acknowledgements. First of all, I would like to thank Johannes Dölling, who offered me the opportunity to do this work and who supervised my thesis. I am also obliged to the other members of the Institute of Linguistics at the University of Leipzig, especially Anita Steube, who agreed to be a member of the committee. And I am indebted to Nicholas Asher, who did not hesitate to act as an external reviewer of this thesis.

The first two years of the work on this project took place at the graduate school "Universality and Diversity: Linguistic Structures and Processes" at the University of Leipzig and the Max Planck Institute for Human Cognitive and Brain Sciences. I would like to thank my colleagues for the splendid working atmosphere, especially Petr Biskup, Kristin Börjesson, Andreas Bulk, Antje Lahne, Susann Lingel Zschernitz, Amelie Mahlstedt, Olav Mueller-Reichau, Stefan Sudhoff, and Helena Trompelt.

I owe special thanks to Doris Schönefeld, who gave me the opportunity to continue my work on this thesis while working for the Institute of English Studies of the University of Leipzig. Thanks go also to Uwe Junghanns, Olav Mueller-Reichau, and Gerhild Zybatow at the Institute of Slavic Studies, where I found the right place for finishing this work.

The scientific advance put forward in this thesis would not have been possible without the numerous critical reviews and comments on the talks I gave on topics of this work. To name a few, valuable comments were given by Mascha Averintseva-Klisch, Anton Benz, Sebastian Bücking, Berry Claus, Manfred Consten, Katja Jasinskaja, Lauri Karttunen, Claudia Maienborn, Louise McNally, Rick Nouwen, Antje Roßdeutscher, Marc Staudacher, and Marianne Vergez-Couret.

I am particularly grateful to Jerry Hobbs, Ariel Cohen, Markus Egg, Gerhard Jäger, and Henk Zeevat, who came to Leipzig for giving inspiring talks on topics related to the subject of this thesis. Furthermore, I would like to thank Henning Korrell and the audience of the "Semantikzirkel" at the Institute of Linguistics for fruitful discussions on parts of this work. I appreciate the help of Andrew McIntyre in putting the manuscript into proper English. Thanks go also to Manfred Stede, Collin Baker, and Carlos Subirats for providing me access to their language resources.

Parts of this work were covered by a scholarship of the German Research Foundation (Deutsche Forschungsgemeinschaft) and by a research grant of the State of Saxony. I would like to thank the responsible persons and authorities for making this possible.

Finally, I would like to thank my family for the continuous support throughout all the time I needed for writing the thesis. I want to thank my parents and my brother for helpful advices and proofreading the manuscript. And, last but by no means least, I am deeply grateful to Elena for supporting and bearing me all this time, and to my daughter Lara, who managed to learn two languages in the time I needed for writing, for providing me a counterbalance to the work on this thesis.

Contents

Introduction

Subject and Goal of the Thesis

Natural language texts and discourses are, in the first place, a means of communication. When an author or speaker makes an utterance, she wants to convey information to the addressee, viz. to the reader or hearer. Importantly, there is almost always more information conveyed than actually written or said. As a consequence hearers must infer the intended meaning of an utterance in its context. In order to interpret an utterance successfully, they must not only process the linguistically encoded information but also consider contextual knowledge from different information sources. Contextual knowledge includes, apart from specific facts of the utterance situation and general encyclopedic knowledge about the world, also the discourse context, which contains preceding utterances and relations between them.

Typically, utterances do not occur in isolation but several connected utterances form a text or a discourse, which is by and large coherent and structured. Let us consider a short example discourse:

(0.1) a. Today, John has visited a new flat.

b. The balcony is very nice,

c. but the bath is too small.

Utterance (0.1a) describes an event which took place on the day of utterance at an earlier time. Two new entities are introduced into the discourse: the individual which is referred to by *John*, and an object which is referred to by the indefinite noun phrase *a new flat*. The utterance (0.1b) describes the state of an object which is referred to by the definite noun phrase *the balcony*. This entity is new in the discourse, but it stands in an implicit relation to the flat introduced in (0.1a): the balcony is obviously part of the flat. This relationship is not expressed directly by linguistic means, but has to be inferred in context by the hearer. Apart from understanding the preceding utterance the speaker presupposes general knowledge about the world, namely that flats have certain properties: usually they have one or more rooms and a kitchen and they can be furnished by a bath and a balcony. Only by means of this additional contextual knowledge, the hearer can successfully interpret the utterance and connect it to the preceding utterance. As the discourse proceeds, the hearer constructs a structured mental representation of the discourse, in which is stored which entities are spoken about and which properties they have.

In our example, utterances (0.1b) and (0.1c) are subordinated to utterance (0.1a). They elaborate the first utterance. Moreover, (0.1b) and (0.1c) express a contrast which is indicated explicitly by the connective *but*. Neither these relations between single utterances nor relations between entities in the discourse model have to be expressed directly by linguistic means. All these relations can be conveyed only implicitly and consequently

have to be inferred by the hearer by means of pragmatic inferences. In this way, if an utterance is successfully interpreted, the inferred information constitutes a part of the discourse model constructed in the course of interpretation.

Pragmatic inferences, in particular bridging inferences as exemplified in (0.1), are the subject of this thesis. My goal is to find a suitable formalization that permits the integration of bridging inferences in the construction of a structured discourse representation.

Contribution of the Dissertation

In the recent literature, there is no uniform theory neither for the formalization of pragmatic inferences nor for the representation of discourse structures. Pragmatic inferences are characterized by a fundamental property: they are defeasible, or in other words, information showing up later can cancel conclusions already drawn. This property makes it difficult to formalize these inferences by means of classical logic. Rather it suggests the use of nonmonotonic inferences developed in research in Artificial Intelligence for formal representation. There are several different theories which are more or less suitable for modeling inferences in natural language interpretation. Likewise, for representing complex discourse structures there are several competing accounts which start from very different assumptions. In particular, theories which restrict themselves to giving a representation of the information conveyed by a linguistic utterance must be distinguished from theories which take intentions of discourse participants into account. Although each of these theories covers some important aspects of the role of bridging inferences in discourse interpretation, the assumed theoretical background assumptions can be quite divergent. Points of view range from psychological and cognitive over computational to formal approaches. It is a challenge to bridge gaps between research undertaken in these areas, bringing together rich cognitive modeling of contextual knowledge with powerful formal mechanisms for representing inferences in discourse interpretation.

This thesis compares existing competing theories and develops a representation of discourse structure that is especially suited for describing bridging inferences. It will be shown how explicit and implicit relationships between entities in the discourse can be integrated into the discourse model. The pragmatic rules to be used will be formalized by means of a nonmonotonic logic, and it will be described how models for discourses can be constructed on this foundation. Particular emphasis is put on bridging inferences involving eventualities or frames and their modeling by means of integrating world knowledge into a structured discourse representation. A second focus lies on the question of how bridging inferences triggered by a specific construction, clitic Left Dislocation in Romance languages, interact with discourse structure.

My starting point is the assumption that underspecified semantic representations, which can be derived from linguistically given information, are further enriched by pragmatic inferences. As a result, a coherent discourse emerges. Discourse interpretation consists in finding a suitable model which yields a structured representation of the discourse. In course of the interpretation process implicit information is made explicit: coherence relations between utterances are established and anaphoric links between entities are detected.

Plan of the Thesis

In the first part of this thesis I will lay out the theoretical foundations and review existing approaches to the core areas of this work. I will examine the role of inferences in discourse interpretation at different levels, and discuss their contribution to cohesion and coherence in structured discourses.

Chapter 1 first introduces the basic theoretical framework of semantic underspecification and pragmatic enrichment, followed by a discussion on how pragmatic inferences are conceived in competing theoretical frameworks. After an elaboration of important properties of pragmatic inferences, formal accounts to defeasible reasoning are introduced and compared. The chapter is closed by an examination of pragmatic inferences in texts and discourses comprising more than one sentence.

Chapter 2 takes a broader view and is concerned with pragmatic inferences as a reflection of cognitive inferences in communication. A central position is occupied by the notion of Common Ground, the shared knowledge base of speaker and hearer. Different approaches and definitions of this representation of contextual knowledge are compared, and insights from psycholinguistic experiments are evaluated. In the second part of this chapter, I will show how intentions can be formally modeled in a game-theoretic approach.

Chapter 3 describes how textual and contextual knowledge can be formally represented in a discourse model. First, I will examine anaphoric phenomena in discourses and discuss the concept of a discourse model. After that, I will introduce the notion of discourse referents and review attempts to explain the availability of referents for anaphoric reference. The last part of this chapter is devoted to theories of anaphora resolution, considering pragmatic, computational, and formal accounts. In particular, the conception of meaning as context update and the dynamic semantic Discourse Representation Theory are introduced.

Chapter 4 is dedicated to discourse structure. First, I will single out the main characteristics of a coherent discourse structure and compare various proposals for the form of discourse structure. Then, I will discuss the notion of discourse relations and its conceptions in different theories. The chapter will be closed by an examination of possible conceptions of discourse topics.

Chapter 5 discusses how the process of discourse interpretation can be treated by formal means. I will present theories of discourse interpretation that are based on formal accounts of defeasible reasoning, including the theories of Interpretation as Abduction, Model Generation, and Segmented Discourse Representation Theory. I will outline differences and commonalities of these theories.

In the second part of the thesis, the focus is directed to bridging inferences, or indirect anaphora, as a particular kind of pragmatic inferences, and their role in discourse interpretation is examined.

Chapter 6 takes a closer look on bridging anaphora. I will review corpus-based and psycholinguistic studies and present a classification of bridging anaphora, before examining which particular bridging relations can be involved in this type of anaphora. Then I will discuss various proposals for the resolution of bridging anaphora, before examining how bridging anaphora attach to the built-up discourse structure.

Chapter 7 presents a new approach to bridging involving eventualities. I will exploit the idea developed in Frame Semantics that world knowledge is organized in frames. With each eventuality introduced in a discourse, a corresponding frame is evoked in the discourse model. I will extend the discourse representation by including possibly underspecified

representations of frame elements, which can give clues for finding suitable antecedents in bridging anaphora.

In chapter 8, the suggested approach to bridging is applied on the resolution of bridging inferences triggered by a specific construction: clitic Left Dislocation. After the treatment of grammatical aspects of this construction, the main focus lies on the integration of these inferences into the constructed discourse model.

Finally, I will summarize and discuss the results of the presented work. Possible applications in computational linguistics and information science will be sketched, and open questions and problems for further research are named.

Chapter 1

Pragmatic Inferences

1.1 Semantic Underspecification and Pragmatic Enrichment

This section introduces some essential aspects of the theoretical framework underlying this thesis. As a starting point I take the observation that natural language utterances almost always transmit more information than actually said. The conveyed information involves more than what is linguistically encoded. The linguistically determined parts of meaning are, traditionally, the subject of natural language semantics. Current semantic theories vary in the way linguistic meaning is to be conceived.

Formal semantic theories describe the meaning of sentences by means of a precisely defined formal meta-language. The most important principle is the principle of compositionality (or Frege's principle): the meaning of complex expressions is composed by the lexical meaning of its constituents, according to the syntactic relationships between them. The semantic form of a sentence obtained in this way is purely linguistically determined and context-invariant. In classical theories of formal semantics (Montague, 1973)[1], the interpretation of sentences consists in assigning them truth conditions, i.e. the meaning of a sentence is determined by the conditions under which it is true. Critics of this approach often object that fine meaning subtleties get lost by reducing semantic content to truth conditions.

Cognitive semantic theories emerged from this criticism, and it seems that the two approaches are quite hostile to each other. Cognitive semanticists view meaning in terms of concepts and categories (cf. Lakoff, 1987). A similar view is taken in Jackendoff's conceptual semantics (Jackendoff, 1983). These approaches often reject or restrict the compositional nature of sentences, and instead of well-defined logical representations of meaning, they use intuitively more appealing, though not precisely definable forms of representation (cf. e.g. Langacker, 1987, 2008).

However, we will see in this thesis that compositional truth-conditional semantics and cognitive theories of human understanding can indeed be combined, yielding both a precise means of computing meaning and cognitive reliability. I can confidently agree with the view that meaning can be identified with conceptual structures. The content of linguistic expressions is not bound to language but exists independently from it. We will see in the next section what role a formal meaning composition plays in this conception. For that, I present an approach which relates formally represented linguistic meaning with conceptual knowledge as part of broader human cognition.

1 See Dowty et al. (1981) for an excellent introduction to Montague semantics.

1.1.1 Underspecified Semantics

A theory which seeks to give a formal foundation for a cognitively oriented semantics is the so-called two-level semantics developed by Bierwisch (1983); Bierwisch and Lang (1989). Although this theory was never very influential in formal semantics in its strict sense, nowadays its main assumptions are shared by many recent theories settled on the borderline between semantics and pragmatics. This theory makes a strict separation between semantic form (SF) and conceptual structure (CS). In the semantic form of an expression, solely that part of the meaning is contained that is determined by the grammar and independent from the context in which the expression is uttered.

At the semantic level SF, from the linguistically encoded data, an underspecified semantic representation is constructed. Underspecification means that only parts of the meaning of utterances are specified. Often, if specific contextual information is not considered, the conditions under which a semantic form of a sentence is true cannot be determined. It is only possible when the specific situation, in which a sentence is uttered, is known. Only then entities that are referred to in an utterance can be specified according to the context and related to conceptual knowledge.

At the conceptual level CS, information about real-world objects and individuals, as well as their properties and relations between them, are stored. This knowledge includes also more complex structures such as the typical course of stereotypical situations and causal dependencies between events. Relationships between concepts holding at this level are reflected at the linguistic level in the use of corresponding expressions. The underspecified semantic structures which are obtained at SF are further enriched by pragmatic inferences at CS. These inferences are a necessary part of the interpretation process: for an utterance to be understood as intended, referring expressions must be linked to entities in the world, and the utterance must be linked to its preceding discourse context. In this thesis, I will have a closer look at these inferences.

The two-level approach to semantics has been criticized in various aspects. Above all, it seems that it involves an unnecessary reduplication of meaning structures. Purely linguistic meaning often is rejected as a constructed layer without any psychological reality. However, the main advantage of the assumption of more than one layer of meaning is the availability of precise formal algorithms for the computation of grammatically determined meaning components. On the one hand, this point is of great importance for computational semantic and language processing theories. On the other hand, it can be argued that also in human language use it is plausible that the computation of grammatically determined meaning representations is carried out by effortless automatic subconscious processes before its integration with conceptual knowledge takes place, involving more complex inferences that require more cognitive effort. Indeed, there is interesting recent psycho- and neuro-linguistic evidence regarding semantic underspecification and pragmatic enrichment, cf. *inter alia* Pylkkänen and McElree (2006); Brennan and Pylkkänen (2008); Frisson (2009).

Techniques of semantic underspecification have been successfully applied in accounting for a broad range of linguistic phenomena, e.g. scope ambiguities, lexical ambiguity, meaning shifts and reinterpretations, etc. Many important generalizations can be drawn, which without the assumption of underspecification are very hard to explain. This becomes particularly clear in cases of systematic lexical ambiguity where underspecification is an efficient alternative to potentially huge and inefficient listings of mutually related

meanings[2]. In sum, I consider underspecification to be a suitable basis on which we can explain the interplay of grammatical and extra-linguistic factors in the interpretation of natural language utterances.

Since the emergence of two-level semantic in the work of Bierwisch, it has undergone many modifications (cf. *inter alia* Dölling, 1997, 2001; Blutner, 1998; Maienborn, 2001). As a response to the above mentioned critique of an independent layer of semantic forms, Dölling (2005, p. 166) argues that also grammatically determined formal meanings are based on conceptual knowledge. Semantic meaning components are immediately and systematically linked to units of conceptualization. As a consequence, semantic forms are seen as constituting a particular subset of conceptual structures.

Dölling proposes a *multi-level model of meaning* (first sketched in Dölling, 1997). The semantic form of an utterance incorporates the purely linguistically determined parts of its meaning, as in Bierwisch's theory. SF is differentiated into various subtypes, where certain structural enrichments can take place. The semantic form is further specified on several levels, where the first one is a parameter-fixed structure. Pragmatic enrichment does not happen in a single step, but rather constitutes an ongoing process of information completion on various layers. Dölling's approach is more precise than Bierwisch's original theory because it is at least partly formalized. It accounts for some key phenomena at the semantics/pragmatics interface such as metonymy and systematic polysemy.

At the level of SF, an underspecified meaning representation is assigned to every lexical unit[3]. This representation, or *basic semantic form* SF_B, is free of extra-linguistic encyclopedic knowledge. According to the principle of compositionality, the semantic form of a complex expression is derived from the semantic forms of its constituents and their syntactic configuration. This representation can contain placeholders for concepts of individuals and situations, as well as constraints on them and relations between them. As an instrument for building underspecified semantic forms serve free variables, which are placeholders for parameters that are not fully specified by the linguistically coded information. SF_B provides the basis for cases of systematic polysemy. For instance, the lexical item "newspaper" can have related, but different meanings according to the sentence context, as in (1.1). It refers to a physical object in (a), to a mental entity in (b), and to an institution in (c).

(1.1) a. This morning, John spilled all his coffee over the newspaper.

b. This morning, the military junta censored the newspaper.

c. The newspaper was founded ten years ago.

If we assume only one underspecified lexical item for "newspaper" in the lexicon, its type must be specified according to the domain in which it is used. Only then, we can successfully integrate it in the representation of the eventualities described by the sentences. At this stage, further operations may be necessary in order to maintain compositionality on the sentence level.

The *inflected semantic form* SF_I emerges from SF_B by means of operations which are obligatory for the semantic type of an expression and which possibly introduce additional

2 For a detailed discussion of this point, see Dölling (2005).

3 This is the core of "radical underspecification" as conceived by Blutner (1998).

parameters. Sometimes, an expression can shift its meaning in certain contexts in a way that it refers to an entity of a domain different from its original domain. This phenomenon is known as metonymy. For instance, the meaning of "the newspaper" in (1.2) is coerced into an author writing for a newspaper.

(1.2) This morning, the newspaper wrote about the riots in the suburbs of Paris.

A third kind of phenomenon, where the semantic type of an expression does not fit in its surrounding configuration, is accounted for by the *coerced semantic form* SF_C, which emerges from SF_B by means of type coercion, i.e. the usual type shift assumed in formal semantics (à la Partee, but different from Pustejovsky's notion of coercion), e.g. a type shift from an entity of type e to a generalized quantifier of type $\langle\langle e,t\rangle,t\rangle$. Details on how inflected and coerced semantic forms are obtained by virtue of applying suitable meaning shifting operators can be found in Dölling (2003, 2005).

In this way, from the semantic form obtained by the composition of lexical units according to their syntactic configurations, together with a restricted amount of pragmatically inferred contextual information, we arrive at the propositional content of an utterance. This first enrichment involves the fixation of open parameters in the semantic form, yielding the *parameter-fixed structure* PFS, which corresponds to a minimally enriched meaning representation. The "literal meaning" of utterances, or "what is said"[4], is the result of a fixation of parameters with default values. Truth values can be assigned to this meaning representation.

Further pragmatic inferences are then needed to obtain "what is meant", that is the communicated content of an utterance, or the intended meaning of a speech act. These inferences take place on additional layers of meaning between semantic form and conceptual content. The next section is devoted to these layers and the inferences mediating between them.

1.1.2 Accounts of Pragmatic Inference

In the recent literature, there is an ongoing discussion on the relationship between semantics and pragmatics (cf. *inter alia* Bach, 1994; Carston, 1999; Levinson, 2000; Recanati, 2004; Dölling, 2005). Various opposing proposals have been made. However, many researchers agree on the idea that two layers have to be distinguished. On the one hand, purely linguistically encoded "semantic" meaning is underspecified or underdetermined in the sense that it cannot be assigned a propositional value. As Carston puts it, the " language code drastically underdetermines the explicitly communicated propositional content of the utterance (hence its truth conditions)" (Carston, 1999, p. 85). On the other hand, *pragmatic* meaning always involves contextual knowledge. Truth-conditions can only be determined if certain context-dependent meaning specifications have been made already. In order to arrive at a propositional level, pragmatic enrichment takes place. In Carston's words: "the derivation of the proposition explicitly communicated is dependent on pragmatic inference" (ibd.). Thus, pragmatic enrichment means that the underdetermined semantic meaning is enriched by additional information that is not linguistically encoded.

4 Unlike Grice (1975) (see below), Dölling subsumes under "what is said" both literal meaning and cases of non-literal meaning, provided that it results from a systematic meaning shift, i.e. metonymy, aspectual shifts, etc.

These enrichments include some rather distinct linguistic phenomena. Most obviously, the referents of referring expressions ("the newspaper") have to be determined. Inherently contextual expressions such as deictical expressions (e.g. "there", "today") must be evaluated with respect to the context. The reference of anaphoric expressions (e.g. "she") has to be resolved. Implicit relationships between clauses (e.g. causal and temporal connections) have to be detected.

Pragmatic inferences are the device by virtue of which pragmatic enrichment takes place. In this thesis, the term "pragmatic inference" refers to any inference a hearer has to draw in order to determine the meaning conveyed by an utterance. Thus, pragmatic inferences, unlike semantic entailments, involve both linguistic and contextual knowledge. A second important difference to semantic entailments is that pragmatic inferences can be withdrawn when information to their contrary becomes available.

Various different conceptions of the nature of pragmatic inferences and the underlying principles governing them can be found in the literature. I will briefly review the main directions followed in recent work on pragmatics, concentrating on the derivation of propositional utterance content. After that, I turn in more detail to properties of pragmatic inferences.

1.1.2.1 Conversational Implicatures

The work of Paul Grice can be considered as the most influential work in pragmatics in the 20th century. Basically, he distinguished between two layers of meaning. First, the sentence meaning, or "what is said", is assumed to be captured by the theory of grammar. Second, the speaker-meaning, or "what is meant", includes inferences made in actual contexts by actual recipients. With his notion of conversational implicatures, he intended to bridge the gap between what is said and what is meant. The basic principle is a general "Cooperative Principle" for conversation participants regarding their contributions (Grice, 1975, p. 26f).

Cooperative Principle Contribute what is required by the accepted purpose of the conversation.

Four *maxims* specify what it means to be cooperative. It is assumed to be common knowledge of all conversation participants that these rules are followed in general.

Maxim of Quality Try to make your contribution one that is true. Do not say what you believe to be false. Do not say that for which you lack adequate evidence.

Maxim of Quantity Make your contribution as informative as is required for the current purposes of the exchange. Do not make your contribution more informative than is required.

Maxim of Relation Be relevant.

Maxim of Manner Be perspicuous: avoid obscurity and ambiguity, and strive for brevity and order.

Grice coined the term *implicature* as a general notion of that part of utterance meaning which goes beyond what is actually said and what does not contribute to the truth-conditional content of an utterance. Implicatures can be *conventional*, that is they are

generated from the meaning of particular lexical items, or *conversational*, that is they are a product of the cooperation principle and the conversational maxims. An important difference between these two is that conventional implicatures cannot be taken back, while conversational implicatures can be cancelled.

Crucial in Grice's theory is the assumption that conversational implicatures can also arise by flouting the maxims. This happens when a speaker says something so immensely violating the maxims that the hearer must infer that the speaker is implying something different. The emergence of a conversational implicature can be characterized as follows (cf. Grice, 1975, p. 31):

- Speaker A says p, and p is not immediately cooperative.
- There is no reason for assuming that S is not cooperative.
- A knows (and knows that B knows that she knows) that B sees that the supposition that A believes q is required.
- Then the utterance of p by S is cooperative in the given context, and q is conversationally implicated.

There are two basic assumptions involved: determinacy, which presupposes that the speakers always cooperate, and mutual knowledge, the assumption that conversation participants make use of complex reasoning about each other's beliefs and intentions. We will discuss different concepts of mutual knowledge in more detail in chapter 2.

Grice's conversational implicatures have two crucial properties: they are inferences in a narrow sense in that participants are aware of them and can draw them consciously. And they are post-propositional: they are drawn on the fact that a speaker has said a proposition. However, these two properties do not hold for the following cases, which nevertheless are widely acknowledged to be cases of conversational implicatures:

(1.3) René left the town and fell in love.

(1.4) René and Claudia have four children.

(1.3) implicates that René first left the town *and then* fell in love. (1.4) implicates that René and Claudia have *exactly* four children. So far, they seem to be conversational implicatures in Grice's sense. But look what happens if they fall within the scope of operators or are embedded in complex constructions:

(1.5) If René left the town and fell in love then his decision to leave the town was fortunate.

(1.6) René and Claudia have three or four children.

Here, it can be seen that the implicatures are drawn before the sentences are completed. The conditional in (1.5) does not make sense if the antecedent is not interpreted as a temporal sequence, and the disjunction in (1.6) is clearly interpreted as "exactly three or exactly four".

These cases are the starting point for a still ongoing discussion in the literature on the distinction between semantics and pragmatics, and several competing accounts have been proposed. Levinson (2000) called them intrusive implicatures, Bach (1994) called them implicitures, Carston (1988) called them explicatures, and Recanati (2003) called them embedded implicatures. In the following, the most important accounts are briefly discussed.

1.1.2.2 Generalized Conversational Implicatures

Although the notion of conversational implicatures as proposed by Grice is intuitively appealing and can explain virtually any type of inferences in linguistic communication, it has been proven to be very difficult to exactly define this notion. Neo-Gricean pragmatic theories have tried to reformulate the Gricean principles.

Horn (1984) took the maxim of Quantity as a starting point and distinguished two mutually related, competing principles, the principles of hearer economy (Q-Principle, "say as much as you can") and of speaker economy (R- or I-principle, "say as much as you must"). Q replaces the first part of Grice's maxim of Quantity, R/I stands for the second part and all the other maxims.

Inferences according to Q are mostly *scalar* and *clausal* implicatures. Scalar implicatures involve a *scale*, an n-tuple of expressions with related meanings, which is partially ordered in such a way that each element logically entails its successors. Examples of scales are <hot, warm>, <all, most, some>, or <know, believe>, or <and, or>. If a weaker expression is used then the Q-implicature arises that the stronger expression is not valid. The typical example for a scalar implicature is (1.7), where the utterance (a) carries the implicature (b).

(1.7) a. John ate some of the cookies.

b. John ate not all of the cookies.

Levinson (2000) extended Horn's reformulation of the Gricean maxims. He coined the term *generalized conversational implicatures* (GCI) as opposed to *particularized conversational implicatures* (PCI).

PCIs are triggered by the particular context of a specific utterance: they are inferences to a particular interpretation which holds only for specific and non-invariant contextual conditions.

GCIs are a kind of preferred or default interpretations, which are made by default in all contexts and only withdrawn in contexts where inconsistencies arise. They establish a third layer of meaning between the meaning of an expression ("what is said") and the meaning of a particular utterance token ("what is meant"), namely the meaning of an utterance type, thus generalizing over particular occurrences of a certain type of utterance. They do not depend on a specific context, rather they are triggered quasi-automatically. They can be seen as a part of the grammar because some conversational implicatures, "intrusive implicatures", contribute to the propositional content of utterances, as already outlined above. Here is another example:

(1.8) If you ate some of the cookies and no one else ate any, then there must still be some left. (Levinson, 2000, p. 205)

In this example, "some" triggers the scalar implicature "not all", without which the conditional does not make sense. The view that GCIs, or default implicatures, are part of the computational system of the grammar, is advocated by Chierchia (2004). He claims that they are "introduced locally and projected upwards in a way that mirrors the standard semantic recursion" (Chierchia, 2004, p. 40).

Levinson assumes a third principle besides Q and I, the M-principle of manner: what is said in a marked form describes a marked situation. M-inferences bears on a notion

of *markedness* instead of Horn scales. Marked forms are generally more complex (syntactically and/or phonetically) and have a less prototypical meaning than their unmarked counterparts (e.g. unmarked "kill" or "pink" vs. marked "cause to die" or "pale white").

The I-principle works in the opposite direction: what is said in an unmarked form refers to a stereotypical situation. This means in particular that as much as possible temporal, causal and referential connections between entities and situations are to be assumed, insofar the particular utterance context does not further specify stereotypical assumptions. I-inferences include conditional completion, conjunction strengthening, bridging anaphora, local coreference, and implicit relations in compound nouns (cf. Levinson, 2000, pp. 112f).

The principles compete with each other: on the one hand, the force of unification tries to minimize the speaker's effort, and on the other hand, the force of diversification minimizes the hearer's effort. Huang (1994, p. 146) states the following preference ranking: Clausal Q-implicatures are ranked higher than clausal Q-implicatures, then follow M-implicatures and I-implicatures. A reformulation of these principles and their interaction in the framework of optimality theory was proposed by Blutner et al. (1996); Blutner (2000).

1.1.2.3 Explicatures

Relevance Theory (Sperber and Wilson, 1986), unlike Grice and Neo-Gricean pragmaticians, does not state various maxims, rather tries to relate interpretation to only one cognitive principle, that of *Relevance*, which can be expressed as "human cognition tends to be geared to the maximization of relevance" (Wilson and Matsui, 1998).

Relevance is defined in terms of cognitive effects and processing effort: The greater the cognitive effects and the smaller the effort needed to achieve those effects, the greater the relevance. An utterance is considered to be optimally relevant to an addressee iff (i) it is relevant enough to be worth the addressee's processing effort, and (ii) it is the most relevant one that is compatible with the speaker's abilities and preferences.

The conception of the interface between semantics and pragmatics within Relevance Theory is centered around the notion of *explicatures* (Carston, 1988, 1999, 2004). Explicatures are the pragmatic contribution to the proposition expressed by an utterance, as opposed to implicatures which are seen as propositions that can be expressed independently from the utterance trigerring them.

Carston assumes that the propositional content of an utterance, or "what is said", is inferred by explicatures from its underspecified grammatically determined meaning. Carston's explicatures involve both obligatory contextual saturation processes like reference assignment and optional processes of "free enrichment", which do not necessarily occur in every context but are possibly influenced by reasoning about the speaker's intentions. In these cases, an expression receives a more specific interpretation that it literally encodes. For example, the following sentence is usually interpreted as "Jane did not have breakfast today".

(1.9) Jane did not have breakfast.

Relevance theorists oppose the view advocated by Levinson that default implicatures arise automatically as associations with certain meanings as part of the grammar, independent of a specific context. Instead, they assume that a stereotypical scenario, that is a minimal, though specific, default context, is always evoked.

1.1.2.4 Primary and Secondary Pragmatic Processes

A different view on the distinction of explicatures and implicatures takes Recanati (2003, 2004). In Recanati's account, explicatures are the output of a system of primary pragmatic processes and serve as input to a system of secondary pragmatic processes, which in turn result in implicatures.

Primary pragmatic processes contribute to the determination of "what is said", operate locally, i.e. at a sub-locutionary level, and are automatic and subconscious processes. They are, in contrast to Carston's conception of explicatures, independent from the speaker's intention. Recanati mentions two main types of primary processes.

First, "saturation" includes the resolution of indexical expressions. This is an inherently pragmatic process: it relies on the speaker's meaning and involves extralinguistic context, but contributes to the determination of the semantic, i.e. truth-conditional, content of an utterance. Saturation is mandatory: indexical expressions must be resolved, if not, an utterance cannot be interpreted.

Second, "modulation" includes meaning shifts such as the different readings of "newspaper" mentioned above in section 1.1.1, as well as Nunberg (1995)'s predicate transfer. In Nunberg's famous example (1.10) , the meaning of "the ham sandwich" must be shifted to the person who ordered the sandwich.

(1.10) The ham sandwich left without paying.

Modulation shifts the meaning of an expression to a pragmatically derived meaning which, nevertheless, contributes to semantic composition. Free enrichment falls under this category, as well.

Secondary pragmatic processes underlie the derivation of implicatures in Grice's sense. They are post-propositional and consciously inferred from the speaker's saying what she says. They involve complex inferencing and rely on reasoning about the intentions of communication participants.

Recanati emphasizes that his primary processes are associative and thus not inferential. This amounts to saying that explicatures are not pragmatic inferences. Although there is clearly a difference between fast, unconscious associations and complex series of conscious conclusions, I will use the term "pragmatic inference" for all types of enrichments involved in determining the conveyed meaning of an utterance beyond purely linguistically encoded information, hence both types of Recanati's pragmatic processes.

1.1.2.5 Abductive Inferences

In addition to the philosophical positions reviewed so far, I would like to mention yet another view on pragmatic inferences advocated in more formally oriented theories. Formal accounts of pragmatic enrichment rely on some form of nonmonotonic reasoning (see section 1.2), although the exact nature of assumed inference mechanisms can be conceived in several distinct ways. A theory of utterance interpretation that assumes only one uniform inference system was proposed by Hobbs et al. (1993). In this approach, sometimes called "local pragmatics", *abductive inference* is seen as the form of human reasoning in general and of language understanding in particular. In general, an abductive explanation for an observation is provided if the observation can be deductively justified on the basis of background knowledge together with possible additional assumptions.

Pragmatic enrichment in this account is the byproduct of finding the best explanation for the grammatically determined meaning of an utterance. Since usually many explanations can be provided, the question is, similar as in Relevance Theory, to find the most effective interpretation. The main advantage of interpretation as abduction is that it provides a formal mechanism for utterance interpretation. However, a drawback is that Hobbs' meaning representation does not distinguish between different knowledge sources, leading to difficulties in determining for each particular utterance which is the relevant contextual information source for its interpretation.

Bearing on this critique of Hobbs' approach, another theory of discourse interpretation, SDRT (segmented discourse representation theory, Asher and Lascarides, 2003), assumes separate logics for constructing and representing enriched semantic forms, as well as for reasoning with different knowledge sources. In section 1.2, the logical foundations of these theories will be laid out.

1.1.2.6 Conclusion

Let us sum up this brief overview of research on the semantics/pragmatics interface. It is important to note that the different types of pragmatic inferences assumed by different theories are not complementary but overlapping. For instance, what Levinson calls GCI covers to a great deal the same phenomena which Carston would subsume under explicature. The major differences are due to distinct theoretical assumptions. If we take for granted that an underspecified grammatically determined meaning is successively enriched by pragmatic inferences it seems useful to distinguish several layers of meaning with differing grades of meaning specification.

The exact number of layers, as well as the manner of their interaction, is still subject of controversial discussions. Bierwisch's two-level semantics focusses on the differentiation between two layers of meaning, SF and CS[5]. Dölling makes the layers more explicit, distinguishing at least SF, PFS, and CS.

Grice assumed two layers of meaning, distinguishing between "what is said" and "what is meant", thus assuming conversational implicatures as one kind of inference mediating between these two layers. Levinson distinguishes between the layer of expression meaning, the layer of utterance type, and the layer of utterance token, telling apart two kinds of pragmatic inferences, GCI and PCI. Carston differentiates two types of inferences, explicature and implicature. Recanati assumes a distinction between primary and secondary processes, where only the latter are mediated by inferences. In difference to the other approaches, Hobbs assumes only one inferential mechanism, namely abduction, for all kinds of pragmatic inferences.

In sum, many researchers, even though working in different frameworks, make a distinction of two basic kinds of pragmatic inferences. A tentative classification is the following. On the one hand, there are (i) salient, automatic enrichments, which take place without any specific context, i.e. they are either completely context-independent or take place assuming a stereotypical minimal default context. Often, they are unconscious and do not require much effort for the hearer. On the other hand, there are (ii) effortful, possibly com-

5 In the original conception of his theory, Bierwisch assumed three layers of meaning, which can be dubbed as sentence meaning, utterance meaning, and communicative sense. However, in the later perception of his theory, the emphasis was laid on the distinction of the first two layers, which correspond to SF and CS.

plex inferences, which take place if a specific context is given. Often, they are conscious and require the consideration of the intentions of the speaker. We have to be cautious with this classification because there is no clear correspondence between the distinct notions assumed by distinct theories and these two classes. Roughly, Levinson's GCIs, Carston's explicatures (as far as they concern the specification of default values for free variables), and Recanati's primary processes fall into the first category, while Grice's conversational implicatures, Levinson's PCIs, Carston's explicatures (if they concern systematic polysemy and meaning shifts) and her implicatures, and Recanati's secondary processes fall into the second category.

It would be nice if the picture were that simple. Facing the still ongoing discussion between these approaches, it seems that we cannot rely on a clear two-way distinction of pragmatic inferences, at least at the moment. Instead, in view of the broad range of proposals made in the literature, let us examine which properties of pragmatic inferences can be made out without bearing on theory-specific assumptions.

1.1.3 Properties of Pragmatic Inferences

A first, obvious property of pragmatic inferences is their *context-dependence*. We have seen that there is a vast range of phenomena where some form of inferencing involving linguistic and non-linguistic contextual knowledge takes place. Important questions are, however, what facets of context contribute to what kind of pragmatic inferences, and to what extent an assumed context must be specific in order to allow drawing these inferences. In chapter 2, I will discuss several aspects of context in discourse.

A second fundamental property common to all types of pragmatic inferences can be made out: they are cancellable or *defeasible*, in contrast to semantic entailments. Pragmatic inferences can be cancelled when inconsistencies arise. Once drawn conclusions sometimes have to be withdrawn in view of contrary information eventually showing up later in a discourse. Many contextual factors can cause the cancellation of already drawn inferences. A typical example of the cancellation of a pragmatic inference is (1.11). The first sentence has the scalar implicature that John did not eat all of the cookies. This inference is explicitly cancelled in the second sentence.

(1.11) John ate some of the cookies. In fact, he ate all of them.

Although it could be argued that explicit cancellability may only hold for Q-based implicatures and PCIs, it is certainly true for all kinds of pragmatic inferences that they do not constitute necessary conclusions (in contrast to semantic entailments), i.e. in general there is no unique way of drawing a pragmatic inference. Take, for instance, the example for an explicature from above, repeated here:

(1.12) Jane did not have breakfast.

In this case, in absence of additional information, one would pragmatically infer the explicature that Jane did not have breakfast *today*. However, if we knew that the discourse was about Jane's experience in prison in the last week, then one would draw a different explicature, namely that she did not have breakfast *in the last week*. Thus, in order to be able to determine the truth-conditional content of an utterance, one of its possible explicatures has to be made, although each of them, regarded in isolation, is defeasible.

All proposals of pragmatic inferences discussed above account for the defeasibility of pragmatic inferences in one or another way. In the Gricean system of maxims, explicit and conscious defeasibility is achieved by the possibility to flout the maxims. Implicit defeasibility is built into the system by the fact that conversational implicatures are conceived, in the first place, as different from semantic entailment, viz. classical implications.

In a neo-Gricean framework, a detailed picture of the process of cancellation of pragmatic inferences is described by Huang (1994), who states that generalized conversational implicatures can be overridden when they turn out to be inconsistent with one of the following factors (as cited by Blackwell, 2003, p. 39ff): (i) background assumptions or world knowledge, (ii) the presumed speaker's intention according to the assumed state of mutual knowledge (Grice's "non-natural" meaning), (iii), semantic entailments, and (iv) what is relevant or salient.

In Relevance Theory, defeasibility is achieved by assuming that Relevance can change as a discourse proceeds.

In the abductive approach and in other formal accounts, defeasibility is a central property of the underlying inference logic. In section 1.2, I will review formal approaches to defeasible reasoning in some detail.

A third property often cited for distinguishing pragmatic inferences from semantic entailments is their *non-detachability*. Conversational implicatures in Gricean and Neo-Gricean theories depend on the explicit content of utterances, and not on their linguistic form. Thus, utterances with the same explicit content will have the same implicatures. Similarly, Relevance-theoretic explicatures cannot be detached from the utterance from which they are derived.

However, this property is not universal to all pragmatic inferences: some M-based implicatures seem to be an exception. The following examples, repeated from above, have the same explicit content but different implicatures.

(1.13) a. René fell in love and left the town.

b. René left the town and fell in love.

Bach (1994) coined the term *impliciture* for inferences that are cancellable and detachable, in contrast to implicatures, which are cancellable and non-detachable. Bach's implicitures correspond, more or less, to Carston's explicatures and Recanati's primary pragmatic processes.

A fourth property attributed to pragmatic inferences is their *calculability*. This property has to be taken cautiously since in the pragmatic literature it usually means calculability in terms of Grice's framework of conversational implicatures in the sense that the addressee can infer the intended implicature on the assumption that the speaker is obeying the cooperative principle and the conversational maxims. However, as already noted in the previous section, Grice's principles have proven to be difficult to be formalized in an exact manner. It is one of the central aims of this thesis to provide a formally more exact way of calculating pragmatic inferences.

To summarize this short discussion of properties of pragmatic inferences, we can observe that two essential properties uncontroversially hold for all types of pragmatic inferences:

- context-dependence and
- defeasibility.

The first property is examined in more detail in chapter 2, whereas the second property is topic of the next section.

1.2 Formal Approaches to Defeasible Reasoning

The main characteristic of pragmatic inferences is their defeasible or *nonmonotonic* character. Reasoning with information that may be changing from one moment to another often allows drawing conclusions only by assuming that things are normal unless not otherwise stated. These default conclusions can be cancelled when information upcoming afterwards indicates that the drawn inference conflicts with additional knowledge. However, in most current theories of formal semantics, starting with Montague (1973), one or another form of classical logic is used to formalize the semantic content. Classical logic is monotonic: an inference once drawn cannot be undone.

Monotonicity

If $A \vdash q$ then $A \cup \{q\} \vdash q$.

With this property, classical logic is not suitable for expressing default inferences. Another form of reasoning must be found in order to deal with defeasible inferences.

Most of the influential formal accounts for representing default inferences were developed in the context of research on Artificial Intelligence independently from each other in the 1980s. The starting point was the aim to model intelligent agents which must be able do make reasonable decisions on the basis of incomplete knowledge, to handle rules with exceptions and to deal with inconsistent information. The search for a logic supporting these features led, almost at the same time, to a series of different accounts[6].

A default is a conclusion that is normally valid. Its justification can only be cancelled by an explicit fact against the inference. A first general formulation of default reasoning was developed by Reiter (1980). The basic idea is the following: a set of default rules is available, and as many conclusions of applicable defaults are added to a set of premises as consistently possible. For formalization, a fixed-point construction is used. A similar construction is also used in modal nonmonotonic logics like the auto-epistemic logic of Moore (1985). These logics, also called consistence-based logics, are presented in section 1.2.1.

Another family of nonmonotonic logics is based on the selection of a preferred set of models of a theory. The method of Circumscription (McCarthy, 1980) is a model-theoretic formalism expressed by higher-order formulae, which minimize the extensions of a set of predicates. Defaults are expressed by abnormality predicates. We will introduce this approach in section 1.2.2. Other logics of model preference are conditional logics like Commonsense Entailment (section 1.2.3), which bear on the concept of possible worlds and a relation of "proximity" between them.

Section 1.2.4 deals with abductive reasoning, another way of accounting for defeasible inferences. An abductive inference is the conclusion from an observation to a possible explanation justifying the observation. The quality of explanations can change in function of the available information. Although this method of reasoning has a long history, its

6 For a broader survey on nonmonotonic paradigms, see Brewka et al. (1997). A good overview on possible uses in natural language interpretation is given in Thomason (1997).

formal and computational properties, as well as its relation to other formalisms, were brought into light only recently.

1.2.1 Default Logic

In Reiter (1980)'s account, the perhaps best-known representation of defaults, defaults are represented by non-classical inference rules, as opposed to classical inference rules like *modus ponens*.

Default Theory A Default Theory is a pair $\langle D, W\rangle$ with

W, a set of classical formulae (propositions), and

D, a set of default rules of the form $\frac{A : B}{C}$ (or $A : B/C$).

In this definition, W represents the knowledge that is reliable, and the defaults in D consist of preconditions A, consistence assumptions B, and consequence C, all of them (sets of) classical propositions. The intuitive interpretation of such a default rule is: if A is derivable, and $\neg B$ cannot be derived, then derive C.

The set of acceptable beliefs that are induced by a given default theory $\langle D, W\rangle$ is called *Extension*. It has the following properties: (i) it contains the reliable knowledge W; (ii) it is closed in the sense of classical logic; (iii) all applicable defaults have been applied; and (iv) it does not contain any proposition that is not derivable from W together with the consequences of applicable defaults D.

A problem with this definition is that the provability of a proposition depends on the non-provability of other propositions. A proof is a series of steps in which the rules have been applied correctly. But in order to know whether a rule has been applied correctly, one has to know already what is provable in order to be able to carry out the consistence check. To avoid this circle, a fixed point can be constructed to define Extensions of a default theory[7].

Extension Let $\langle D, W\rangle$ be a default theory, and S a set of propositions. Γ is an operator such that $\Gamma(S)$ is the smallest set for which the following holds:

1. $W \subseteq \Gamma(S)$
2. $Th(\Gamma(S)) = \Gamma(S)$[8]
3. if $A : B_1, ..., B_n/C \in D$ and $A \in \Gamma(S)$ and $\neg B_i \notin S$ $(1 \leqslant i \leqslant n)$, then $C \in S$.

E is *Extension* of $\langle D, W\rangle$ iff E is fixed point of Γ, i.e. if holds $\Gamma(E) = E$.[9]

A simple example for a default rule is: if birds normally fly, and Tweety is a bird, then it can be concluded that Tweety flies. In Reiter's Default Logic, this rule can be represented as follows:

7 Definitions are taken from Brewka et al. (1997).

8 $Th(A) = \{p | A \vdash p\}$

9 There is an alternative, nearly inductive, but not constructive definition: Be E a set of propositions. A set of propositions can be defined as follows:

1. $E_0 = W$, and for $i \geqslant 0$
2. $E_{i+1} = Th(E_i) \cup \{C | A : B_1, ..., B_n/C \in D, \quad A \in E_i, \quad \neg B_i \notin E\}$

E is *Extension* of $\langle D, W\rangle$ iff $E = \bigcup_{i=0}^{\infty} E_i$.

$$(1.14)\quad D = \left\{ \frac{Bird(x) : Flies(x)}{Flies(x)} \right\},\ W = \{Bird(Tweety)\}$$

The only extension of this default theory is $Th(W \cup \{Flies(Tweety)\})$.

Another inference pattern, known as the *Nixon Diamond*, deals with the case of conflicts between default rules. This happens when several competing defaults with mutually independent preconditions can be applied at the same time. In this case, the conflict cannot be resolved.

$$(1.15)\quad D = \left\{ \begin{array}{c} \frac{Quaker(x) : Pacifist(x)}{Pacifist(x)} \\ \\ \frac{Republican(x) : \neg Pacifist(x)}{\neg Pacifist(x)} \end{array} \right\},\ W = \left\{ \begin{array}{l} Republican(Nixon), \\ Quaker(Nixon) \end{array} \right\}$$

For the theory in (1.15), there are two extensions: $Th(W \cup \{Pacifist(Nixon)\})$ and $Th(W \cup \{\neg Pacifist(Nixon)\})$. If Nixon was not only Quaker but also Republican, then it is not decidable whether he was a pacifist or not.

There are several possible strategies for deciding which extension should be chosen. Skeptical and credulous reasoning can be distinguished. *Skeptical Reasoning* accepts only what is true in *all* extensions. For the Nixon diamond, it means: Nixon was neither pacifist nor Republican. *Credulous Reasoning* chooses *one* extension following certain principles and accepts the conclusions of this extension. In the Nixon diamond, this kind of reasoning requires a decision supporting or rejecting the belief that Nixon was a pacifist, e.g. on the basis of Nixon's speeches and actions.

To represent preferences between defaults, the preconditions of a preferred rule can be included in the other rules as an additional consistence condition.

$$(1.16)\quad D = \left\{ \begin{array}{c} \frac{Bird(x) : Flies(x) \wedge \neg Penguin(x)}{Flies(x)} \\ \\ \frac{Penguin(x) : \neg Flies(x)}{\neg Flies(x)} \end{array} \right\},\ W = \left\{ \begin{array}{l} Bird(Tweety), \\ Penguin(Tweety) \end{array} \right\}$$

The theory (1.16) has a unique extension $Th(W \cup \{\neg Flies(Tweety)\})$. The second rule is applied. When preferences are defined following the criterion of specifity, we speak of the *Penguin Principle* [10].

Default logic is more powerful than other formalisms like e.g. Circumscription. In order to reduce the expressive power, only *semi-normal* defaults of the form $A : B \wedge C/C$ can be applied. With them, the representation of preferences between defaults is still possible. More restricted are *normal* defaults $A : B/B$, which guarantee the existence of extensions.

This inference scheme allows the expression of pragmatic inferences in a direct way. Levinson (2000) sketches a possible formalization of scalar implicatures using Default Logic. For example, (1.17b) is an implicature of (1.17a), based on $\langle all, some \rangle$ as an instance of a Horn-scale[11] $\langle STRONG, WEAK \rangle$. This implicature can be expressed by the default rule in (1.18), where $\alpha(w)$ means that the expression w is part of the sentence frame α.

(1.17) a. Some students sleep.

10 This principle holds in nonmonotonic conditional logics which is the basis for SDRT (Lascarides and Asher (1991); Asher and Lascarides (2003)), see example (1.29) below.

11 see section 1.1.2.2 on page 19.

b. Not all students sleep.

(1.18) $$\frac{\alpha(some) : \alpha(not\ \ all)}{\alpha(not\ \ all)}$$

Default logic is more expressive than other approaches. It subsumes the semantics of logic programming languages. For example, Prolog[12] has a built-in default operator, "negation as failure". While its semantics is defined procedurally in standard dialects of Prolog, there are other, model-theoretically defined logic programming paradigms. For instance, stable models for extended logic programs, *Answer Sets*, are a special case of Reiter's Default Logic. This paradigm is characterized by the existence of two kinds of negation: Classical negation ($\neg$) and default negation `not b` (b is not believed). A rule $\texttt{c} :- \texttt{a}_1, ..., \texttt{a}_n, \texttt{ not b}_1, ..., \texttt{not b}_m$ corresponds to the default $a_1 \wedge ... \wedge a_n : \neg b_1, ..., \neg b_m / c$. For more on Answer Set programming, see e.g. Brewka and Eiter (1999).

1.2.2 Circumscription

Another kind of nonmonotonic reasoning goes back to McCarthy (1980)[13]. It is an inference pattern that is based on a simple idea: there are properties which normally hold only for a few entities, e.g. among all persons, there are only a few color-blind persons. Thus, the expression $\neg colorblind(John)$ should be derivable as long as $colorblind(John)$ is not provable.

To account for this idea, a certain notion of minimality of models is defined. The extension of a given predicate P is *minimized*, i.e. the number of entities for which P is true. Additionally, a preference relation $\leq_P$ on the models of a set of propositions is defined. In preferred models, the chosen predicate has an extension as small as possible. Derivability then is not defined as validity in all models, but as validity in the minimal models with respect to this preference relation. Semantically, we are only concerned with certain models, namely with those which have a minimal extension of P. The preference relation $\leq_P$ is defined as follows:

(1.19) $M_1 \leq_P M_2$ iff

M_1 and M_2 coincide in the extensions of all predicates but P, and

the interpretations of P in M_1 are a subset of those of P in M_2.

(1.20) M is *minimal* wrt. $\leq_P$ iff there is no other model M' such that $M' \leq_P M$.

(1.21) A proposition ϕ ist *derivable* from a set of propositions Φ iff ϕ is true in all minimal models of Φ.

Syntactically, Circumscription consists of transformations of logical formulae. To a set of propositions Φ, a second order formula is added. This formula eliminates all models of Φ in which the extension of the predicate to be minimized is not minimal. Note that quantification over predicates is necessary to express minimality, yielding a second order formula. For the sake of better readability, I will use the following set notation:

12 A recent popular introduction to Prolog is Blackburn et al. (2006).
13 The description in this section relies on Brewka et al. (1997).

(1.22) a. $P = Q$ iff $\forall x[P(x) \leftrightarrow Q(x)]$

b. $P \subseteq Q$ iff $\forall x[P(x) \rightarrow Q(x)]$

c. $P \subset Q$ iff $P \subseteq Q \wedge \neg(Q \subseteq P)$

Now the Circumscription of a predicate in a formula can be defined as follows:

Circumscription Let ϕ be a proposition and P a predicate symbol. The *Circumscription* of P in ϕ, $Circ[\phi, P]$ is the formula

(1.23) a. $\phi \wedge \neg\exists\mathfrak{Q}[\phi[P/\mathfrak{Q}] \wedge \mathfrak{Q} \subset P]$, or equivalently,
b. $\phi \wedge \forall\mathfrak{Q}[(\phi[P/\mathfrak{Q}] \wedge \mathfrak{Q} \subseteq P) \rightarrow (P \subseteq \mathfrak{Q})]$

Here, $\phi[P/\mathfrak{Q}]$ is the formula that results from replacing in ϕ every instance of P by the predicate variable $\mathfrak{Q}$.

A simple example: be $\phi = colorblind(John) \wedge colorblind(Tom)$ and $P = colorblind$. In accordance with the schema (1.23b), the Circumscription of P in ϕ can be equated with the following formula:

(1.24) $Circ[\phi, P] = \phi \wedge \forall\mathfrak{Q}[(\mathfrak{Q}(John) \wedge \mathfrak{Q}(Tom) \wedge \mathfrak{Q} \subseteq P) \rightarrow (P \subseteq \mathfrak{Q})]$

The substitution of $\mathfrak{Q}$ by $Q(x) := [x = John \vee x = Tom]$ yields for $\phi[P/\mathfrak{Q}]$ the true expression $(John = John \vee John = Tom) \wedge (Tom = John \vee Tom = Tom)$. With that we get as Circumscription the following formula, according to the definition of $\subseteq$:

(1.25) $Circ[\phi, P] = colorblind(John) \wedge colorblind(Tom) \wedge$
$[\forall x[(x = John \vee x = Tom) \rightarrow colorblind(x)]$
$\rightarrow \forall x[colorblind(x) \rightarrow (x = John \vee x = Tom)]]$

Hence, Tom and John are the only color-blind persons.

This technique is exploited by introducing a predicate *ab* ("abnormal"), which is minimized by Circumscription. For instance, the assertion that birds normally fly, can be expressed by the proposition (1.26).

(1.26) $\forall x[bird(x) \wedge \neg ab_1(x) \rightarrow flies(x)]$

The intuitive meaning of this formula is: Birds which are not abnormal can fly. Since entities can be normal in one respect and abnormal in another, we need a distinct abnormality predicate for each default. Circumscription is then used to minimize the extensions of *ab*-predicates.

For more information on the logical foundations of Circumscription and some useful extensions, see Lifschitz (1989). For us, this approach to nonmonotonic reasoning is very interesting because *minimal models* are the key notion of Circumscription. We will explore this in more detail in chapter 5 in section 5.2.

1.2.3 Commonsense Entailment

A family of accounts to nonmonotonic reasoning extends classical logic by a special inference operator '>'. Intuitively, $A > B$ means: if A then normally B. Its semantics is based on possible worlds. The expression $A > B$ is true in a possible world w iff in all worlds "close" to w in which A is true, B is true, too. This kind of nonmonotonic conditional logics was originally developed by Stalnaker (1968) and Lewis (1973) in order to formalize contrafactual conditionals, e.g.

(1.27) If Nixon had not been president, Watergate would never had occurred.

If we want to interpret this utterance, we have to look at all worlds which are similar to our world and in which Nixon was not president, and examine these worlds if Watergate had occurred in them. If not, then the conditional expression is true.

Among the existing conditional logics, I want to focus on a logic called *Commonsense Entailment* (Asher and Morreau, 1991), which is most interesting for linguistic applications. This logic contains complex operators for constructing extensions of possible worlds, similar to the fixed point constructions of Default Logic. This theory consists of two stages: (i) a monotonic modal logic is constructed to which (ii) a nonmonotonic consequence relation is added. The basic idea is to maximize the number of inferences drawn from the defaults while maintaining consistence with respect to the monotonic modal logic. It is assumed that as many things as possible are as normal as possible without causing inconsistencies.

In the modal logic, the operator '>' is interpreted by means of a function $*$ from possible worlds and sets of possible worlds into sets of possible worlds. The operator has the following model-theoretic interpretation with respect to a model M, a world w, and a variable assignment g (a function which assigns an entity to each entity variable in the model):

(1.28) $M, w, g \models A > B$ iff $*(w, [\![A]\!]_{M,g}) \subseteq [\![B]\!]_{M,g}$

The expression $A > B$ is true in a world w iff $*(w, [\![A]\!]_{M,g}) \subseteq [\![B]\!]_{M,g}$ is true. $*(w, p)$ returns all worlds which are normal with respect to p, i.e. all worlds in which p holds with all consequences typical for w.

This operator gives rise to a nonmonotonic consequence relation $\mid\!\sim$ (to be distinguished from the classical monotonic consequence relation $\vdash$), which has some interesting properties. First of all, *modus ponens* is cancellable, in contrast to its equivalent in classical logic:

Defeasible Modus Ponens

$A, A > B \mid\!\sim B$

$A, A > B, \neg B \not\mid\!\sim B$

If it is known that birds normally fly ($A > B$) and that Tweety is a bird (A), then we can conclude that Tweety flies (B). If there is additional information that Tweety cannot fly ($\neg B$), then the conclusion B cannot be drawn.

Another basic property is the Specifity Principle or Penguin Principle (cf. (1.16) above): If more than one defaults are applicable, and there is a specifity relation between them, the most specific default rule applies (1.29):

Penguin Principle

If $\vdash A \rightarrow C$ then $A > \neg B, C > B, A \mid\sim \neg B$.

(1.29) All penguins are birds.
Birds fly.
Penguins don't fly.
Tweety is a Penguin.
Hence: Tweety doesn't fly.

Finally, Commonsense Entailment shares a property with Default Logic, the Nixon Diamond (cf. (1.15)): If two defaults compete whose preconditions do not stand in a logical entailment relation, then the conflict cannot be resolved.

Nixon Diamond

$A > B, C > \neg B, C, A \not\mid\sim B$
$A > B, C > \neg B, C, A \not\mid\sim \neg B$

(1.30) Quakers are pacifists.
Republicans are not pacifists.
Nixon is Quaker.
Nixon is Republican.
Hence: It cannot be decided whether Nixon is pacifist or not.

This logic is fundamental to SDRT (Asher and Lascarides, 2003). There, it is the core of the "glue logic". We will come back to it in chapter 5 in section 5.3.

1.2.4 Abductive Reasoning

The philosopher Charles Sanders Peirce (Peirce, 1903) distinguished three basic kinds of reasoning: deduction, induction and abduction. First, classical logic can prove by deduction that something *must be* the case. As shown in the inference schema (1.31), a result can be derived from a particular case applied to a given rule. Peirce illustrated the different inferences with the help of a bag full of beans.

(1.31)
$$\frac{\forall x[p(x) \rightarrow q(x)] \quad p(A)}{q(A)}$$

(1.32) a. Rule: All beans in the bag are white.

b. Case: These beans are from this bag.

c. Result: These beans are white.

Induction shows that something *is actually* the case. From a series of instances of p and q, a general law is derived (1.33). Learning is based mostly on induction, and for the empirical sciences this kind of reasoning is essential.

(1.33)
$$\frac{p(A_1), ..., p(A_n) \quad q(A_1), ..., q(A_n)}{\forall x[p(x) \rightarrow q(x)]}$$

(1.34) a. Case: These beans are from this bag.

b. Result: These beans are white.

c. Rule: All beans in the bag are white.

Abduction finally suspects that something *might be* the case. A given result p could possibly brought up by a rule applied to a case q (1.35).

$$\begin{array}{l} q(A) \\ \forall x[p(x) \rightarrow q(x)] \\ \hline p(A) \end{array} \tag{1.35}$$

(1.36) a. Result: These beans are white.

b. Rule: All beans in the bag are white.

c. Case: These beans are from this bag.

In the words of Peirce, the abductive principle can be expressed as follows:

> "The surprising fact, C, is observed;
> But if A were true, C would be a matter of course,
> Hence, there is reason to suspect that A is true. " (Peirce, 1934, 5.161-212)

Eco (1985) classifies various types of abduction. In *overcoded abduction*, a case is derived from a result by a kind of automatic selection of rules or codes. Examples include the recognition of an animal's scent or a symptom of a disease.

Undercoded abduction is inference to a law. From a set of possible laws in a knowledge base, the most plausible law is chosen, depending on the situation. Scientific research mostly follows this paradigm. Scientists construct abstract theories capable of deriving or predicting observable events. An example is Kepler's discovery of the elliptical course of Mars. In natural language comprehension, an ambiguous utterance requires a process of disambiguation involving an inference to the most plausible explanation of the utterance.

In *creative abduction*, a law is invented *ex novo*. This inference goes beyond preceding and already coded experiences. Revolutionary discoveries leading to changes in an established scientific paradigm usually involve the introduction of previously unknown rules and laws. More examples of this pattern include the solution of puzzling detective novels and the interpretation of poetic texts.

The inference pattern of abductive reasoning is not valid in the sense of classical logic: normally, there are many possible explanations for an observation. When me make the observation that the lawn is wet in the morning, there can be several possible hypotheses explaining the observation: possibly it has rained before, or the sprinkler has been in action. Drawn conclusions are just assumptions, and possibly they have to be withdrawn later if new facts to the contrary become known. In this sense, this kind of reasoning is nonmonotonic. If the street next to the lawn is not wet, we cannot hold the assumption that it has rained before. The following scheme characterizes abductive reasoning.

Abductive Explanation Given a set F of facts (background knowledge), a set H of hypotheses, and a set O of observations to be explained, then a set of propositions E is an *Abductive Explanation* if it fulfils the following criterions:

1. $E \subseteq H$
2. $F \cup E \vdash O$
3. $F \cup E$ is consistent.

An observation O cannot be explained by F alone, but it can be explained by a consistent set extending the facts F by the hypotheses E. For a simple example, suppose that the following facts, observations, and hypotheses are given.

$$(1.37)\quad F = \left\{ \begin{array}{l} rain \rightarrow wet\_lawn \\ sprinkler \rightarrow wet\_lawn \\ \neg(sun \wedge rain) \end{array} \right\},\ H = \{rain, sprinkler\},\ O = \{wet\_lawn\}$$

There are three explanations for the observation O:

(1.38) a. $\{rain\} \cup F \vdash wet\_lawn$

b. $\{sprinkler\} \cup F \vdash wet\_lawn$

c. $\{rain, sprinkler\} \cup F \vdash wet\_lawn$

If the proposition *sun* is added to the theory, only explanation (b) remains, because *rain* is inconsistent with *sun*.

This picture is attractive when we only have a few possible explanations. But normally, the search space for abductive hypotheses is enormous. Methods of constraining it use a metrics on the quality of explanations. *Cost-based abduction* (Charniak and Shimony, 1994) prefers proofs of observations with minimal costs based on a probabilistic Bayesian network. *Weighted abduction* (Stickel, 1991; Hobbs et al., 1993) has a model-theoretic semantics basing on model preference. In chapter 5 in section 5.1, I will present and discuss the latter method and its application to natural language interpretation.

1.2.5 Conclusion

The above presented formalizations of nonmonotonic reasoning represent defaults and explanations by virtue of distinct formal means. There is no uniform approach, although representations can be transferred, at least partly, from one theory to another. The approaches deviate from classical logic in distinct ways in order to obtain nonmonoticity. Reiter's Default Logic uses non-classical inference rules and needs a fixed point construction in its formal definition. Circumscription makes use of a higher order logic and abnormality predicates. Conditional logics add a special consequence relation to a modal logic. Abductive reasoning uses the inversion of classical logic's *modus ponens*.

If there is more than one explanation for a fact, or if more than one default rule can be applied, then we have to select the appropriate rule or explanation. In Default Logic, preferences between defaults are expressed by adding the consequence of the preferred rule to the consistence conditions in the other rules. Preferences in circumscriptive theories can be handled by assuming a notion of minimality of models, which I will introduce in section 5.2. In conditional logics, there is the specifity principle, which can be applied if there are entailment relations between antecedents of conflicting rules. The question of choosing among competing abductive explanations will be addressed in section 5.1.

Despite the differences between the presented approaches, there are some commonalities. Crucially, any kind of nonmonotonic reasoning needs a consistency check. In Default Logic, it is inherent in the fixed-point constructions. In Circumscription, it is done by the higher order logic, and in Commonsense Entailment, it is part of the notion of normality of worlds. Also abduction requires abducing a consistent set of assumptions.

The nonmonotonic logics discussed here are not decidable, that is there is no correct and complete universal proof procedure. If they are to be implemented, the logics have to be restricted to computable subsets, or approximations have to be found.

To conclude, it seems that it does not matter which of the formalizations is used in a theory of discourse interpretation. Consistency checks are not avoidable, and most defaults can be expressed in any of the formalisms. Importantly, one has to account for nonmonoticity in order to be able to deal with cancellable conclusions. In chapter 5, I will discuss some theories of discourse interpretation which make use of the logics discussed in this section.

1.3 Pragmatic Inferences Beyond the Sentence Level

So far, we have examined pragmatic inferences with respect to isolated utterances and sentences. However, sentences usually are not uttered in isolation, rather they are embedded into larger units, viz. a text or a discourse. In this section, the discussion of pragmatic inferences is extended onto the discourse level. After some introductory remarks on the nature of texts and discourses, their main characteristics are elaborated: cohesion and coherence. It will be examined which inferences have to be drawn in order to reach these criteria for textuality and to what extent they share properties with the inferences examined so far. The section will be closed by presenting a general model of discourse structure and interpretation subsuming most of the factors that are relevant for understanding texts.

1.3.1 Text and Discourse

Generally, a *text* or a *discourse* is a sequence of natural language utterances. It can be a spoken or written monologue, or also involve more than one speaker, like in dialogues. As can be seen from this definition, I do not distinguish between typically written monologic texts and typically spoken dialogic discourses, and unless otherwise indicated, the two terms can be interchanged. The underlying reason is that I am convinced that spontaneous spoken discourses are the most natural setting for the occurance of natural language. As a consequence, a theory of discourse interpretation cannot be restricted to written monologic texts only. However, most of the research results on the interpretation of written texts can be naturally extended to the understanding of discourses.

The term *text* has its origin in the latin word “textum” which means “weave” or “fabric”. This association points to two main characteristics of texts: parts of a text stand in mutual relationships, and texts have a structure.

The aims of classical text linguistics in the philological tradition (cf. e.g. de Beaugrande and Dressler, 1981; Heinemann and Heinemann, 2002) are to define what is a text, to classify different types of texts, and to examine the communicative functions of texts. Although these are very ambitious goals, text linguistics mostly remains at a descriptive level and seldom makes use of formal methods.

In order to distinguish meaningful texts from arbitrary conglomerations of chunks of natural language strings, text linguistics has established a series of criteria for textuality. As soon as a text does not fulfill one of these criteria, the text is not communicative and considered a non-text. De Beaugrande and Dressler (1981) enumerate seven criteria: cohesion, coherence, inentionality, acceptability, informativity, situationality, and intertextuality. Although this listing is not very structured, it well characterizes the main properties of texts.

For illustration, take the following excerpt from a famous fairytale[14].

(1.39) a. It was the middle of winter, and the snow flakes were falling like feathers from the sky,

b. and a queen sat at her window working, and the embroidery frame was of ebony.

c. And as she worked, gazing at times out on the snow,

d. she pricked her finger,

e. and there fell from it three drops of blood on the snow.

f. And when she saw how bright and red it looked,

g. she said to herself, "Oh that I had a child as white as snow, as red as blood, and as black as the wood of the embroidery frame!"

h. Not very long after she had a daughter, with a skin as white as snow, lips as red as blood, and hair as black as ebony, and she was named Snow-white.

(The brothers Grimm: Snow-white, KHM 053)

In this text, at first sight we can observe two kinds of relationships which hold the text together: cohesion, exemplified by the relation between "a queen" in (b) and "she" in (c), and coherence, indicated, for instance, by the causal relation between (c) and (d). I will turn to each of these properties of texts in sections 1.3.2 and 1.3.3, respectively.

The other criteria for textuality mentioned by de Beaugrande and Dressler (1981) point at important issues which I will also deal with in course of this thesis. Intentionality means that, normally, texts are produced with the purpose to arrive at a goal in a plan of the speaker. The linguistic means to arrive at this goal are cohesive and coherent texts. On the recipient's side, acceptability refers to the attitude of the speaker to expect a coherent and cohesive text. The speaker has to take such attitudes into account in order to arrive at her communicative goals. Informativity means that texts normally convey something new. Crucially, new information must be connected to already known information in order to ensure that a text can be properly understood. Situationality refers to the fact that knowing the situation of an utterance of a text, i.e. place, time, social situation, etc., is often essential for understanding. Finally, intertextuality refers to a common property of texts to make reference to other texts, e.g. in legal texts.

14 The story was collected by the brothers Grimm in the 19th century; the English translation of the originally German story is taken from a parallel corpus of fairytales (http://www.grimmstories.com/language.php?grimm=053&l=en&r=de).

1.3.2 Cohesion

Cohesion emerges from relationships between entities mentioned in the text. The most important cohesive device are *anaphora*. The term anaphor (from Greek ἀναφορά "to carry back") originally refers to a rhetorical figure consisting in emphasizing words by repeating them at the beginnings of adjacent clauses. A famous example of this figure of speech was given by Winston Churchill in his speech to the House of Commons of the British Parliament on June 4, 1940:

(1.40) We shall go on to the end, we shall fight in France, we shall fight on the seas and oceans, we shall fight with growing confidence and growing strength in the air, we shall defend our Island, whatever the cost may be, we shall fight on the beaches, we shall fight on the landing grounds, we shall fight in the fields and in the streets, we shall fight in the hills; we shall never surrender.

In linguistics, however, the term *anaphor* has a more general sense and corresponds to an expression referring to an entity or a concept mentioned before, the *antecedent*. For example, in (1.41), anaphoric relations are marked by assigning the same index to anaphor and antecedent[15].

(1.41) *A queen*$_1$ sat at her$_1$ window working, and the embroidery-frame was of ebony. And as she$_1$ worked, gazing at times out on the snow, she$_1$ pricked her$_1$ *finger*$_2$, and there fell from it$_2$ three drops of blood on the snow.

A recipient of a text containing anaphoric expressions has to find their antecedents in order to correctly interpret the text. This process is called *anaphora resolution*. Although a considerable variety of anaphoric relations can be resolved by means of purely linguistic information encoded in a text, in many cases additional information is necessary in order to establish anaphoric links.

This can be the case for *direct anaphora*, which bear coreference of expressions. For example, "it" in the first sentence of (1.42) refers to the finger, whereas "it" in the second sentence refers to the blood. While a grammatical explanation for the first coreference could be thought of, it does not seem possible to establish the second coreference on purely linguistic grounds.

(1.42) She pricked her finger$_1$, and there fell from it$_1$ three drops of blood$_2$ on the snow. And when she saw how bright and red it$_2$ looked, she said to herself, ...

While these were direct anaphoric relations, also *indirect anaphora* can be frequently found in texts. Such indirect, or *bridging anaphora*, are exemplified in (1.43): the expression "embroidery frame" is indirectly related to the event of "working" (as being the working instrument). The former is referred to as the anaphoric expression and the latter is the bridging *anchor*. In the examples in this thesis, bridging anaphora are typeset in **bold face**, and their anchors are underlined.

(1.43) A queen sat at her window working, and **the embroidery frame** was of ebony.

15 A more thorough discussion of anaphoric phenomena can be found in section 3.1.

In these cases, the relationship is indirect: it is distinct from coreference, it was not explicitly mentioned before, and the hearer has to infer it by further inferences in order to establish the anaphoric link and to make sense of the text.

In sum, anaphora resolution exhibits the main properties of pragmatic inferences. Context-dependency is an intrinsic property of anaphoric relations since they establish a link between an expression and some other entity in the context. Anaphoric relations are established by default. For example, without further context, the pronouns "he" and "him" in the main clause of (1.44b) are interpreted by default as referring to Max and John, respectively.

(1.44) a. Max met John.

b. He asked him if he had finally decided whether he would come to the dinner.

This preference can be explained in terms of the grammatical status of the antecedents. "Max" is the subject of (1.44a) and thus the most probable antecedent of a subsequent pronoun in subject position. Another explanation draws from the salience, or activation, of entities in a discourse. If a discourse is about Max, then he corresponds to a salient entity in the discourse, which is in the focus of attention of discourse partipants. It can be argued that pronouns in subject position are likely to be interpreted as referring to the most salient discourse entity.

However, if the discourse is continued, it may become necessary that the already established anaphoric relation has to be revised. For example, after the continuation (1.44c), "he" and "him" in (1.44b) must be taken as referring to John and Max.

(1.44) c. Max answered no.

As a consequence, from the observation that anaphora resolution is defeasible and context-dependent can be concluded that establishing anaphoric relations is a case of pragmatic inference.

Apart from anaphora, also other lexical and syntactic means contribute to the cohesion of a text. Among these are ellipses, i.e. incomplete expressions which can be completed by already available text material. For example, in (1.45), "did" is a placeholder for "run about after every wild humble-bee".

(1.45) Send our youngest child out with my dinner to-day, she has always been good and obedient, and will stay in the right path, and not run about after every wild humble-bee, as her sisters did. (The brothers Grimm: The Hut in The Forest, KHM 169)

Other cohesive devices are discourse connectors. These can be of more logical nature, e.g. conjunction ("and"), disjunction ("or"), indicate causalities ("because". "although"), or temporal cohesive devices, e.g. "not very long after" in (1.46).

(1.46) A queen sat at her window working [...]. *Not very long after* she had a daughter [...]

These cohesive phenomena will not be in the center of attention in this thesis, although many findings on anaphora also apply to them.

In the course of this thesis, I will repeatedly come back to cohesion and anaphora. Chapter 3 contains a more detailed discussion of anaphoric phenomena and presents some accounts of anaphora resolution. Chapter 6 is devoted to indirect bridging anaphora and their resolution.

1.3.3 Coherence

Coherence is established by virtue of *rhetorical relations*, which mark a relationship between chunks of texts, and not between single entities[16]. For instance, the text (1.39), repeated here as (1.47), can be assigned the structure shown in Fig. 1.1.

(1.47) a. It was the middle of winter, and the snow flakes were falling like feathers from the sky,

b. and a queen sat at her window working, and the embroidery frame was of ebony.

c. And as she worked, gazing at times out on the snow,

d. she pricked her finger,

e. and there fell from it three drops of blood on the snow.

f. And when she saw how bright and red it looked,

g. she said to herself, "Oh that I had a child as white as snow, as red as blood, and as black as the wood of the embroidery frame!"

h. Not very long after she had a daughter, with a skin as white as snow, lips as red as blood, and hair as black as ebony, and she was named Snow-white.

Rhetorical relations can be expressed by cohesive means such as discourse connectors (e.g. "and", "but"), but in many cases they are not explicitly marked. Text recipients always try to establish a rhetorical relation between parts of a text in order to conceive the text as coherent. Hence, rhetorical relations are established by default. They are context-dependent: as a discourse proceeds, once established relations may possibly be retracted. Consider the following example, from Lascarides and Asher (1991).

(1.48) a. Max fell.

b. John pushed him.

Without a specific context, hearers interpret (1.48b) as an explanation for (1.48a). The rhetorical relation supposed to hold is Explanation. However, if the context is further specified by additional information showing up later in the discourse, it may become necessary to retract the already drawn conclusion. If the discourse is continued by (1.48c), then it becomes clear that the sentences (a), (b), and (c) form a narrative sequence bearing a Narration relation.

(1.48) c. Max rolled over the edge of the cliff.

16 Rhetorical relations are central to chapter 4, especially section 4.2.

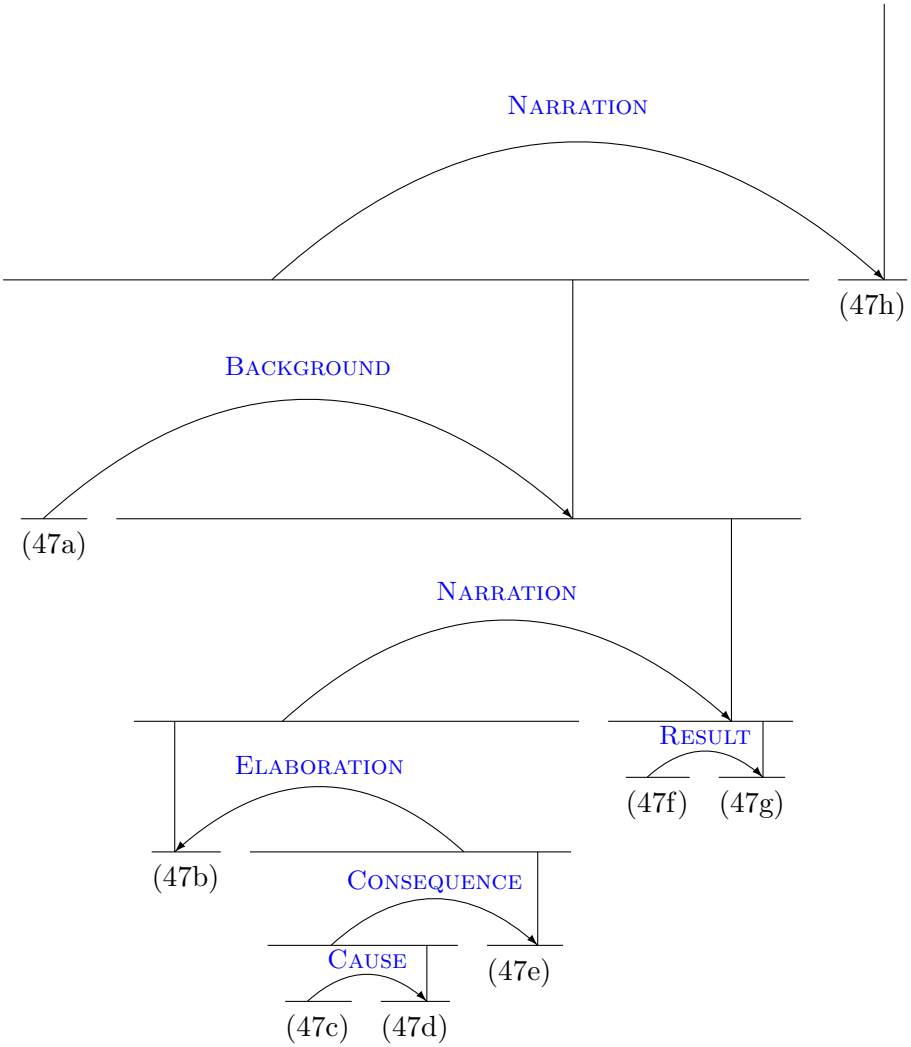


Figure 1.1: Discourse structure of (1.39/1.47)

With regard to the discussion of properties of pragmatic inferences in section 1.1.3, establishing rhetorical relations, being defeasible and context-dependent, clearly is a case of pragmatic inference.

In the literature, there are competing views on the question what kind of knowledge is involved in this process. Two main types of approaches to discourse coherence can be distinguished. There are approaches which try determine the part of utterance meaning that is independent from particular intentions of discourse participants. And there are approaches which take a broader view and take intentions of discourse participants into account. Hobbs (1996) called these accounts of discourse interpretation *informational* and *intentional* accounts. We will discuss both types of approaches in sections 1.3.3.2 and 1.3.3.1, respectively.

1.3.3.1 An Intentional View on Discourse Coherence

In an *intentional* view on discourse coherence, one tries to answer the question what a speaker or writer wants when she produces a discourse or a text, or in other words: why does the speaker want to convey the content of her utterance? Every utterance is based on a certain intention of the speaker, and thus in order to catch all aspects of the meaning of an utterance, intentions of both speaker and addressees have to be taken into account. For an addressee, discourse interpretation consists of recognizing the intentions of the speaker. Coherence relations are assumed to hold between the underlying intentions, or the purposes, of involved utterances.

In intentional accounts, in Hobbs' terms, the interpretation of an utterance is seen as the search for the best explanation why it was made. In natural language pragmatics, intentions have been central to the theories of Grice (1975) and Searle (1969). Intentions play a key role in Grosz and Sidner (1986)'s theory of discourse structure, to which I will turn below. Intentions in discourse interpretation are dealt with in chapter 2.

1.3.3.2 An Informational View on Discourse Coherence

An *informational* view on discourse structure and interpretation is taken by theories which aim to explain form and meaning of natural language discourses while restricting themselves to informational aspects of discourse meaning, excluding any intentional components specific to particular discourse participants. Discourse coherence is defined in terms of relationships between the information contained in successive utterances. The question posed by these approaches is: what is the specific content of a discourse?

In informational accounts, in Hobbs' terms, the interpretation of an utterance is the search for the best explanation for the information explicitly conveyed by the utterance, according to background knowledge which the hearer assumes to be shared. Coherence relations reflect the relationships between the speaker-independent meanings conveyed by successive discourse segments.

Usually, not only purely linguistically encoded information is considered, but also contextual information is taken into account, though restricted to a speaker-independent level of meaning.

It is clear that an informational interpretation is part of the broader intentional perspective. Nevertheless, Hobbs (1996) motivates an informational perspective by making the following points. First, the speaker's plan plays only an indirect role in the interpretation process. The hearer has no direct access to the speaker's intentions. Second, the speaker's intention is not relevant for determining the speaker-independent meaning conveyed by an utterance. Third, hearers have their own intentions which do not need to have anything in common with the speaker's plan, and interpretation consists primarily in relating an utterance to them. Informational accounts to discourse interpretation are subject of chapter 5.

Relating the two views to the discussion of the semantics / pragmatics distinction sketched in section 1.1, the upshot is as follows. The informational view deals with the utterance meaning, i.e. "what is said". The intentional view takes the speaker's menaning into account, i.e. "what is meant". In Bierwisch's terminology, this is the level of communicative sense, that is the third level of meaning.

1.3.4 Looking Ahead

The main devices to make a sequence of natural language utterances a text are cohesion and coherence. Both criteria have to be taken into account in order to explain how texts can be understood. In both cases, pragmatic inferences are involved. The two text phenomena often occur intertwined: establishing coherence depends on cohesion and vice versa. To illustrate the view of these phenomena taken in this thesis, consider Fig. 1.2[17]. Suppose that u_1 and u_2 are consecutive utterance. In the upper part of the picture symbolizing coherence, R_i designate rhetorical relations between the utterances. The lower part symbolizes cohesive links between events e_j and entities x_k, y_l which are mentioned in the utterances. Note that relationships between events (i.e. a special kind of entities) are, in a strict sense, cohesive relations, although they often indicate a coherence relation between the corresponding utterances.

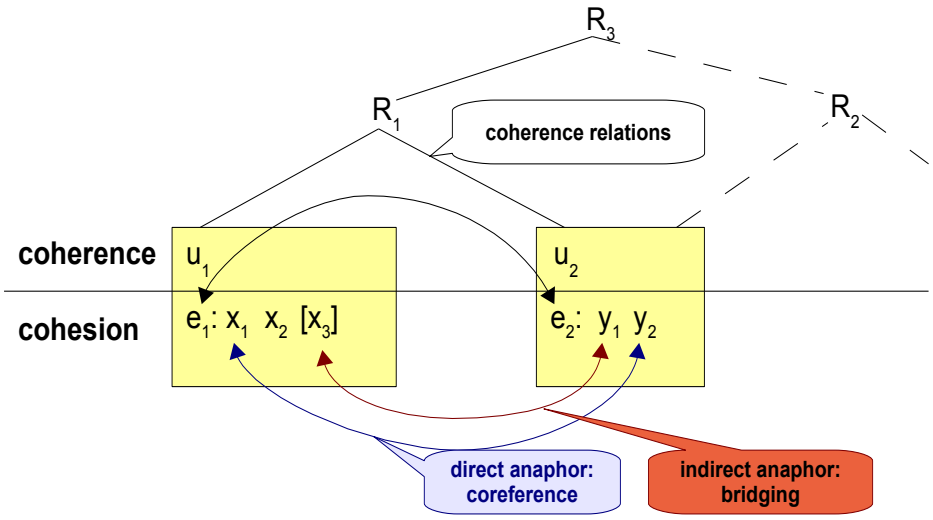


Figure 1.2: Coherence and cohesion

In course of this thesis, I will treat the characterizing properties of texts in the following order. I will deal with situationality, informativity, intentionality, and acceptability in chapter 2. Aspects of cohesion in texts are discussed in chapter 3. Coherence is subject of chapter 4.

Let us close this chapter by sketching the influential theory of discourse structure and interpretation presented by Grosz and Sidner (1986). The reason is that this theory provides a general framework covering all important factors that contribute to discourse structure, thereby providing a useful overview over the processes involved in discourse interpretation. Grosz and Sidner (1986) take an intentional perspective, drawing from insights made in plan-based reasoning developed in the field of Artificial Intelligence. The basic idea is that the speaker is executing a plan, and utterances are actions in this plan. This theory aims not only at accounting for monologic discourses but naturally extends to dialogues involving more than one speaker. A tripartite structure of discourses in general is assumed, consisting of a linguistic, an intentional, and an attentional level.

The *linguistic structure* consists of the linguistic structure of successive utterances and

17 I owe the idea to this picture to Markus Egg from a talk held at the University of Leipzig in April 2007.

their connection, and is originally assumed to be determined by cue markers. We will see later that there are many cases where discourse structure is not indicated directly on the surface.

The *intentional structure* which is assumed to be isomorphic to the linguistic structure deals with the purpose of discourse segments. Grosz and Sidner (1986) assume that each discourse segment is assigned a particular *discourse segment purpose*. The intentional structure then describes how the purpose of a discourse segment is related to the purposes of other segments.

The *attentional state* is determined by the salience of entities in a discourse segment. Salience is defined in terms of the focus of attention of the conversation participants[18]. This layer involves a stack containing the discourse entities and the degree to which they are in the focus of attention of the discourse participants.

The first two structures are responsible for coherence of a discourse, as discussed in the previous section. The third structure deals with cohesion of a discourse. Although many departures from Grosz & Sidner's original theory have been made since then in one detail or another, most theories of discourse structure adopt a particular view of this general picture. Roughly, I can state the following generalization:

Intentional approaches (cf. section 1.3.3.2) deal with the intentional structure and the purposes of discourse segments. Recently, a series of accounts have tried to give exact formulations of the role of intentions in utterance and discourse interpretation, using quite different formalisms, among them optimality theoretic accounts (Blutner, 2000; Blutner and Zeevat, 2004; Zeevat, 2006, 2009) and game theoretic accounts (Parikh, 2001; Benz et al., 2006). Intentions in discourse interpretation are subject of chapter 2.

Attentional approaches to discourse cohesion (cf. section 1.3.2) are centered on the attentional states of discourse participants and the activation or salience of discourse entities. Closely connected with the latter is the problem of anaphora resolution. Important salience-based computational theories of anaphora resolution are the Focus Theory of Sidner (1981) and Centering Theory (Grosz et al., 1995). I will discuss aspects of cohesion in discourse in chapter 3.

Informational approaches (cf. section 1.3.3.1) account for the linguistic structure of discourses. The most influential theories in this area are Interpretation as Abduction (Hobbs et al., 1993) and SDRT (Asher and Lascarides, 2003). I will discuss characteristics of discourse structure in chapter 4, and informational theories of discourse interpretation in chapter 5.

18 This conception of *focus* as used in the computational linguistic and psycholinguistic literature in the 70's and 80's (e.g. Sidner, 1981; Garrod and Sanford, 1982) should not be confused with the information structural term used by e.g. Rooth (1985); Hajičová et al. (1998).

Chapter 2

The Common Ground and Intentions in Conversations

2.1 The Common Ground

Successful communication in a conversation bears on shared knowledge of the participants. Conversation participants must know the language in use, they need to know the situation of communication and the preceding discourse. They need both general knowledge about the world and specific facts about the participants.

The widely used notion of *Common Ground* is commonly understood as a continuously changing body of public information that is used to keep track of what has happened in the conversation and to delimit the range of possible further utterances which have to be evaluated against it.

In section 2.1.1, I will review a variety of different definitions of the Common Ground that can be found in the literature. In section 2.1.2, I will look into some psycholinguistic experiments made in order to identify the actual use of the Common Ground in conversation. In section 2.1.3, I will discuss which processes play a role in establishing and structuring the Common Ground, before I conclude in section 2.1.4.

2.1.1 Definitions of Shared Knowledge

The conception of Common Ground takes a central position in the philosophical work of Robert Stalnaker. He puts *presuppositions* on a par with the Common Ground as the shared knowledge of speaker and hearer in a conversation:

> "Roughly speaking, the presuppositions of a speaker are the propositions whose truth he takes for granted as part of the background of the conversation. A proposition is presupposed if the speaker is disposed to act as if he assumes or believes that the proposition is true, and as if he assumes or believes that his audience assumes or believes that it is true as well. Presuppositions are what is taken by the speaker to be the Common Ground of the participants in the conversation, what is treated as their common knowledge or mutual knowledge."
>
> (Stalnaker, 1978, p. 320)

In this view, the Common Ground is a set of propositions believed and accepted by the conversation participants. This idea is traditionally treated formally by making use of a concept of *possible worlds*. The world is, following Wittgenstein, everything that is the

case, and a possible world is all that would be the case if the considered world were the real one. Following this idea, the context of an utterance can be considered – in a very general manner – as a set of possible worlds, the situations or worlds that are compatible with the information conveyed by the utterance. The worlds in this set, or context set (Stalnaker, 1974), have in common the information assumed to be shared by the discourse participants. An utterance can be conceived of a proposal to change the context. The content of the utterance is added to the knowledge base defining the context. In other words, all the worlds are removed from the context set in which the proposition expressed by the utterance is false[1].

Another view is taken by Lewis (1979), who compares the Common Ground with a "score board" in a baseball game where information about the players and their actions is displayed and continuously updated. This metaphor was very influential for research on dialogues and in game theoretic approaches to language use. I will discuss a particular framework that is based on this idea in section 2.2.1.

However, Clark (1996) argues that shared information in this form is not available in conversations. There is no public space where this information could be stored. The Common Ground has to be seen as an inherent part of the beliefs and knowledge of the participants because it does not exist independently from them. Therefore, the public Common Ground is at most a very abstract layer and needs good motivations for being assumed.

Before reviewing some formal attempts to define mutual knowledge, let us perform a small thought experiment. It is a well-known problem in decision theory: the Problem of Coordinated Attack or the Two Generals' Paradox(Akkoyunlu et al., 1975).

The Problem of Coordinated Attack

> Two generals who are leading two armies want to defeat a common enemy. They can only be successful if both attack at the same time. If they don't succeed, the enemy will destroy both armies. General A knows that and sends a message to B saying "I will attack at dawn". As the messenger has to cross the enemy country, A cannot be sure that the message will arrive. Hence B returns a message saying "I know that you will attack at dawn". General A receives the message. Now he knows that B knows what A wants to do, but B doesn't know yet that A knows that. General A cannot attack without running a risk. He could send back another message, but even that won't be enough. As long as the information "A will attack at dawn" and the fact that both know that is not believed by both at the same time, a decision to attack cannot be free of risk.

Communication is an activity similar to the case of the generals. The speaker sends a message that she wants the hearer to interpret as intended, and the hearer wants to understand what the speaker had in mind. Language, like a messenger in an enemy country, is an unreliable medium. Hence the speaker has to coordinate her action with the action of the hearer, and the other way round.

The scenario of the attacking generals shows how difficult can it be to reach shared or mutual knowledge. Indeed, various proposals can be found in the literature. Barwise (1988)

1 I will take up again the notion of possible worlds in section 3.4.3.

distinguishes three non-equivalent definitions of mutual knowledge: mutual knowledge as iterated propositions, as a fixed point definition, and as a shared basis.

Mutual Knowledge: Iterated Propositions

ϕ is mutual knowledge between A and B iff

1. A knows that ϕ,
2. B knows that ϕ,
3. A knows that B knows that ϕ, and B knows that A knows that ϕ, and A knows that B knows that A knows that ϕ, ... *ad infinitum* (Lewis, 1969)

This definition gives rise to the **Mutual Knowledge Paradox** (Clark and Marshall, 1981). In order to use definite descriptions correctly (e.g. "the movie" in "the movie shown tonight in the Capitol"), speaker and hearer must have mutual knowledge. If the verification of each of the above conditions takes some amount of time, then the verification of all (infinitely many) conditions would last forever. But natural language users have no problems using definite descriptions.

Mutual Knowledge: Fixed Point Definition

ϕ is mutual knowledge between A and B iff (τ) A and B know that ϕ and τ are true. (Harman, 1977)

In this definition, the infinite regression of the iterated propositions approach is replaced by self-reference. This definition entails that if there is shared knowledge then A and B know that. However, the decision whether ϕ is mutual knowledge requires the existence of an omniscient being: if ϕ is not mutual knowledge, then there is no fixed point, and this cannot be decided within a finite number of iteration steps.

A way out of this circle was proposed by Lewis (1969). The solution presupposes that some additional assumptions are made.

Mutual Knowledge by Induction: Shared Basis

ϕ is mutual knowledge between A and B iff there is a situation S so that the following holds:

1. A and B believe that S holds.
2. S indicates to A and B that A and B believe that S holds.
3. S indicates to A and B that ϕ holds.

"S indicates to x that ϕ" means: from "x believes that S holds" follows "x believes ϕ". S is called *shared basis* (Lewis, 1969).

The classical example to illustrate this definition is the following: S is a situation in which A and B are sitting around a table with a burning candle on it. The fact that there is a burning candle on the table is shared knowledge because both can see it. With that, the paradox seems to be resolved. It is sufficient to provide a basis S and mutual knowledge follows from it. It can be checked in finite time that ϕ is shared knowledge. But: it cannot be decided whether ϕ is not mutual knowledge because it cannot be excluded that there might be (potentially infinitely many) non-intended bases S' on which ϕ is shared knowledge (Barwise, 1988).

The shared basis approach seems to be the most appropriate definition of mutual knowledge because it avoids the mutual knowledge paradox without the need of an omniscient being. However, in order to find a suitable shared basis, additional assumptions are necessary. Clark and Marshall (1981) propose to use as a basis a few heuristics that are taken to underly comprehension. These heuristics are given by physical, linguistic and cultural copresence.

Physical copresence means that communication participants must be copresent in the communication situation. For instance, in order to use a deictic expression, the conversation participants must know that the object being referred to is present in the discourse situation. This kind of context is commonly known as **situational context**.

Linguistic copresence consists in that speaker and hearer must know the language being used. Moreover, conversation participants know what has been said in course of the conversation so far at any given point. Successful use of anaphoric expression relies on the knowledge of the preceding discourse. This kind of context is also known as **discourse context**.

Cultural copresence means that communication participants must be part of the same community in order to understand each other correctly. For instance, in a specific conversation about a football game, the participants need to know the rules of this game. General knowledge about the world and more specific knowledge about the culture in which a conversation takes place is needed to understand most conversations, in particular if proper names (e.g. "Noam Chomsky") are used. This kind of context is often referred to as **encyclopedic context** or general **world knowledge**.

Clark (1996) argues that from these copresence heuristics follows mutual knowledge. In an upshot, Clark divides the Common Ground into three parts: (i) the discourse context consisting of linguistic knowledge (linguistic copresence), (ii) the situational context consisting of specific facts regarding the utterance situation (physical copresence), and (iii) world knowledge, i.e. nonspecific knowledge of facts and rules in the world (cultural copresence).

With regard to the use of referring expressions, Clark and Marshall (1981) propose that mutual knowledge is a gradual notion, depending on the type of evidence for mutual knowledge. The strongest evidence for mutual knowledge is physical copresence while linguistic copresence provides weaker evidence. Different means of reference rely on the different types of context. The linguistic context enables the use of definite expressions, the situational context is reflected by demonstratives and indexical expressions, and the encyclopedic context gives the background for proper names and some definite descriptions.

2.1.2 The Use of the Common Ground in Conversation

As can be concluded from the last section, it is by no means easy to give a simple but satisfying definition of mutual knowledge. As the reviewed definitions are of theoretical nature, I will now turn to experimental evidence regarding the use of Common Ground for identifying the referents of referring expressions. The main question I am concerned with is: Is mutual knowledge really a precondition for successful communication? I will look at experiments dealing with both language production and language comprehension.

2.1.2.1 Experiments in Language Production

The model of speech production developed by (Levelt, 1989) distinguishes three phases: Planning, Formulating and Articulation (with Monitoring). The question I am interested in here is in which phase the Common Ground is established. Two hypotheses can be assumed. The *Initial Design Model* maintains the assumption that the Common Ground is already used when an utterance is planned. In this model it is assumed that people are interacting, in contrast to the *Monitoring and Adjustment Model* that assumes people to be egocentric. In this model, Common Ground is taken into account only in the process of adjustment when monitoring during articulation shows that the initial planning hasn't been sufficient.

An experiment in order to decide between these two models was made by Horton and Keysar (1996). In this experiment, it is shown that speakers do less rely on private contextual information than on shared contextual information – provided that they have all the time they need to carry out a speech act. In contrast, when they are pressed for time, they rely equally on both types of contextual information. This result can be interpreted as in favour of Monitoring and Adjustment: with time pressure, speakers revert to an initial plan that did not take into account the Common Ground. In this case, Monitoring is left out.

2.1.2.2 Experiments in Language Comprehension

Let us turn now to an experiment regarding language comprehension carried out by Keysar et al. (2000). The hypothesis is that the strategy applied by hearers in the search for referents of definite descriptions is based on an egocentric heuristic: they consider potential referents which are not part of the Common Ground but which can be considered from their own perspective. Hearers use shared knowledge only to correct errors which occur using the egocentric heuristics.

An eye-tracking experiment was carried out on 20 English native speakers. Two subjects, a "director" and an "addressee", were seated in front of each other. The addressees were told that the distribution of roles was random. The task of the addressees was to arrange a couple of objects following the instructions of the director.

The critical instruction was ambiguous: the shared perspective contained one potential referent, and the addressee's perspective contained another potential referent, which, in turn, was not visible for the director. In the control condition, the object not visible to the director was changed in a way that it did not represent a potential referent in the critical utterance any more. The time frame for observation of eye-fixations was from the beginning of the utterance of the word representing the shared referent until the referent was identified.

The experiment yielded the following results: even though addressees noted the intended object (by initial eye-fixation), for their decision that it was the correct referent they needed more time in the test condition than in the control condition (2,092 ms vs. 1,146 ms). In 23% of the cases in the test condition addressees reached for the hidden object, in 6% of the cases the decision was still corrected, and 17% moved the object.

In a second experiment, the addressees had to hide themselves the objects so that it was guaranteed that they knew perfectly what was hidden and what not. The results were similar: in 20% of the cases they reached for the hidden object (15%) or even moved it

(5%).

The conclusion that can be drawn from these experiments is the following: the egocentric interpretation can be so strong that it has the potential to overwrite the knowledge that the other participant cannot see the hidden object. If shared knowledge has an influence in the process of comprehension, then at most partially. The Common Ground plays no or only a partial role in the initial phase of the comprehension process, but it surely is used later for error correction.

2.1.2.3 Discussion and Conclusion

Language users do not always take the Common Ground into account while generating or interpreting referring expressions. They do so only in processes of monitoring or correction. Establishing the Common Ground needs the conversation participants to build and maintain a complex model of the conversation situation in which two things have to be considered: (i) the own knowledge, and (ii) knowledge assumed to be shared with the other participants. Hence, taking into account the Common Ground requires a lot of effort and processing costs and is not necessary for many conversations. Nevertheless, communication succeeds normally. A plausible reason could be that the speaker's perspective normally does not differ much from the shared perspective. In other words, in most conversations, the speaker's view of the world is not very different from the hearer's view of the world. As a result, even if a speaker does not take into account that the addressee might have a different perspective, she will be understood, at least sufficiently for the purpose of the conversation.

A recent psycholinguistic model that takes these findings into account is the *Interactive Alignment Model* (Pickering and Garrod, 2004). These authors start from the observation that spontaneous spoken language is the most natural form of language use. Linguistic representations of conversation participants are aligned on various representational levels. Alignment takes place on all linguistic levels from phonetic and phonological representations over syntactical representations up to lexical and semantic representations, and also on the level of the situation model.

Alignment of the discourse situation models is an automatic process following from the alignment on lower levels. Conversation participants tend to have similar situation models to make the same changes on it. Two forms of establishment of the Common Ground can be distinguished. On the one hand, an implicit Common Ground is established automatically. On the other hand, the full Common Ground is established in a controlled way only if the situation models of the conversation participants are differing too much. But "simple" conversation works without these strategies.

In dialogues, an interactive process of error correction takes place if the implicit Common Ground is erroneous: the hearer checks if he can interpret an utterance with regard to his own representations. If this fails, the utterance must be rephrased in a way that the implicit Common Ground can be established.

In monologues, automatic alignment and interactive repair is not possible. In case of misunderstandings, the hearer can try to draw inferences from the perspective assumed to be shared. The speaker has to model the cognitive state of the hearer and must be able to put herself in the hearer's perspective.

Let us summarize the findings of this section. Mutual knowledge does not always play a role in producing and interpreting utterances. In normal conversation situations, the

shared perspective of all participants overlaps with the own perspective for the most part. It seems that participants do not explicitly model the cognitive state of conversation partners. Mutual knowledge is used for error correction if communication fails and if a speaker wants to be sure that her utterance succeeds.

With this result, we have to assume a distinct representation of the Common Ground for each conversation participant. The knowledge in these representations – be them as similar or equal as they may – can be divided into knowledge of the situational context, the discourse context, and world knowledge.

2.1.3 Establishing the Common Ground

I turn now to the way how pieces of knowledge enter the Common Ground and how they can change it. As a conversation proceeds or a text is read, the Common Ground accumulates (section 2.1.3.1) in a process of Grounding (section 2.1.3.2). A way of structuring the resulting Common Ground is presented in section 2.1.3.3.

2.1.3.1 Accumulation

Every communicative action is affected by the Common Ground and, at the same time, affects the Common Ground. The Common Ground is continuously changing during a conversation. As Lewis (1979, p. 339) puts it, "presuppositions can be created or destroyed in the course of a conversation". Different types of speech acts change the Common Ground in different ways: Assertions "change the presuppositions of the participants in the conversation by adding the content of what is asserted to what is presupposed. This effect is avoided only if the assertion is rejected" (Stalnaker, 1978, p. 323). Questions can be seen as a request to update the Common Ground, and other speech acts affect the Common Ground in other ways.

As a discourse proceeds, conversation participants keep track of the Common Ground and update it moment by moment, and as a result they come to share more and more information. In this way, the Common Ground *accumulates* in course of a conversation. Common Ground can thus be viewed as a dynamic concept. It is important to note that even if presuppositions are destroyed, i.e. they do not become part of the Common Ground, the participants know that they have been destroyed, and this knowledge itself becomes part of the Common Ground. So accumulation means that there is no way that a Common Ground gets smaller.

An essential precondition for the accumulation of the Common Ground is that the participants make sure that each utterance is understood as intended. And, most crucially, the content of an utterance is not added automatically to the Common Ground, but the discourse participants have to make positive steps in order to integrate the content of an utterance into the Common Ground.

2.1.3.2 Grounding

Grounding is understood as the process of establishing contents as part of the Common Ground (Clark and Schaefer, 1989; Clark and Brennan, 1991). Conversation participants try to make clear that what is said is being understood correctly. They try to reach a state in which they both believe of each other that the other has understood what was meant by an utterance. Clark and Schaefer (1989, p. 262) define the *grounding criterion* as follows:

"The contributor and the partners mutually believe that the partners have understood what the contributor meant to a criterion sufficient for current purposes". Grounding then is the collective process by which the participants try to reach this criterion.

In the following paragraphs, I will present two approaches to grounding in conversation. First, a psychologically oriented account of *contributions*, and second, a speech act based account of *grounding acts*.

Contributions and Acknowledgments

Following Clark and Schaefer (1989), a conversation consists of *contributions*. Contributing to a discourse requires more than just uttering the right words at the right time. It consists of collective acts performed by the participants working together. A contribution can be divided into two phases: a presentation phase and an acceptance phase.

In the *presentation phase*, A presents to B the utterance u. This happens under the assumption that A can believe that B understands what A has meant by uttering u if B gives at least evidence (feedback) e.

In the *acceptance phase*, B accepts the utterance u by giving evidence e' that he believes that he understands what A has meant by her utterance u. This happens under the assumption that if A notes the evidence e', she will also believe that B understands u. In this phase, two cases are possible: (i) B indicates that he has understood, or (ii) B indicates difficulties in understanding.

In the first case, the hearer can signal his understanding in a strong, explicit form by making an acknowledgement (e.g. "o.k."), or in a weaker, implicit way by showing continued attention or by taking the next turn. Which type of evidence B should give? The acceptance process is recursive: B's evidence as answer to A's presentation is itself another presentation which has to be accepted. Where does this recursion lead to? Clark and Schaefer (1989) state the *Principle of Strength of Evidence*: The contribution participants expect the following: if evidence e_0 is needed for the acceptance of presentation u_0, and e_1 for accepting e_0, then e_1 is weaker than e_0.

In the second case, the hearer signals difficulties in understanding. For every utterance u which is part of A's presentation, B believes that he is in one of four states of understanding (Clark and Schaefer, 1989, p. 268).

- State 0: B didn't notice that A uttered any utterance u.
- State 1: B noticed that A uttered some u (but wasn't in state 2).
- State 2: B correctly heard u (but wasn't in state 3).
- State 3: B understood what A meant by u.

The goal of A and B is to reach the mutual belief that B is in state 3. The *Principle of Least Collaborative Effort* helps to reach this state: participants in a conversation try to minimize the effort needed for the contribution, in both presentation and acceptance phases. The more effort is spent for the design of a suitable presentation, the less effort is needed for acceptance.

Grounding Acts

Any action in a conversation becomes part of the Common Ground. Stalnaker (1978, p.

323) notes that "the fact that a speaker is speaking [...] is a fact that is usually accessible to everyone present". Traum and Allen (1992) present an account of grounding based on speech acts. Following them, a conversation consists of *conversation acts*. They distinguish core speech acts (inform, yes-no-question, check, suggest, request, accept, reject, ...), acts for turn-taking (take-turn, keep-turn, ...), acts for argumentation (elaborate, summarize, clarify, convince, ...) and acts for grounding (initiate, continue, acknowledge, repair, request-repair, req-acknowledgement, cancel).

When a speaker makes an utterance, apart from performing a core speech act, she also performs one or more *grounding acts*: operations that either extend the part of the discourse model containing *ungrounded* information (unacknowledged statements), or change the state of contents from *ungrounded* to *grounded* (acknowledgements). Presented material that can be confirmed together (e.g. by a single "o.k.") is grouped together in a discourse unit.

A model of discourse and dialogue interpretation based on these considerations was proposed by Poesio and Traum (1997). It is an extension of a compositional variant of DRT (see section 3.4.3) where interpretation includes the recognition of the illocutionary speech act performed by an utterance. The Common Ground is seen as a protocol of the discourse situation, in contrast to the protocol of the described situation used in standard DRT (Kamp and Reyle, 1993). It consists of two parts: the grounded root DRS represents the actual Common Ground, and the ungrounded root DRS is an extension of it containing all discourse entities which are ungrounded. Each of these discourse entities represents the view of a participant of what should become part of the Common Ground.

2.1.3.3 Structuring the Common Ground

As Ginzburg (1996) pointed out, Stalnaker limited himself to the contextual change brought about by assertions, thus seeing the context as a set of propositions, or commonly accepted facts. This is certainly one component of the Common Ground.

Moreover, as seen in the last sections, an assertion does not automatically add a fact to the Common Ground, rather this is done by a Grounding process. The latest move made in a conversation imposes restrictions on the available reactions or continuations in the next move. Hence, the Common Ground should contain information about the last move, or even keep track of all the moves made in the conversation so far.

Apart from assertions and their establishment in the Common Ground, questions also play an important role in structuring conversations. In fact, many theorists both in formal semantics and conversational analysis assume that a question together with its answer form a single discourse unit. This is the starting point Ginzburg (1996) takes in his approach. In order to keep track of the questions that are under discussion at a given point, Ginzburg assumes that the Common Ground, in addition to the facts and information about the latest move, must also contain a stack, or, more general, a partially ordered set of questions under discussions, or QUD. The maximal element of QUD corresponds to the discourse topic.

This way, the following picture of a conversation participant emerges. Each participant's mental state or her *information state* is made up of a private, unpublicized mental situation and a public game board. The game board represents her view of the Common Ground and consists of (i) FACTS: a set of facts commonly agreed upon, (ii) MOVES: the moves made in the conversation so far, and within that, the content of the latest move made, LATEST-

MOVE, and (iii) QUD: a partially ordered set that specifies the currently discussable questions.

I will not dive further into this theory at this point, since it would lead us too far away from the main concerns of this thesis. However, I will come back to this approach in chapter 4 when discussing approaches to discourse structuring (section 4.3.3.4).

2.1.4 Conclusion

Linguistic communication relies on shared knowledge of discourse participants. However, speakers and hearers do not always take into account what is mutually believed. For most conversations it is sufficient to assume that the shared perspective of all participants is very close to the perspective of a particular participant.

The process of Grounding makes sure that contents are in fact shared knowledge. Heuristics of copresence can serve as a basis for shared knowledge. Generally speaking, the context of an utterance is what is assumed to be shared knowledge. Three types of context can be distinguished: the situational context, the discourse context, and general world knowledge.

What consequences for formalization and implementation can be drawn? The fact that conversation participants do not take into account the cognitive state of their partners in the first place suggests a level of representation without this component. However, as the Common Ground can be used in both production and interpretation, a formal theory of discourse interpretation cannot avoid including a representation of the Common Ground. Formal models should make a distinction between contents of which can be assumed that they are shared knowledge (i.e. grounded) and contents of which this condition is still unclear (i.e. ungrounded).

To sum up our discussion of the Common Ground, I will cite Ginzburg (2009) who enumerates possible models of what the information state of conversation participants A and B should consist of:

(2.1) a. A: $\langle A.private \rangle$, B: $\langle B.private \rangle$

b. A: $\langle pub \rangle$, B: $\langle pub \rangle$

c. A: $\langle A.private, A.B.private \rangle$, B: $\langle B.private, B.A.private \rangle$

d. A: $\langle pub, A.private \rangle$, B: $\langle pub, B.private \rangle$

e. A: $\langle A.pub, A.private \rangle$, B: $\langle B.pub, B.private \rangle$

Position (2.1a) is the approach taken in classical Artificial Intelligence. It does not reflect the distinction between private and shared knowledge that is useful for explaining linguistic data.

Position (2.1b) is the approach taken by most formal semantics theories, assuming that all information is public and shared. This view can be modelled straightforwardly by the concept of possible worlds (cf. sections 2.1.1 and 3.4.3). However, this picture is certainly oversimplifying real conversation processes.

Position (2.1c) assumes that conversation participants keep track of both their own private information and the information that is private to the other participants. As discussed in section 2.1.1 above, this assumption leads to the mutual knowledge paradox,

and moreover, as seen in section 2.1.2, most conversations do not always rely on such a sophisticated view of the information states of participants.

Position (2.1d) makes a distinction between information that is private to each conversation participant and shared public knowledge. This is the view adopted by Lewis (1979) and brought forward by Parikh (2001)'s game theoretic account of language use, which I will discuss below in section 2.2.1. As noted above, it is unclear where in the world an objective and independent scoreboard representing the context should exist. This argument leads us to

position (2.1e) in which each participant has her own view on public knowledge. This approach is taken by Poesio and Traum (1997) and Ginzburg (1996, 2009). The latter account was sketched briefly in section 2.1.3.3.

2.2 Modelling Intentions in Discourses

In section 1, I introduced Grosz and Sidner (1986)'s general model of discourse structure and interpretation. According to this view, discourses consist of a linguistic structure, an intentional structure, and attentional states. Before drawing the attention to the attentional and linguistic levels in the following chapters, I will dedicate the remainder of this chapter to the role of intentions in discourse interpretation.

Recently, a series of accounts have tried to formulate the role of intentions in discourse interpretation in a more exact way by virtue of using quite different formalisms, among them optimality theoretical accounts (Blutner, 2000; Blutner and Zeevat, 2004; Zeevat, 2006, 2009) and game theoretical accounts (Parikh, 2001; Benz et al., 2006). Here, I want to sketch an approach to language use based on game theory.

2.2.1 Linguistic Communication as a Game

At various points in this chapter, I mentioned the suggestion that a conversation can be seen as a kind of game. Think of utterances as moves of players. The aim of the game is, for the speaker, to convey information as intended, and for the hearer, to recognize the intention behind the utterance. In this section, I will present how the information flow in communication can be represented in game-theoretic terms.

2.2.1.1 Basics of Game Theory

In a game, players have to make decisions in their own interest. Crucially, the outcome of a game is influenced by decisions other players make. Modelling games is the aim of game theory. Linguistic communication can be seen as a game in which the conversation participants act as players. The speaker makes an utterance and the hearer tries to interpret it. They take possible actions of the others into account before choosing which utterance they make or which interpretation is to be preferred.

Before sketching how linguistic communication can be modelled as a game, I will introduce a few basic notions of game theory.

Rational Decisions

Players are rational agents who have to choose an action from a set of possible actions. They have preferences for certain actions which underlie their decisions. Every action has

a *payoff* which can be measured on a scale of *utility*. A rational agent chooses the action with the highest utility. Decisions can be made with certainty, at risk, or with uncertainty.

Decisions at Risk

Normally, a player is uncertain about the outcome of her actions because she has only restricted information about the real state of the world. A player has nothing but certain expectations about possible outcomes. She assigns a probability distribution to possible outcomes, and a payoff to every outcome. She decides on the action with the highest *expected* utility.

Strategic Interaction

The actions of other agents influence the agent's payoff. Payoffs are functions of the actions of all participants. A rational agent draws a *strategic inference* when she takes into account possible actions of the other rational agents, as well as her knowledge and beliefs, in particular the knowledge of the situation shared with the others.

Equilibrium

A *Nash equilibrium* exists if the combination of the decisions of two or more agents is such that no agent has an incentive to change her actions. The basic idea behind an equilibrium for several agents corresponds to what is optimality for one agent.

Cooperative Games

To get an idea how these notions can be applied, I will consider a simple example known as the "battle of sexes". Suppose that Adam and Eve want to go out together. Unfortunately, they forgot to agree on the location. There are two possibilities: a football match or a concert. Each player has to make his/her decision on his/her own, independently of the other. Adam prefers to go to the concert, and Eve likes to watch the football match. As both of them prefer to spend the evening together rather than staying alone, they should avoid missing each other. If Eve goes to the football match, Adam should go there, as well. Conversely, if Adam goes to the concert, Eve should also go there. The two strategies resulting in meeting each other yield the highest payoff. These are Nash equilibria.

The game can be represented by a payoff matrix as shown in Fig. 2.1. A strategy is represented as a tuple of numbers representing the payoffs for each player. Eve is the "row player", and Adam is the "column player". An alternative sequential representation in extensive form as a game tree is given in Fig. 2.2. Here, the state labelled with "E" corresponds to Eve's choice, and the states labelled with "A" correspond to Adam's choice. Adam does not know in which of the both states he is in. In the game tree, the two states which cannot be distinguished by Adam are marked by the oval around both states.

Eve \ Adam	Football	Concert
Football	$(3;1)$	$(0;0)$
Concert	$(0;0)$	$(1;3)$

Figure 2.1: *Battle of Sexes* as payoff matrix

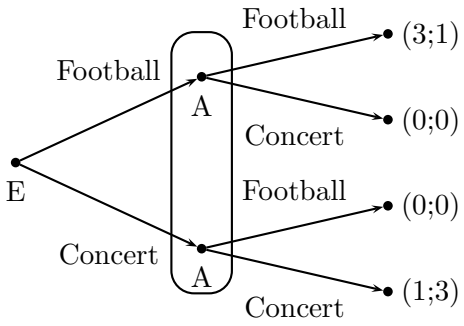


Figure 2.2: *Battle of Sexes* in extensive form

2.2.1.2 Communication as a Game with Partial Information

Prashant Parikh was one of the first who applied ideas of game theory in linguistics. He extended the idea of communication as a signaling game as proposed by Lewis (1969) and presented a model of communication as a sequential cooperative game with partial information (Parikh, 2001). As his approach is intuitively accessible and rather straightforward, I will present it here in some detail. Other, in some respect more sophisticated models, as well as an excellent introduction to game theory can be found in Benz et al. (2006). In his book, Parikh claims to provide solutions to some semantic and pragmatic problems such as quantifier scope, lexical and structural ambiguity, resolution of indexicals, as well as resolution of pronoun and noun phrase reference.

The Discourse Situation

The discourse situation assumed by Parikh starts from two rational agents A and B who have shared knowledge about their rationality. A utters sentence ϕ in a discourse situation d in order to convey information p to B. B in turn tries to interpret A's utterance in d. He uses his linguistic knowledge for determining its meaning $m(\phi)$. The *meaning* of a sentence is a collection of possible *contents* of a sentence, and the content of an utterance is specific to a situation.

Public knowledge accessible to all agents includes the agent's state before the utterance, the uttered sentence after the utterance, and the purely linguistically determined meaning of the sentence. Private knowledge are beliefs and intentions of speaker and hearer. With this distinction, Parikh's model of the information state of conversation participants is essentially that of Lewis (1979), distinguishing between private and public knowledge, and taking public knowledge as an abstraction over knowledge assumed to be shared by all participants[2].

Strategic Inference and Partial Information

Communication is a special case of information flow. There are many ways information can flow. For example, smoke means fire. But smoke does not communicate that there is a fire. It does not have the *intention* to convey this information. Following Grice, in communication, the speaker has the intention to convey the relevant information, and the

2 See the discussion of models of the Common Ground in section 2.1.4 on page 52.

addressee wants to recognize this intention. A this point, game theory comes in. Every intended information flow between the agents involves a strategic interaction between them. If this strategic interaction is shared knowledge, then the information flow is communicative. This case can be seen as a game with a unique solution. Thus we can say that A communicates with B if and only if there is a game between A and B. Crucially, every piece of communicated information requires its own strategic inference. As a consequence, every utterance demands several distinct strategic inferences which are simultaneous and which affect each other.

As this picture turns out to be quite complex, Parikh makes a simplification. He considers only one strategic inference in isolation, assuming that B has already the partial information of the other inferences, enabling him to infer the intended content from this and the utterance situation. The communicated content depends on (i) what has been uttered, (ii) what the speaker *could* have said but did not say, and (iii) the shared information about these decisions. Thus, the constructed structure is a *game of partial information*. Note that the game is not played explicitly, but is intended to be understood as model reflecting the logic underlying communication.

In order to get a feeling how the proposed model works, let us go through an example of Parikh's. Suppose that A utters ϕ (2.2) to B in situation d.

(2.2) Every ten minutes a man gets mugged in New York.

This sentence has two readings:

- p: Some man or other gets mugged every ten minutes.
- p': A particular man gets mugged every ten minutes.

Situation d enables B to disambiguate ϕ. A communicates p to B if the following conditions are met:

1. Background Assumptions
   - A and B are rational,
   - L is a common language,
   - m is the meaning function of L: a function from sentences to propositions,
   - these assumptions are shared knowledge between A and B.
2. Specific Assumptions
   - A has the intention to convey p and utters ϕ,
   - B has the intention to interpret ϕ, and receives and interprets ϕ,
   - $m(\phi) = \{p, p'\}$,
   - p' is relatively improbable,
   - expressing p or p' unambiguously is more effortful than using the ambiguous form,
   - everything but the intentions is common knowledge.

The course of the game is depicted in Fig. 2.3. There are two initial situations, s and s'. In s, the speaker wants to convey p, and in s' she wants to convey p'. In both cases, she

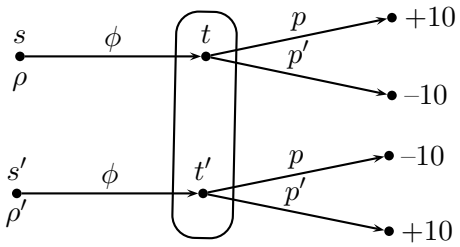


Figure 2.3: Game for uttering (2.2) (cf. Parikh, 2001, Fig. 3)

can make use of utterance ϕ. After the utterance there are again two situations, t and t'. A knows her intuitions, and also in which situation she is in. B cannot distinguish between the two situations. As in the previous example, this feature is marked in the game tree by the oval around the indistinguishable states. B assigns to s and s' the probabilities ρ and ρ', respectively. Then he chooses an interpretation, p or p'.

The decision preferable for B depends on information B does not have. If he is in situation t, he should prefer p. If he is in t', he prefers p'. The payoff for successful communication is, say, +10, and for miscommunication –10.
B is in the dilemma that he cannot disambiguate the sentence by means of the information considered so far. An additional assumption is necessary. As noted above, successful communication requires taking alternative utterances into account. The addressee has to compare the ambiguous utterance with unambiguous alternatives in order to make sure that the former is more efficient. Utterances which unambiguously convey the two readings of ϕ are the following:

- μ: Every ten minutes some man or other gets mugged in New York.
- μ': Every ten minutes a particular man gets mugged in New York.

For these utterances, the following holds: $m(\mu) = \{p\}$ and $m(\mu') = \{p'\}$. With that, we get an extended game tree as shown in Fig. 2.4.

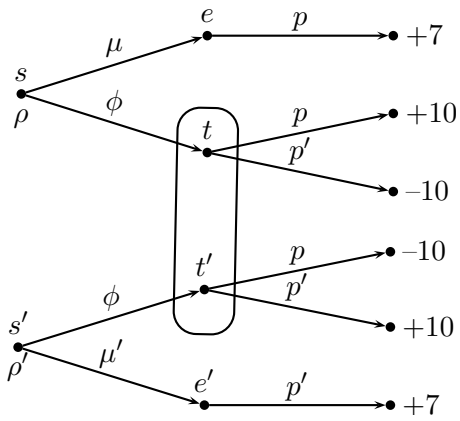


Figure 2.4: Game for uttering (2.2) with alternatives (cf. Parikh, 2001, Fig. 5)

The alternatives are not ambiguous. In situation e, B's only choice is to interpret the utterance as conveying p, and in e', as p'. The additional cost for making an unambiguous utterance is assessed at –3. Utterance ϕ has a higher payoff than μ and is to be preferred by the speaker.

B's problem is the following: is there enough information in the game to rule out t' as a possible situation and to be able to choose p?

A has the problem that the question whether her choice is optimal depends also from the question whether B has enough information to solve the problem.

Game Solutions

The applied game theoretic solution concept is known as Nash equilibrium with Pareto dominance. A *strategy* specifies the actions for a player in the situations in which she has to make a decision. It is a function from a set of information states in a set of actions. A strategy is a *Nash equilibrium* if no player has a reason to deviate unilaterally from this strategy.

Possible strategies for A are the following: (i) $\langle(s,\phi),(s',\mu')\rangle$, (ii) $\langle(s,\phi),(s',\phi)\rangle$, (iii) $\langle(s,\mu),(s',\phi)\rangle$, and (iv) $\langle(s,\mu),(s,\mu')\rangle$. For B, there are two strategies: $\langle(\{t,t'\},p)\rangle$ and $\langle(\{t,t'\},p')\rangle$.

The strategies of both players together span a strategy space with 8 strategies, two of which are Nash equilibria:

- $N_1 = \langle(s,\phi),(s',\mu');(\{t,t'\},p)\rangle$ and
- $N_2 = \langle(s,\mu),(s',\phi);(\{t,t'\},p')\rangle$.

Now we have more than one solution for the game. It arises from Parikh's modeling of communication as a cooperative game, which, in general, allows for multiple solutions. Parikh's solution for this problem is to exploit the concept of Pareto dominance to exclude undesirable Nash equilibria. *Pareto dominance* is defined as follows: if one of two strategies in a game yields higher payoffs for all players, the other one can be eliminated. In this way, the Nash equilibrium with the highest expected payoff is selected. In our example, let be $\rho = 0.9$ and $\rho' = 0.1$, then we get

- for N_1: $0.9 * 10 + 0.1 * 7 = 9.7$,
- for N_2: $0.9 * 7 + 0.1 * 10 = 7.3$.

Thus, the intuitively correct solution N_1 pareto-dominates N_2. The optimal expected payoff is much higher than in a decision by chance. So far, we have a nice account for resolving scope ambiguities like in example (2.2), taking into account intentions of discourse participants. Now consider the following ambiguous sentence:

(2.3) A comet appears every ten years.

In this case, none of the two readings is more probable than the other one, and the second solution cannot be eliminated. Hence, in a null context, the ambiguity in (2.3) cannot be resolved. This example points to the crucial problem for all classical game theoretic approaches: from where shall we take the probabilities? In the discussed example, ρ was the probability that the speaker wants to convey p in situation s. Following Parikh (2001),

ρ depends on the probability that p is true. This probability is in general very subjective and varies from one speaker to another. Allott (2006) presents an alternative view assuming that ρ reflects the salience, or activation in psycholinguistic terms, of a reading. It is not fixed *a priori*, but enters into the payoff as some sort of cost. Less activation of a reading yields bigger costs. This view is indeed very plausible, but still does not account for the numerical values we need for calculating the payoffs.

2.2.1.3 Conclusion

Parikh's account exploits classical game theory for modeling communication. It is a precise game theoretic formulation of pragmatic resolution of semantic underspecification. However, this model has to make some rather strong underlying assumptions. These are the following: (i) conversation participants are rational, and rationality is applied directly on the discourse situation, (ii) speaker and hearer are cooperative, (iii) language is assumed to be given, (iv) all knowledge but the intentions is public and equal for all participants, and (v) conversation participants have mutual knowledge of the ability to draw complex inferences about the behaviour of the others. There have been attempts to soften the idealistic assumption of perfect rationality of communication participants, among them, perhaps most appealing, evolutionary game theory (cf. e.g. Jäger, 2007). I will not go into further details here. A compact introduction to evolutionary game theory and its application in linguistics is presented in Benz et al. (2006).

Chapter 3

The Discourse Model and Discourse Anaphora

This chapter deals with cohesion and, closely connected to that, discourse representation. Cohesive phenomena (cf. section 1.3.2) like anaphoric relations between entities mentioned in a text play a key role in the search for a discourse representation which can provide a suitable basis for interpreting texts. Such a representation, the discourse model, contains the entities anaphoric expressions refer to. These entities, discourse referents, can serve as antecedents for subsequent anaphora to a variable degree. Theories of anaphora resolution, i.e. the process of finding the correct antecedents for anaphora, can profit from a suitable representation of a discourse model and discourse referents therein.

Section 3.1 places the focus of attention on discourse anaphora, the major device for establishing cohesion in a discourse. I will first determine various types of anaphoric expressions and examine their distribution. Then, different syntactic and semantic notions of anaphora will be reviewed in order to clarify the term discourse anaphora as used in in this thesis.

Subject of section 3.2 is the discourse model. This notion will be reviewed from different points of view. I will present a psycholinguistic attempt to explain how texts are represented in human minds as mental models, make some general remarks on formal ways of representing the meaning of discourses, and determine the basic requirements on discourse models.

Section 3.3 is devoted to discourse referents. First, I will introduce this notion in section 3.3.1, and then, in section 3.3.2, I will review competing notions used for explaining how discourse referents can be accessed as antecedents for discourse anaphora.

In section 3.4, I will discuss approaches to anaphora resolution from pragmatic (section 3.4.1), computational (section 3.4.2) and formal points of view (section 3.4.3).

3.1 Discourse Anaphora

Typically, a natural language discourse is full of anaphoric relations, both within a single utterance and spanning bigger distances. In chapter 1, they were introduced as being responsible for cohesion in a text. Recall that an anaphoric relation holds if an anaphoric expression makes reference to a concept already introduced in the text, the antecedent. There is a wide range of anaphoric phenomena, some of which can be observed in the following excerpt from a fairytale.

(3.1) a. There was once a miller who was poor, but he had one beautiful daughter.

b. It happened one day that he came to speak with the king, and, to give himself consequence, he told him that he had a daughter who could spin gold out of straw.

c. The king said to the miller,

d. "That is an art that pleases me well;

e. if thy daughter is as clever as you say, bring her to my castle to-morrow, that I may put her to the proof."

(The brothers Grimm: Rumpelstiltskin, KHM 055)

A variety of expressions can be anaphoric, among them personal pronouns, reflexive and reciprocal pronouns, null pronouns, definite noun phrases, and even other noun phrases such as indefinite NPs and proper names. Let us examine some occurences of these types in text (3.1).

3.1.1 Types of Anaphoric Expressions

Personal Pronouns
In the first clause of (3.1a), the noun phrase "a miller" introduces an entity into the discourse. This entity serves as the antecedent for the personal pronoun "he" in the second clause. Personal pronouns like these are the canonical form of anaphoric expressions. Note that personal pronouns can have antecedents in other sentences. This is the case of the personal pronoun "he" in (3.1b).

Reflexive, Reciprocal, and Null Pronouns
There are also other kinds of anaphoric pronouns. In the second clause of (3.1b), we can find three instances of anaphoric pronouns: "he", "him", and "himself". "He", as before, refers to the miller, whereas "him" refers to the king. The reciprocal pronoun "himself" refers to the subject of the embedded clause which is not expressed but unequivocally meant to refer to the miller. In these cases, many theories speak of null pronouns (e.g. PRO). In the following examples, anaphoric expressions and their antecedents are marked by the same index, i .e. they are coindexed.

(3.2) To PRO_1 give himself_1 consequence, he_1 ...

Crucially, if a pronoun wants to access an antecedent in the same core clause it must be expressed as a reflexive or reciprocal pronoun. Look again at (3.1b), repeated as (3.3), where expressions referring to the miller are marked by the index 1, and those referring to the king by 2. Here, a reading in which "himself" is coreferential with the king is not possible. Hence, there is a division of labour between personal pronouns and reflexive / reciprocal pronouns.

(3.3) To give PRO_1 $\text{himself}_{1/*2}$ consequence, he_1 told $\text{him}_{*1/2}$ that he_1 had a daughter ...

Definite Descriptions

Not only pronouns but also other definite noun phrases are anaphoric, as can be observed in (3.1c). Both "the king" and "the miller" refer to entities introduced before. The antecedents of definite NPs are usually in preceding sentences. Definite NPs where antecedent and anaphor are in the same sentence are, although grammatical, highly marked. For example, (3.4) does not sound very well.

(3.4) # There was once a miller$_1$ who was poor, but the miller$_1$ had one beautiful daughter.

A feature that distinguishes the anaphoric behaviour of definite descriptions from that of pronouns is the fact that definite descriptions can be *indirect anaphora* or *bridging anaphora* (cf. section 1.3.2). In example (3.5), "the door" bears an indirect anaphoric relationship to "a room".

(3.5) When the girl was brought to him, he led her into a room that was quite full of straw, and gave her a wheel and spindle, and said, "Now set to work, and if by the early morning thou hast not spun this straw to gold thou shalt die." And he shut **the door** himself, and left her there alone.

Proper Names

Even proper names are, when they are repeated, anaphoric. The occurence of "Rapunzel" in example (3.6b) is anaphoric to the child introduced in (3.6a).

(3.6) a. When the time came when the child was born the witch appeared, and, giving the child the name of Rapunzel, she took it away with her.

b. **Rapunzel** was the most beautiful child in the world.

(The brothers Grimm: Rapunzel, KHM 012)

Indefinite Noun Phrases

Indefinite noun phrases are prototypically used to introduce a new entity in a discourse. However, they can be anaphoric under certain circumstances. In the first place, this is the case of indirect anaphora. For example, "a window" in (3.7) is, although new in the discourse, indirectly related to the previously introduced "the chapel".

(3.7) When the boar caught sight of the tailor he ran at him with foaming mouth and gleaming tusks to bear him to the ground, but the nimble hero rushed into a chapel which chanced to be near, and jumped quickly out of **a window** on the other side.

(The brothers Grimm: The Gallant Tailor, KHM 020)

Deictic Expressions

Yet another kind of anaphoric expressions are *deictic* pronouns, e.g. "here", "there". They usually take their referents from the utterance situation, e.g. the text (3.8) on a traffic sign.

(3.8) Stop here on red!

Note that, however, the antecedent of deictic pronouns can frequently be found in the linguistic context, as in (3.9).

(3.9) Hans went to a bar_1. There_1, he ordered a beer.

In fact, almost always, a referential expression (e.g. a pronoun or a definite NP) can have a deictic reading, i.e. with an antecedent in the utterance situation. For example, in a comic strip, the following sentence can be read on a Wanted poster together with a picture of Spiderman. Addressees can identify the referent of "he" although Spiderman was not linguistically introduced.

(3.10) Caution: He is dangerous.

(Stan Lee (1963): The Amazing Spider-Man, Vol. 1(1))

Although deictic readings of anaphoric expressions are often possible, these readings are usually not the preferred readings if a suitable antecedent can be found in the linguistic context.

Distribution of Anaphoric Expressions

It can be observed that anaphoric expressions have the following distributional pattern: reflexive and reciprocal pronouns must have an antecedent in the same core sentence, i.e. the same clause. Personal pronouns can have antecedents both in the same sentence (although not in the same core sentence) or in other sentences. Definite NPs usually have an antecedent in another sentence.

Pronouns are typically direct anaphora while definite descriptions (and sometimes indefinite NPs) can be indirect anaphora. All types of anaphoric expressions (except reflexive and reciprocal pronouns) can have a deictic reading. The distribution of anaphoric expressions is summarized in Table 3.1.

	direct antecedent			indirect antecedent	deictic use
	same core sentence	same sentence	other sentence		
reflexive pronouns	√	*	*	*	*
personal pronouns	*	√	√	(√)	√
definite descriptions	*	(√)	√	√	√
proper names	*	(√)	√	*	√
indefinite descriptions	*	*	(√)	(√)	√
deictic pronouns	*	*	√	*	√

Table 3.1: Distribution of anaphoric expressions

3.1.2 Syntactic and Semantic Notions of Anaphora

Deep vs. Surface Anaphora

There have been various attempts to classify anaphora in order to account for these distributional facts. A categorical distinction was made by Hankamer and Sag (1976) between *deep* and *surface* anaphora.

Surface anaphora are syntactically controlled and require as antecedent a linguistic expression of a particular syntactic form. Surface anaphora can be interpreted from a representation of the surface form of texts (e.g. Chomsky's LF), without being referentially

anchored. Usually, antecedent and anaphor occur both within a single sentence[1]. Typical examples for surface anaphora are (3.12) and (3.13).

(3.12) The queen$_1$ looked at her$_1$ mirror.

(3.13) Anyone$_1$ who eats that will lose his$_1$ hair. (Hankamer and Sag, 1976, p. 391)

Deep anaphora are either semantically controlled when they require an antecedent of a particular semantic type, or pragmatically controlled when there is no linguistic introduction at all. Deep anaphora are interpreted as referentially anchored to the real world or a representation of it. This type of anaphora occurs both intra- and intersententially. For example, in first sentence of the fairytale (3.1), repeated here, "he" is a deep anaphor and its antecedent is "a miller".

(3.14) There was once a miller$_1$ who was poor, but he$_1$ had one beautiful daughter.

The classification of Hankamer and Sag (1976) has been questioned by various scholars, *inter alia* Garnham and Oakhill (1996), who point out that the picture seems to be more complex. For instance, the null pronoun in (3.2), repeated below, is not realized at the surface, although it is syntactically controlled.

(3.2) To PRO$_1$ give himself$_1$ consequence, he$_1$ ...

Coreference vs. Coindexing

In generative grammar, syntactically controlled anaphora are a thoroughly studied subject. Especially in Chomsky's government and binding theory (Chomsky, 1981), the central notions are *binding* and the binding conditions. Büring (2005) delivered an excellent introduction to binding theory. He draws a detailed picture of differences between anaphoric relations. Basically, he distinguishes syntactic from semantic relations between noun phrases. Syntactic relations are coindexing and syntactic binding, and semantic relations are coreference and semantic binding.

From a syntactic perspective, anaphoric phenomena are identified by coindexing, i.e. antecedent and anaphoric expression bear the same index. Coindexed NPs are cases of *syntactic binding* if and only if one NP c-commands the other. An example for syntactic binding is (3.12), as well as (3.2) above.

(3.12) The queen$_1$ looked at her$_1$ mirror.

From a semantic perspective, (i) cases of anaphora without reference are distinguished from (ii) cases of anaphora involving reference. In the former case (i), Büring speaks of *semantic binding*. This is the case if the antecedent (the "binder") is a quantifying NP like in (3.13) above, repeated here, where "anyone who eats that" is not referential, i.e. there is no referent this noun phrase could refer to. Such anaphora show a similar behaviour to variable binding in logical languages.

1 Surface anaphora are not necessarily limited to sentence boundaries, and often involve the omission of a second occurrence of material that has already been spelled out, as the case of *ellipsis* in example (3.11).

(3.11) Someone else has got the job. I wonder who. (Garnham and Oakhill, 1996, p. 330)

(3.13) Anyone$_1$ who eats that will lose his$_1$ hair.

Of more interest in this thesis is the latter case (ii). Within this class of anaphora involving reference, coreference can be distinguished from non-coreference. This distinction was made already in chapter 1, section 1.3.2, where direct anaphora were told apart from indirect anaphora. Since this is a very important distinction in this thesis, it is illustrated here once more.

Anaphora involving coreference, e.g. (3.14) above, repeated here as (3.15), are *direct anaphora*.

(3.14) There was once a miller who was poor, but **he** had one beautiful daughter.

Anaphora involving non-coreference, e.g. between "the door" and "a room" in (3.5), repeated here, are *indirect* or *bridging anaphora*.

(3.5) When the girl was brought to him, he led her into a room that was quite full of straw, and gave her a wheel and spindle, and said, "Now set to work, and if by the early morning thou hast not spun this straw to gold thou shalt die." And he shut **the door** himself, and left her there alone.

Note that the two term pairs coindexing / coreference and syntactic / semantic binding do not have a one-to-one correspondence. Syntactic coindexing involves cases of both semantic binding and coreference, and there are cases of syntactic binding without semantic binding, although semantic binding implies syntactic binding. More details can be found in Büring (2005, ch. 4). Büring's classification is summarized in Table 3.2 (adapted from Büring, 2005, Table 4.1).

syntax	coindexing		non-coindexing
	syntactic binding	no syntactic binding	
semantics	semantic binding	coreference	non-coreference
	non-reference	reference	

Table 3.2: Syntactic and semantic relations between anaphor and antecedent (adapted from Büring, 2005)

In this thesis, I will deal with anaphora involving reference (this amounts roughly to Hankamer and Sag's deep anaphora), thus leaving aside cases of non-referential, purely syntactically controlled anaphora (i.e. Hankamer and Sag's surface anaphora). I will use the term *discourse anaphora* for direct and indirect anaphora. They can have antecedents in the same sentence, though typically they span bigger distances than sentences. Crucially, discourse anaphora involve reference to entities that exist independently from language. These entities, or referents, can be understood as elements of the extralinguistic reality. However, from a more cognitive point of view, referents can also be understood as elements of the conceptual system in the conscience of speaker and hearer. This issue takes us to the next section.

3.2 The Discourse Model

3.2.1 Conceptions of Discourse Models

A common view of cognitive processes in human thinking in general and in language understanding in particular relies on the notion of a *mental model*. In general, mental models can be thought of as representations of real world phenomena in human minds. A mental model emerges as a result of thinking processes. This idea has its antecedents in 19th century philosophy in the work of Charles Sanders Peirce (Peirce, 1903). In the 1970s, both psychological and formal research led to a resurgence of the conception of mental models, which had a great impact on many linguistic theories.

I will first present the psychological view on mental models, then briefly sketch the cognitive concept of a situation model, and finally look at models from a formal perspective. All these views converge on some essential features of discourse models as suitable representations of discourses.

Mental Models

Among psycholinguists (cf. *inter alia* Garrod and Sanford, 1982; Johnson-Laird, 1983; Garnham and Oakhill, 1990; Britton and Graesser, 1996), there is a broad consensus that during text comprehension, the reader builds multiple representations of texts. Often, a three-way distinction is made between a short-lived representation of exact linguistic material, a text base containing the propositions expressed by sentences and their interrelations, and a mental model where linguistic material is integrated with background knowledge. The central role in text comprehension plays the mental model.

In the theory of Johnson-Laird (1983, 2005), mental models are seen as the basic structure of cognition: "[...] mental models play a central and unifying role in representing objects, states of affairs" (Johnson-Laird, 1983, p. 397). The theory of Johnson-Laird relies on four principles, which I will briefly discuss.

- The principle of iconicity

Discourses are a way of experiencing the world in detail. It seems reasonable that when we comprehend a discourse, we try to build a model of the world corresponding to the discourse content. A mental model has a structure that corresponds to the known structure of the part of the world it represents. These models are iconic in the following way:

> "[Mental models] contain a token for each referent in the discourse, properties corresponding to the properties of the referents, and relations corresponding to the relations among the referents."
>
> (Johnson-Laird, 2005, p. 187)

Crucially, mental models are a mapping from the real world to an abstract representation of the world. It is important to note that entities in the mental model are distinct from real world entities on the one hand, and from linguistic expressions on the other hand.

Johnson-Laird states a second principle, which is closely related to Stalnaker's notion of possible worlds (cf. sections 2.1.1 above and 3.4.3 below).

- The principle of possibilities

Each mental model represents a possibility of how the content of a discourse is to be embedded in the world. Of course, often exist many possible ways the world could be in order to fulfill the presuppositions made by a discourse. Experimental evidence has shown that people normally build models of what is true, not what is false. As a consequence, people commit systematic fallacies in reasoning. Hence, this principle can be refined as follows: a mental model represents a true possibility, and it represents a clause in the premises only when the clause is true in the possibility.

- The principle of truth

This principle yields a considerable simplification in the conception of mental models when compared to a conception that takes all possibilities into account. Supported by psychological research, Johnson-Laird claims that reasoning is easier assuming just one model than assuming multiple models. People tend to focus on one of the possible models for a given phenomenon. This leads to irrational decisions and erroneous conclusions that possibly have to be withdrawn afterwards. This fact points to an important feature of inferences in discourse interpretation: defeasibility. Consequently, Johnson-Laird's last principle deals with the manipulation of mental representations, although its original formulation is quite vague.

- The principle of strategic variation

From exploring manipulations of models, reasoners develop a variety of strategies. A more detailed psycholinguistic discussion of inferences in mental models can be found in Garnham and Oakhill (1996).

Up to here, this is a very general conception of mental models in cognition, which can be easily applied to other cognitive domains than language, as well. The claim that mental models play a central role in language comprehension is supported by psycholinguistic research showing that people rapidly forget the surface form of sentences, keeping in mind only the facts which were talked about. Consider the following sentences from Gernsbacher (1996). A short time after reading sentence (3.15), many people incorrectly accept the sentence (3.16) as the sentence they read before.

(3.15) Three turtles rested on a floating log and a fish swam beneath them.

(3.16) Three turtles rested on a floating log and a fish swam beneath it.

An obvious explanation for this phenomenon is that both sentences express similar physical situations, and readers or listeners build mental models on the basis of the propositional content of uttered sentences rather than from their exact linguistic structure.

Johnson-Laird (1983, p. 165) makes a three-way distinction between propositions, i.e. "strips of symbols that correspond to natural language", mental models, i.e. "structural analogues of the world", and images, i.e. "the perceptual correlates of mental models from a particular point of view".

Mental models as described above are defined as a mental construction describing the knowledge a person has of a particular domain of the world. This definition corresponds to a general concept of background knowledge. But something more is desirable from a theory

of mental models: a tool for representing and formalizing the way background knowledge comes into play in human language processing.

Garnham and Oakhill (1990, 1996) elaborated the theory of mental models and presented an account that considers text comprehension as a process of constructing a model of the situation the text is about. In the construction of this model, information from different parts of the text is linked together and integrated in a mental representation that is different from a representation of the linguistic forms occuring in the text.

Situation Models

A cognitive theory of text comprehension was developed in the work of van Dijk and Kintsch (1983). The central notion, *situation model* is quite similar to Johnson-Laird's mental model. Three levels of representations are distinguished: verbatim forms, propositions, and situational models.

> "A situational model is the cognitive representation of the events, actions, persons, and in general the situation that a text is about."
>
> (van Dijk and Kintsch, 1983, p. 11)

Situation models are constructed by producing and interpreting linguistic expressions. Therefore, expressions can be thought of as bearing a procedural meaning. This view is taken in the theory of mental spaces of Fauconnier (1994).

A related concept to situation models is the idea of a scenario, a script (Schank and Abelson, 1977) or a frame (Fillmore, 1976). The notion *frame* will be introduced in detail in section 7.1.

As the theory of van Dijk and Kintsch (1983) is rooted in a cognitive linguistics setting, it is, despite the considerable amount of literature citing this work, a very informal framework. But, as van Dijk himself notes, some of the processes described by his theory may be modeled more explicitly in *formal* models of context. Van Dijk recognizes that

> "[...] if language users construct mental models of communicative events, some of the properties of these events may be accounted for in formal models that are more explicit theoretical frameworks for the structure of such models."
>
> (van Dijk, 2008, p. 109)

Formal Models

In formal semantic theories, the notion of a model can be found in model-theoretic approaches, which are characterized by the idea that the interpretation of linguistic material consists in evaluating it with respect to a model. In model-theoretic semantics, the semantics of a language has two parts: The first part is a definition of what a model or an interpretation for the language consists of: that is, minimally, a set of entities (the domain) and an interpretation function providing assertions about these entities. The second part consists of rules which determine truth conditions for all sentences of the language relative to such a model.

In general, in logically oriented theories of natural language semantics, two prevailing views on meaning can be distinguished (cf. Abbott, 1997). One is often referred to as 'realism', where the meaning of an utterance is taken to involve relationships between

linguistic expressions and things in the real world. This view on meaning can be understood model-theoretically if the entities in the domain of the model are taken to be real-world entities.

A second view on meaning is dubbed 'representationalism', 'conceptualism', or 'semantic idealism'. Here, meaning consists of relationships between linguistic expressions and mind-internal entities, or 'semantic representations', which exist in the consciousness of speaker and hearer. Model-theoretically, the entities in the domain are mental representations of real-world entities. These representations can be seen as being part of the mental model in Johnson-Laird's sense. Examples for representational semantic theories are DRT (Kamp and Reyle, 1993) (cf. section 3.4.3 below) or Bierwisch's two-level semantics (Bierwisch, 1983) (cf. section 1.1.1).

3.2.2 Requirements on Discourse Models

Although the conception of formal models applied in logically oriented semantic theories is considerably more abstract than Johnson-Laird's cognitively rich mental models or van Dijk's detailed models of situations, they all share the basic idea that text interpretation involves the construction of a discourse model, which contains objects and relations referred to in the text, but not the linguistic structure of the text itself. Be a model thought of as a purely formal device or as a rich cognitive representation of the world, the theories converge in the idea that a discourse is understood (in terms of psychology), or appropriate (in terms of pragmatics), or true (in terms of formal semantics), if it can be embedded in a model, or in other words, if it satisfies a model.

What are the requirements on a discourse model as a suitable representation of the content of a discourse? An essential feature of a discourse representation is its ability to change *dynamically* as a discourse proceeds. This is, obviously, a central property of dynamic semantic theories (Kamp, 1981; Heim, 1982; Groenendijk and Stokhof, 1984), which are characterized by the position that linguistic expressions are instructions to update the discourse model. I will turn to these theories in more detail in section 3.4.3. Also psycholinguists agree on this view, exemplified by the following quote.

> "While a reader proceeds through the text, the activation of concepts, facts, and events as part of a discourse representation fluctuates constantly."
>
> (Sanders and Spooren, 2001, p. 5)

I have already discussed in section 1.1 that linguistic information often does not fully specify a discourse model. Text representations can remain *underspecified*. This is a key feature in many recent works in logically oriented semantics, e.g. Bierwisch (1983); Bos (2001); Dölling (2005); Egg (2005).

This idea is not unfamiliar to psycholinguists:

> "A mental model goes beyond the literal meaning of the discourse because it embodies inferences, instantiations, and references."
>
> (Johnson-Laird, 1983, p. 245)

It is also recognized in cognitive linguistics:

> "Language does not itself do the cognitive building – it 'just' gives us minimal, but sufficient, clues for finding the domains and principles appropriate for building in a given situation. Once these clues are combined with already existing configurations, available cognitive principles, and background framing, the appropriate construction can take place, and the result far exceeds any overt explicit information."
>
> (Fauconnier, 1994, p. 3)

Thus, shallow linguistic information together with inferencing yields a mental representation. The basic, obvious, argument for this view is that the same linguistic expression can be interpreted differently in different contexts. For such a view to be complete, apart from a suitable representation of linguistic information, rich inferential mechanisms are needed in order to construct discourse models not only from linguistic input and the meaning of sentences, but also from contextual knowledge sources such as the discourse situation and general knowledge about the world.

As a summary of this section and as a general principle in this thesis, I assume that discourse interpretation involves incrementally constructing a structured mental representation of the discourse. In the words of Cornish (1999), the notion of a discourse model can be characterized as follows:

> "[The discourse model is] a constantly evolving representation of the entities, propositions, eventualities, properties, and states, as well as their interrelations, which are introduced into the discourse, or are assumed already to exist therein, at particular points."
>
> (Cornish, 1999, p. 150)

Crucially, in a successful interpretation of a discourse, all information, not only directly expressed but also indirectly inferred by means of pragmatic inferences, are part of the discourse model constructed by the hearer in course of interpretation.

3.3 Discourse Referents

3.3.1 Introducing Discourse Referents in the Discourse Model

Karttunen (1976) first introduced the notion *discourse referents* for conceptual entities in a discourse model representing persons or things in the world. Characteristically, the speaker can assume that the real-world entities discourse referents stand for are known to the hearer. The entities can be referred to later in a discourse by means of anaphoric relations, e.g. pronouns or definite NPs. Discourse referents stand not only for really existing objects, but also for other entities that are localizable in space and time, such as situations or eventualities (i.e. events, states, processes and actions). Furthermore, discourse referents can stand for concepts or only intensionally existing entities like 'unicorn' or 'Santa Claus', as well as abstract entities like entity types or kinds, types of eventualities, facts, circumstances, etc.

In Karttunen's dissertation, as well as in later works on the subject (e.g. Heim, 1982; Ariel, 1990), it is assumed that indefinite noun phrases are the canonical form of introducing new discourse referents. But there are many other ways of introducing discourse referents into the discourse model (cf. e.g. Cornish, 1999). They can be introduced

(i) by explicit linguistic means,

(ii) by extra-linguistic means, or

(iii) by implicit inferential means.

An explicit linguistic introduction (i) typically occurs in a presentational or existential construction involving indefinite NPs, proper names, demonstratives, or definite NPs. Introduction by non-linguistic means (ii) can be observed in spoken interaction, when an utterance of an expression of the above-mentioned types is accompanied by a gesture. Sometimes, there is only a gesture, e.g. a pointing gesture, without any explicit linguistic expression. Also, other kinds of indexical devices are possible[2]. Implicit inferential introduction (iii) occurs when discourse referents are not explicitly introduced but can be inferred by virtue of pragmatic inferences.

Once a discourse referent is established in the discourse model, it becomes available to be accessed later by anaphoric expressions. This crucial property of discourse referents, also known as availability, accessibility, or salience, is constrained by various factors. (Karttunen, 1976) designated with the "life span" of discourse referent the discourse segment in which an introduced discourse referent can be accessed. He distinguished permanent and temporary discourse referents.

Permanent discourse referents are introduced in referentially transparent contexts. They can be anaphorically accessed in subsequent utterances, e.g. (3.17).

(3.17) Bill saw a unicorn$_i$. The unicorn$_i$ had a golden mane. (Karttunen, 1976, p. 366)

Temporary or "short-term" discourse referents are introduced in referentially opaque contexts, which are created by modal operators such as negation, world-creating verbs, contrafactual constructions, and certain quantifiers. The life span of temporary referents is restricted to the duration of the referentially opaque context. For a modal operator, it is the length of its scope. (3.18) with a coreferential reading of the two occurrences of "unicorn" is not a well-formed discourse.

(3.18) Bill didn't see a unicorn$_i$. *The unicorn$_i$ had a golden mane. (Karttunen, 1976, p. 366)

Nevertheless, the opaque context can be extended afterwards by modally equivalent operators in later utterances, thus making subsequent anaphoric reference possible.

(3.19) Bill didn't see a unicorn$_i$. If he had, it$_i$ would certainly have a golden mane.

Crucially, Karttunen pointed out that the introduction of a referent and its availability for anaphoric reference are interrelated: Anaphoric reference is possible if and only if a referent has been introduced.

3.3.2 Accessing Discourse Referents as Antecedents for Anaphora

The presence or absence of discourse referents in a discourse model alone is not always strong enough to uniquely determine the antecedent of an anaphoric expression. In many

2 Here, index refers to Pierce's notion of index (Peirce, 1903).

texts, pronouns like "it" are used and unambiguously understood although more than one potential antecedent is present in the discourse model.

Many approaches to solve this problem have invoked a device that restricts the relevant context for the interpretation of an anaphoric expression. Unfortunately, in the numerous attempts to account for anaphor resolution there is no uniform definition of such a device. In this section, I will review the most important notions that have been proposed, including Prince's familiarity, Gundel's givenness, Ariel's accessibility, the Prague school's salience, and psychological activation.

3.3.2.1 Familiarity

Familiarity means that a discourse referent in question must be known to the hearer. A well-known Familiarity scale was given by Prince (1981), distinguishing between *new* (brand-new or unused), *inferrable*, and *evoked* (textually or situationally). This scale is depicted in Fig. 3.1.

brand-new $<$ unused $<$ inferrable $<$ textually evoked $<$ situationally evoked

Figure 3.1: The familiarity scale (Prince, 1981)

In later work, Prince (1992) made another distinction based on a cross-classification between *discourse-old* vs. *discourse-new* on the one hand, and *hearer-old* vs. *hearer-new* on the other hand. This classification is given in Table 3.3.

	discourse-new	discourse-old
hearer-new	brand-new	-
hearer-old	unused	evoked

Table 3.3: Discourse states of referring expressions (Prince, 1992)

Prince mentions that a class of discourse referents, which she calls *inferrables*, does not fit into this scheme.

> "[D]iscourse entities may be of a third category, Inferrable, where they are technically Hearer-new and Discourse-new but depend upon beliefs assumed to be Hearer-old, and where these beliefs crucially involve some trigger entity, which is itself Discourse-old, and where they themselves are being treated as though they were Hearer-old and possibly also Discourse-old".
>
> (Prince, 1992)

This class will be the main concern of chapter 6.

3.3.2.2 Givenness

Gundel et al. (1993) proposed a hierarchy of six cognitive statuses of discourse referents with the aim to explain the use of referring expressions (cf. Fig. 3.2). *In focus* means that the referent is in the center of attention, and thus anaphoric reference via pronouns ("it") is allowed. *Activated* means that the referent is present in short-term memory although not in

the center of attention. Hence more complex anaphoric expressions must be used ("that" or "this" or "this N"). Similar definitions are given for the other statuses. Each of the statuses is meant to be a necessary and sufficient condition for the use of its corresponding form. The statuses are related by implication, that is by using a particular form, the speaker implies that the statuses of the right of the used form are also fulfilled.

in focus ⊃	activated ⊃	familiar ⊃	uniquely identifiable ⊃	referential ⊃	type identifiable
it	*this/that/this* N	*that* N	*the* N	indefinite *this* N	*a* N

Figure 3.2: The givenness hierarchy (Gundel et al., 1993)

The difference to Prince (1981)'s Familiarity scale is that the statuses in the Givenness hierarchy are ordered by an entailment relation, while the statuses in the Familiarity scale are mutually exclusive.

3.3.2.3 Accessibility

According to Ariel (1990, 2001), what is linguistically expressed signals the referent's degree of *accessibility* in memory. Accessibility is claimed to correlate with memory structure and is defined in terms of four primitives:

- *Distance* between antecedent and anaphoric expression,
- *Competition* between possible alternative antecedents,
- *Salience* of the referent in terms of topicality,
- *Unity* of the frame/scenario between the anaphor and its antecedent.

Ariel does not explicitly distinguish between three types of context, as made in Clark and Marshall (1981)[3]. Instead, she claims that referential expressions, or as she calls it, accessibility markers, code a particular degree of mental accessibility. They obey the accessibility marking hierarchy (cf. Fig. 3.3). The same hierarchy is assumed to be valid in all languages although particular markers can differ in their functions. The accessibility scale ranges from pronouns, which are highly accessible and low in informativity (typically retrieved from linguistic context), over demonstratives and indexicals, which are intermediately accessible (e.g. from situational context), to proper names and highly informative definite descriptions that are rigid in designation and thus low in accessibility (typically retrieved from world knowledge).

A particular referential expression is chosen according to the assessed degree of accessibility of the discourse referent corresponding to it. The degree of accessibility depends on the inherent salience of the referent within the discourse model and on the unity between antecedent and anaphor.

With respect to inferred entities, or inferrables, Ariel states that they "come in different degrees of accessibility" (Ariel, 2001, p. 33). Entities that can be assumed to be evoked by a particular scenario or frame (e.g. the waiter in a restaurant) are more accessible than

3 cf. section 2.1.1 above.

⟵ high			accessibility		low ⟶
null	clitic pronouns	stressed	demonstratives	definite NPs	proper names

Figure 3.3: The accessibility scale (Ariel, 1990)

inferrable entities that are not salient or necessary in a specific frame (e.g. an umbrella in a restaurant).

The principles governing accessibility are assumed to be a universal cognitive device. Some aspects of it found their way into the grammar via language-specific grammaticalizations of accessibility markers but other aspects follow from an inferential interaction of linguistic and conceptual knowledge. Ariel (2001, pp. 40f) admits that accessibility theory can not exclusively account for referential choices, only for a default choice of referential expressions which can possibly be overridden by pragmatic inferences such as conversational implicatures. The lexical meaning of a referential expression gives clues to its accessibility; however, accessibility is finally determined by context.

3.3.2.4 Salience

Salience-based approaches are based on the cognitive ability to dynamically change the focus of attention. Discourse participants draw their attention to that part of knowledge which is relevant in a given situation. The Prague school of information structuring (cf. Kruijff-Korbayová and Hajičová, 1997) models the attentional state of discourse participants in terms of the *salience* or activation of entities in the "stock of shared knowledge".

The stock of shared knowledge corresponds to the Common Ground (cf. chapter 2) and is a dynamically defined set of discourse referents, which are partially ordered according to their salience. The salience of a discourse referent reflects its immediate accessibility in the hearer's memory. The salience scale ranges from "highly salient" to "faded away". The degree of activation is highly dependent on linguistic form and the topic-focus articulation of sentences (cf. Hajičová et al., 1998).

In the Praguian model, sentences are divided into a topic and a focus part. The *topic* is the part of the sentence structure that is being presented by the speaker as readily available in the hearer's memory. The *focus* is what is being asserted about the topic.

Note that the Praguian concept of focus is not identical with the concept of focus in Sidner's focus theory (section 3.4.2.1). The term *focus* in the Prague school corresponds to what is known in other theories of information structure as "comment", or "rheme", or new information[4]. Likewise, *topic* is also known as "theme" or given information[5].

It is assumed that salience and topic-focus articulation are represented on a separate "tectogrammatical" layer of language structure, where every element is either contextually bound or nonbound. Contextually bound items roughly correspond to the topic of the sentence; the rest constitutes the focus. For example, the topic-focus partition of the answer in the following examples depends on the particular utterance context.

4 Note that focus should not be confused with an intonationally highlighted item. Intonational highlighting is merely a means of indicating focus when it is in a marked position. Another way of marking focus is word order. For more details, see Hajičová et al. (1998).

5 see also section 4.3

(3.20) Q: What about dogs?

A: [$_{Topic}$ Dogs][$_{Focus}$ must be CARRIED.]

(3.21) Q: What must be carried?

A: [$_{Focus}$ DOGS][$_{Topic}$ must be carried.]

(3.22) Q: What must we do in order to take the metro?

A: [$_{Focus}$ DOGS must be carried.]

In an upshot, the approach accounts for anaphoric reference as follows. Non-salient entities can be referred to only in the focus part of the sentence, while activated entities can be referred to in both the topic and the focus part of a sentence. The essential idea is that the degree of activation of entities in the stock of shared knowledge depends on whether an entity is referred to in the topic or the focus part of an utterance. The activation of discourse entities can change as the discourse progresses. The relationship between the form of referring expressions and the salience of the corresponding discourse referents is determined by principles of cooperative communication.

3.3.2.5 Activation

In the psychological and psycholinguistic literature, the terms *activation* and *activation pattern* can be found frequently. Numerous models of text comprehension making use of one or another form of these notions have been proposed. We have already seen the theories of mental models (Johnson-Laird, 1983; Garnham and Oakhill, 1990); other theories include the structure building model (Gernsbacher, 1996), and the landscape model of reading (van den Broek et al., 1996; Gaddy et al., 2001).

In this section, the notion of activation is illustrated by sketching the landscape model of reading, a theory that is based on the idea of mental models (cf. section 3.2) in earlier cognitive research on attention in comprehension (van Dijk and Kintsch, 1983).

The process of reading is conceptualized as a cyclic process. With each consecutive reading *cycle*, a new text segment and its constituent concepts are processed. New information enters the working memory (the attentional stack) and is activated.

> "[O]ver the course of reading, individual concepts fluctuate in their activation as the reader proceeds from cycle to cycle: some concepts come into the focus of attention, others fade, and yet others remain in working memory but fall and rise in the level of their activation."
>
> (Gaddy et al., 2001, p. 90)

The "reading landscape" consists of a matrix with 3 dimensions. Values in the first horizontal dimension correspond to reading cycles. The second horizontal dimension is formed by the major concepts relevant to the text. On the vertical dimension, the degree of activation of concepts in each cycle is plotted. The patterns of activation across reading cycles form the basis for the memory representation that the reader constructs of the text, i.e. the discourse model.

A further feature of this theory is that when two concepts are activated simultaneously, a connection is built between them in the reader's episodic memory. The strength of this connection is a function of the amount of activation of concepts involved. The evolving representation affects subsequent reading cycles: activated concepts in the current cycle elicit potential activation of the concepts that have become associated to them in the current representation of background knowledge ("cohort activation").

The fluctuation in activation (from 0 to 5 points) results from the application of 5 rules (van den Broek et al., 1996, p. 171):

1. explicitly mentioned concepts get 5 points,

2. concepts required to restore cohesion[6] get 4 points,

3. concepts contributing to a causal explanation get 4 points if they are proper causes, or 3 points if they merely enable a causation,

4. statements may elicit associative elaborations, i.e. inferences drawn from background knowledge, and get 2 points, and finally

5. unless an activated concept is repeated or reinstated, its activation returns to 0.

The landscape model provides a vivid picture of a discourse model. However, it is by no means clear in what manner lexical expressions evoke concepts, and how do we determine what are the "major concepts relevant to a text". The assignment of numbers to activations seems rather ad hoc, although probabilistic values could be retrieved from applying the model to large corpora. Furthermore, an associative activation pattern as predicted by the theory does not explain how the meaning of texts is integrated with contextual knowledge. The model does not explain influences on interpretation and anaphora resolution beyond the lexical level.

3.3.3 Conclusion

This section introduced the basic term of a discourse referent as standing for an entity in the discourse model. To be underlined is the ontological status of discourse referents to be located on a representational level, distinct from both real world entities and linguistic expressions. The essential reason for assuming discourse referents in the first place is that they can serve as antecedents for discourse anaphora. Consequently, I have tried to shed some light on the numerous notions used to explain why some discourse referents are better antecedents than others. We have seen that the availability of referents for anaphoric reference can be modelled in terms of focus, familiarity, givenness, accessibility, or salience. None of these concepts gives a thoroughly satisfying account of anaphoric reference. A theory which wants to account for discourse interpretation and, as a consequence, for anaphora resolution, can rely on any of these notions, although, obviously, it must go further and take additional considerations into account.

6 van den Broek et al. (1996) speak of "anaphoric coherence" for what I have called "cohesion".

3.4 Theories of Anaphora Resolution

This section presents theories of anaphora resolution from different points of view. First, I will briefly describe pragmatic constraints in in a neo-Gricean framework. Then, I will review accounts rooted in computational linguistics. I will sketch Sidner's Focus Theory and describe its successor Centering Theory. Finally, I will close this section and the chapter with an introduction to Discourse Representation Theory (DRT), emphasizing the treatment of anaphora in this theory.

3.4.1 A Pragmatic Account

A pragmatic account of anaphora interpretation in a neo-Gricean framework (cf. section 1.1.2.2) was presented by Huang (1994). His theory of anaphora starts from the assumption that anaphora interpretation can be explained by a small number of general pragmatic interpretation principles, which are expressed as a combination of neo-Gricean conversational implicatures. These principles have to be further constrained by a set of consistency constraints which organize their interation. In short, the theory can be summarized as follows:

> "In this theory, anaphora is largely determined by the systematic interaction of two neo-Gricean principles, namely the M- and I-principles (in that order of priority), constrained by a Disjoint Reference Presumption[7], information saliency and general consistency conditions on conversational implicature."
>
> (Huang, 1994, p. 115)

Huang argues that the following general pattern, which can be observed in the behavior of anaphora, is an instantiation of a systematic interaction of neo-Gricean pragmatic principles.

> "Reduced, semantically general anaphoric expressions tend to favour locally coreferential interpretations; full, semantically specific anaphoric expressions tend to favour locally non-coreferential interpretations."
>
> (Huang, 1994, p. 119)

This pattern is illustrated by (3.23) and (3.24). A "vehicle" is semantically more general than a "bus", because "bus" is a hyponym of "vehicle". On the one hand, when "vehicle" follows "bus", a locally coreferential interpretation is encouraged, as in (3.23). On the other hand, when the ordering is reverse, a locally non-coreferential interpretation is favoured, as in (3.24), just as predicted by the general pattern of anaphora.

(3.23) a. The bus$_1$ came trundling round the bend.

b. The vehicle$_1$ almost flattened a pedestrian.

(3.24) a. The vehicle$_1$ almost flattened a pedestrian.

b. The bus$_2$ came trundling round the bend. (Huang, 1994, p. 119)

7 Disjoint Reference Presumption: The arguments of a predicate are intended to be disjoint, unless marked otherwise (Huang, 1994, p. 129).

Here, I do not want to go into more details on the interaction of pragmatic principles in a neo-Gricean theory. Instead, I will focus on the constraints at work. Huang sees anaphora resolution as an instance of a generalized conversational implicature (cf. section 1.1 above). The specification of an anaphoric relationship can be overridden when it turns out to be inconsistent with one of the following factors (as cited by Blackwell 2003, p. 39ff):

1. background assumptions or world knowledge,
2. the presumed speaker's intention according to the assumed state of mutual knowledge (Grice's meaning-nn),
3. semantic entailments, and
4. what is relevant or salient.

The Background Knowledge Constraint
Pragmatic inferences in general, and a specific anaphora resolution in particular, can be cancelled when they run contrary to background assumptions or world knowledge. This constraint includes general encyclopedic knowledge about entities in the world, properties and relations between them, the occurence of events and actions, causal relations, etc. Drawing on our knowledge of previous experiences we have assumptions and expectations about specific situations. In a given situation, some events and actions can be more likely to occur than in another situation. We can imagine specific scripts or schemas of stereotypical situations or scenes. Fillmore (1976) uses the term *frame* to refer to a system of linguistic choices that can get associated with prototypical instances of scenes, coherent segments of human actions, beliefs, experiences or imaginings. This idea will be presented and exploited for anaphora resolution in chapter 7.

The Mutual Knowledge Constraint
According to Huang (1994), a second constraint deals with intentions and beliefs of conversational participants. Conversational implicatures, and in particular anaphora resolution, must be consistent with what the speaker might clearly intend given the assumed state of mutual knowledge (cf. section 2.1.1 above). Note that the notion of mutual knowledge on a shared basis incorporates shared background knowledge. Seen in this way, this constraint subsumes the background knowledge constraint.

Semantic Entailments
Huang's third constraint is concerned with semantic entailment, which is usually defined as follows: p semantically entails q iff in all worlds in which p is true, q is true. This constraint is violated in cases of logical inconsistency like the following:

(3.25) # John$_i$ is a musician. He$_i$ is not a musician.

Semantic entailments can also depend on general ontological properties of kinds and entitites. Consider the following Spanish example from Blackwell (2003).

(3.26) a. Juan habló con su esposa$_i$. La mujer$_{ij}$ le dio un beso.
Juan speak.PAST with his wife. The woman him give.PAST a kiss
'Juan spoke with his wife$_i$. The woman$_{ij}$ gave him a kiss.'

b. Juan habló con la mujer$_i$. Su esposa$_j$ le dio un beso.
Juan speak.PAST with his wife. The woman him give.PAST a kiss
'Juan spoke with the woman$_i$. His wife$_j$ gave him a kiss.'

In (3.26a), the antecedent "esposa" ("wife") semantically entails the anaphoric expression "mujer" ("woman"), because the first one is a hyponym of the second. Hence, "la mujer" is preferently interpreted as coreferring with "su esposa". In contrast, in (3.26b), the antecedent "mujer" does not semantically entail the anaphoric expression "esposa", thus forbidding a coreferential reading.

As mentioned above, this kind of semantic entailments relies on general semantic ontologies. There are different answers to the question if this knowledge is part of the lexicon, e.g. in form of semantic features of lexical entries, or rather to be regarded as extralinguistic knowledge of the world. If the latter view is taken, semantic entailments are a special case of background knowledge.

The Antecedent Salience Constraint

Pragmatic inferences can be overridden when there is inconsistency with what is relevant or salient. Specifically for pronominal anaphora resolution, the intended referent must be the most salient potential referent in the discourse universe at the moment of the utterance. In the last section (3.3.2), a row of competing notions of salience was discussed.

Huang (1994) points out that listeners tend to favour as antecendent of pronouns topics over subjects, subjects over objects, and objects over any other noun phrases. Blackwell (2003) reports experimental evidence from Spanish regarding this preference. Consider the following example. Since Spanish is a pro-drop language, personal pronouns need not to be spelled out explicitly. Nevertheless, the reference of the nonarticulated subject of (b) must be resolved.

(3.27) a. Al llegar a casa Juan$_i$ besó a su$_i$ mujer$_j$.
Upon arrive at home Juan kiss-PAST.3sg to his wife
'Upon arriving at home, John$_i$ kissed his$_i$ wife$_{ij}$'

b. y entonces $\emptyset_{ij}$ se puso a preparar la cena.
and then [he/she] REFL start-PAST.3sg to prepare the dinner
'and then [he$_i$/she$_j$] started to prepare dinner.'

In a multiple choice questionnaire, 88 out of 105 subjects answered the question "¿Quién preparó la cena?" ("Who prepared the dinner?") by "Juan", while 11 selected "su esposa" ("his wife") as antecedent. Thus, in absence of additional background information, e.g. the knowledge that Juan hates cooking, listeners tend to choose the most salient potential antecedent – in this case the subject. In sum, although possibly be overridden by other consistency constraints, antecedent saliency is an important factor in the choice of antecedents of anaphoric expressions.

Based on Huang's proposal, Blackwell (2003) summarizes the consistency constraints on anaphora resolution as follows: apart from grammatical conditions (agreement, c-command and other binding constraints), anaphora resolution is constrained by (i) antecedent salience, (ii) mutual knowledge (including (presumably shared) background knowledge) and (iii) semantic constraints (lexical constraints, semantic entailments).

Pragmatic principles and constraints certainly play an important role in anaphora resolution and can explain many anaphoric phenomena. However, since these principles, as stated in pragmatic theories of anaphora interpretation, are far from being easily formalized or implemented, let us turn now to formally more explicit theories of anaphora resolution. Nevertheless, the discussion of pragmatic constraints will be continued in section 7.3 in connection with the resolution of bridging anaphora.

3.4.2 Computational Accounts

3.4.2.1 Focus Theory

The Focus Theory developed by Sidner (1981) provides an algorithm that is aimed at predicting the antecedents of referring expressions. The most salient entities in an utterance are the preferred antecedents for the resolution of anaphora in a subsequent utterance. If these fail to be possible antecedents, a suitable antecedent must be chosen from the set of referents introduced in the last utterance.

The *focus* is the discourse referent around which the attention of the speaker is centered. Two types of attentional centers, or foci, are defined, *agent focus* and *discourse focus*. What constitutes the focus is determined by syntactic clues (subject, object, clefts, existential constructions) and thematic information (agent, theme).

The idea of two foci in one utterance is suitable to account for ambiguous anaphora. In these cases, each possible antecedent corresponds to one type of focus. However, this concept can be generalized and more than two types of attentional states can be assumed, as we will see in the following sections. Moreover, as Sidner's proposal is procedural, it is desirable to rely anaphora resolution on independently motivated principles.

An account of bridging anaphora resolution based on a combination of Focus Theory with DRT was presented by Freitas (2005). This approach will be reviewed in section 6.3.2.2.

3.4.2.2 Centering Theory

Centering Theory (Grosz et al., 1995; Brennan et al., 1987; Walker et al., 1998) is a framework for modeling cohesion in discourses. It examines the relationship between the focus or the center of attention and the determination of the form of referential expressions. In Grosz and Sidner (1986)'s general model of discourse structure (cf. section 1.3.4), it can be seen as a specification of the level of attentional state, or – as the authors call it – the *local* coherence within a discourse segment. Centering Theory formally analyzes a discourse participant's choice of referential expressions. In particular, pronominalization is a means to draw the focus of attention towards particular entitites in a discourse. Thus, Centering Theory is suited for the analysis of anaphoric relations in discourses.

The center of attention of an utterance is an abstract entity that connects an utterance with other utterances. Centering Theory divides the center of attention into a backward-looking center and a forward-looking center. An utterance U has at most one *backward-looking center* Cb, which constitutes a connection to the previous utterance, while the same utterance can have more than one *forward-looking centers* Cf, which are the entities that are potentially salient in subsequent utterances. Thus, Cf is defined as a partially ordered list of forward-looking centers. Cf contains all entities which are referrred to in an utterance U. The position of an entity in this list indicates the probability with which

it will be the focus of the following utterance. The highest-ranked element in this list is called the *preferred center* Cp, which can be seen as a prediction regarding the Cb of the next utterance. These properties can be summarized by a small number of constraints as follows (cf. Grosz et al., 1995, p. 210):

Centering Constraints For every utterance U_i holds:

1. there is at most one $Cb(U_i)$,[8]
2. every element of $Cf(U_i)$ must be realized in U_i,
3. $Cb(U_i)$ corresponds to the highest-ranked element of $Cf(U_{i-1})$ that is realized in U_i.

The first constraint tells us that in an utterance there is a central entity which the utterance is about.

The second constraint depends on the definition of the notion *realize*. The realization of an entity is an operation that generates from a mental representation (i.e. an entity in the discourse model) a linguistic expression. This relation can be seen as the inversion of constructing a mental representation from a linguistic expression. The speaker decides which linguistic expression to choose: a pronoun, a definite description, or a proper name.

The third constraint presupposes a ranking of the elements in Cf. This ranking is a central parameter of Centering Theory (Walker et al., 1998): it can vary across languages. Various proposals have been made, among them the word order on the sentence surface, thematic roles, or grammatical status. For English, Brennan et al. (1987, p. 156) have suggested an ordering according to grammatical function: the entity realized as subject is most salient, followed by the objects, and finally adjuncts. In languages with relatively free word order, other criteria have to be taken into account. For German, Strube and Hahn (1999) have proposed an ordering according to the functional information structure of a sentence.

Crucially, the third constraint expresses the locality condition of Centering Theory: the backward-looking center of an utterance must be an element of the list of forward-looking centers of the immediately preceding utterance. This principle guarantees that in a maximally coherent discourse segment, there is an entity that keeps being salient. Although certain deviations are allowed, they diminish the local coherence of the discourse. These deviations, or *center transitions*, depend on whether the Cb is maintained or changed from one utterance to another. According to Brennan et al. (1987), four possible transitions can be distinguished, as shown in Table 3.4.

	$Cb(U_i) = Cb(U_{i-1})$ or $Cb(U_{i-1}) = \varnothing$	$Cb(U_i) \neq Cb(U_{i-1})$
$Cb(U_i) = Cp(U_i)$	CONTINUE	SMOOTH-SHIFT
$Cb(U_i) \neq Cp(U_i)$	RETAIN	ROUGH-SHIFT

Table 3.4: Centering transitions

8 Obviously, the first utterance in a discourse has no Cb.

With respect to the choice of referring expressions and preferences among transition states, Centering Theory defines two basic rules (Grosz et al., 1995, p. 214):

Centering Rules For every utterance U_i holds:

1. if an element of $Cf(U_{i-1})$ is realized in U_i as a pronoun, then $Cb(U_i)$ must be realized as a pronoun, as well,
2. the transition states between utterances are ordered:
   Continue > Retain > Smooth-shift > Rough-shift.

Rule 1 is based on the assumption that a discourse entity realized by a pronoun is more salient than an enity realized by a definite description or a proper name. If there are more than one pronoun in an utterance, then its Cb must be realized as a pronoun. If there is only one pronoun, then it is the realization of the Cb.

The preferences of rule 2 allow to judge a discourse that maintains an entity in the center of attention as more coherent than a discourse that does not so. Firstly, if the Cb is maintained, we have a Continue or a Retain, if not, we have a Shift. Secondly, the transition type depends on whether Cb is identical with the Cp of the preceding utterance. Thus, if the Cbs of two consecutive utterances are distinct and the new Cb is different from the old Cp, then we have a Rough-shift. In this case, a less salient entity gets into the center of attention, rendering the discourse as less coherent.

The combination of Centering rules and constraints allow a precise prediction about the local coherence (i.e. cohesion) within a discourse segment. Let us illustrate this with some examples.

(3.28) a. John has been acting quite odd recently.

b. He called up Mike yesterday.

c. He wanted to meet him urgently.

This often cited example is quite coherent. John is introduced in the first utterance and is kept in the center of attention. The transition states between (a) and (b) and from (b) to (c) are both Continue. The centers in this discourse are given in Table 3.5.

Utterance	Cb	Cf	Transition
(3.28a)	$\varnothing$	$\langle[John]\rangle$	
(3.28b)	$[John]$	$\langle[John], [Mike]\rangle$	Continue
(3.28c)	$[John]$	$\langle[John], [Mike]\rangle$	Continue

Table 3.5: Centers for discourse (3.28)

This discourse shows how Centering Theory accounts for the resolution of anaphoric pronouns: “he” in (c) is correctly predicted to refer to John, while “him” is interpreted as referring to Mike. Now, if this discourse is slightly changed by replacing (c) by (c’), it is commonly perceived as less coherent, although neither Cbs nor Cfs are changed.

(3.28) c’. #John wanted to meet him urgently.

Centering Theory accounts for the oddness of this example because rule 1 is violated in (c'). Mike is referred to via the only pronoun in the utterance without being the backward-looking center.

The following discourse exhibits a phenomenon commonly known as a "garden path effect".

(3.29) a. Max was thinking if he had taken everything into account for the preparation of the party.

b. He had forgotten to invite Moritz.

c. He called him up.

d. #He had planned nothing for the evening and wanted to come.

The centers and transition states are given in Table 3.6.

Utterance	Cb	Cf	Transition
(3.29a)	$\varnothing$	$\langle[Max]\rangle$	
(3.29b)	$[Max]$	$\langle[Max], [Moritz]\rangle$	CONTINUE
(3.29c)	$[Max]$	$\langle[Max], [Moritz]\rangle$	CONTINUE
(3.29d)	$[Moritz]$	$\langle[Moritz]\rangle$	ROUGH-SHIFT

Table 3.6: Centers for discourse (3.29)

As can be seen from this table, there is a ROUGH-SHIFT from (c) to (d), violating the preferences of rule 2. Although Max is expected to be in the center of attention, the pronoun "he" in (d) can only refer to Moritz, thus producing the garden-path effect that is misleading the addressee until the end of the utterance. Centering Theory correctly predicts the oddness of this example.

There are several ways of implementing the Centering principles. A first comprehensive procedural algorithm was presented by Brennan et al. (1987). However, as this account strictly adheres to the principles as presented here, it obeys the above mentioned locality condition, which constrains the theory to anaphoric relationships between two subsequent utterances. The search for possible antecedents of anaphoric expressions is restricted to the immediately preceding utterance. This constraint leads to failures in discourses like (3.30) , where no conclusive statement about the local coherence can be made. Centers and transitions are given in Table 3.7.

(3.30) a. Peter likes Paul's Canelloni.

b. Yesterday, he ate twenty of them.

c. Paul is an excellent cook.

Here, the intervening utterance (b) does not mention Paul, so he falls out of the Cf list. However, he is in the center of attention in (c). Although this discourse is not very odd, Centering Theory, as it is presented above according to Grosz et al. (1995) and its implementation by Brennan et al. (1987), cannot interpret it. However, Centering Theory

Utterance	Cb	Cf	Transition
(3.30a)	$\varnothing$	$\langle[Peter], [Paul]\rangle$	
(3.30b)	$[Peter]$	$\langle[Peter]\rangle$	CONTINUE
(3.30c)	?	$\langle[Paul]\rangle$	?

Table 3.7: Centers for discourse (3.30)

was originally not conceived as a standalone theory, but rather as part of a more global coherence theory such as Grosz and Sidner (1986)'s, within which its task was to explain coreference relations on a local level. Consequently, in case that the global theory provides a way to account for long-distance anaphora, the principle of locality of Cb could be maintained on the local level of attentional state.

An extension that overcomes this constraint within a local theory of attentional state was proposed by Strube and Hahn (1999), who adapted Centering to German texts based on the information structure of sentences, taking the assumed familiarity hierarchy of Prince (1992) (cf. section 3.3.2.1) into account, and as a consequence being able to account for relationships between entities with longer distances.

A recast of Centering in Optimality Theory (cf. Prince and Smolensky, 1993) was developed by Beaver (2004). This account offers a declarative formulation of Centering constraints and exhibits a greater flexibility by allowing certain violations of the principles.

Centering Theory, in spite of its limitations, has had a great perception in various subfields of linguistics. Due to its exact predictions for pronoun resolution, it is well suited to be adopted as part of more comprehensive discourse theories. An example for such an attempt is Veins Theory (Cristea et al., 1998), a combination of Centering Thory with a theory of discourse structure, RST, which I will present in section 4.2.3 in the next chapter.

3.4.3 Dynamic Semantics and Discourse Representation Theory

In this section, I will present a formal representation of discourse referents within a discourse model according to a dynamic approach to formal semantics, which takes the discourse context into account and thus accounts for certain anaphoric relationships between discourse referents.

In the classical formal semantic theory of Montague (1973), the interpretation of utterances consists in assigning them a truth-conditional content. In terms of Stalnaker (1974)'s conception (cf. section 2.1.1) of possible worlds, the meaning of a sentence is the characteristic function of the set of worlds in which its truth conditions are met. If the discourse structure is determined only by the syntactic structures of its constituting sentences, then the meaning of a discourse D is a set of worlds obtained by the intersection of the meaning of sentences $S_1, S_2, ..., S_n$, i.e. $[\![D]\!] = [\![S_1]\!] \cap [\![S_2]\!] \cap ... \cap [\![S_n]\!]$.

3.4.3.1 Context as Index

In his paper on demonstratives, Kaplan (1977) sketched a theory of context in the model-theoretic tradition of formal semantics. The context of an utterance is a set of features of the world. The characteristic features for an utterance are that it has a speaker and a recipient, and that it is localizable in space and time. A *Kaplan Context* is a tuple

$c = \langle c_S, c_H, c_T, c_P, c_W \rangle$ with the speaker c_S, the recipient c_H, the utterance time c_T, the utterance location c_P, and the world of the utterance c_W.

An essential feature of Kaplan's theory is the distinction between content and character. The *character* of an expression is the function which returns at each index its *content*, that is the intension or the sense of this expression. The character is determined by linguistic conventions, thus part of context-independent meaning.

The *content* of an expression in a particular context is the intension defined by the character. For sentences, it is a proposition, and for indexical expressions, it is an individual concept. If the same content is to be expressed in different contexts, indexical expressions have to be adapted. For instance, if somebody wants to utter today the same as yesterday, she must replace the occurrences of "today" with "yesterday". The result of evaluating the content is its extension: for sentences, it is a truth value, and for indexical expressions, it is the proper referent itself. The content can be seen as a function from possible worlds into extensions.

Sag (1981) made a proposal to extend Montagovian model-theoretic semantics by Kaplan contexts. His major contribution was to take into account not only demonstratives and other indexical expressions but also context-dependent meaning shifts such as metonymy. Technically, he extends Kaplan contexts by additional functions c_f, which map demonstratives to individuals, and sense transfer functions c_{ST}, which map senses of predicates to shifted predicate senses. For example, in (3.31), c_f maps "this" to the traffic light, and in Nunberg's example (3.32), c_{ST} maps the sense of ham sandwich to the orderer of the ham sandwich.

(3.31) This is red.

[accompanied by a pointing gesture towards a traffic light.]

(3.32) The ham sandwich has left without paying.

However, these functions have to be constrained by information that is not part of linguistic knowledge. The theory does not give any clues to determine when these functions apply and when not. The non-linguistic context has to provide them, and a pragmatic theory is necessary to supplement the formal semantic theory.

In the mentalistic view of Bierwisch (1983)'s two-level semantics (cf. chapter 1), the character corresponds to the layer of semantic form, SF. The content corresponds to the level of conceptual structure, CS. The difference is that while CS is a mental representation of the world, Kaplan's content, viz. intension, is a function, which can be seen as corresponding to an objective idea, i.e. something that is independent from human cognition.

3.4.3.2 Dynamic Semantics

The emphasis of Montagovian semantics lay on the meaning of single sentences, leaving out, apart from contextual phenomena such as indexical expressions, relationships between sentences. If a sentence contains an anaphor, its reference has to be resolved in order to determine the truth value of the sentence. This works well if the antecedent is in the same sentence, but not if its antecedent is not in the same sentence. Consider (3.33).

(3.33) A man entered. He smiled.

A Montague grammar would represent the first sentence as an existentially bound formula and the second sentence into an independent formula such that the pronoun "he" is represented as a free variable, without imposing constraints on its binding. But crucially, in this example, "he" unambiguously refers to the entity introduced by "a man" in the first sentence. Even if we assume that a Montagovian grammar could be extended to cope with this kind of intersentential anaphora (but see Kamp et al., 2005, p. 7 for counter-arguments), there are more serious anaphoric phenomena which do not permit a straightforward solution within Montagovian semantics. There are cases in which the binding of pronouns is blocked. In the second sentence in (3.34), "it" cannot be interpreted as coreferential with "a donkey".

(3.34) Every farmer who owns a donkey$_i$ beats it$_i$. *It$_i$ lives in a stable.

Dynamic semantic theories emerged in order to solve this kind of problems. These approaches, unlike Montague, consider the meaning of a sentence as dependent on the discourse context. Meaning is seen as a relationship between an input context, which consists of the discourse content before the sentence was uttered, and an output context, which extends the input context by the content of the utterance. This context in turn is the input context for subsequent utterances. Heim (1982) called this relational notion of meaning the *context change potential* of a sentence. She proposed to see the context as a set of file cards where discourse information is stored and updated. In general, context in this view is a world-assignment pair, and discourse meaning is a relation between contexts. Hence, the meaning of a discourse D consisting of constituent sentences $S_1, S_2, ..., S_n$ is determined by consecutively connecting the sentences to the prior context, i.e. $[\![D]\!] = [\![S_1]\!] \circ [\![S_2]\!] \circ ... \circ [\![S_n]\!]$. The main difference to Stalnaker's conception is the fact that the order of sentences matters: $[\![S_i]\!] \circ [\![S_j]\!] \neq [\![S_j]\!] \circ [\![S_i]\!]$.

3.4.3.3 Discourse Representation Theory

Various dynamic approaches to formal semantics have been proposed (Kamp, 1981; Heim, 1982; Groenendijk and Stokhof, 1984, 1991). Although Discourse Representation Theory (DRT, Kamp, 1981; Kamp and Reyle, 1993; Kamp et al., 2005) in its original conception is not a dynamic semantic theory in its strict sense, it is the most detailed theory with the broadest reception in the literature. In this theory, sentences and texts are represented as *Discourse Representation Structures* (DRSs).

Syntax of DRSs

A DRS is a pair $K = \langle U_K, C_K \rangle$, where

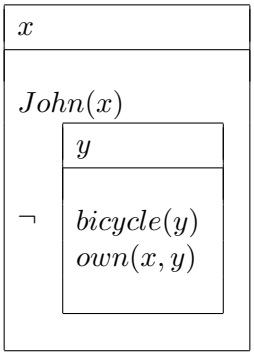
1. the universe U_K is a set of discourse referents;
2. C_K is a set of DRS-conditions: if K, K_1, K_2 are DRSs, $x_1, ..., x_n$ discourse referents, and P an n-place predicate, then $P(x_1, ..., x_n)$, $\neg K$, $K_1 \Rightarrow K_2$ are DRS-conditions;
3. if K_1 and K_2 are DRSs, then $K_1 \oplus K_2 = \langle U_{K_1} \cup U_{K_2}, C_{K_1} \cup C_{K_2} \rangle$.

According to this definition, a DRS consists of a domain, or a universe U_K, and a set C_K of conditions on the referents in this set. The third clause defines a *merge* operation $\oplus$

on DRSs.[9] Typically, a graphical representation of a DRS is a horizontally divided box, which contains elements of U_K in its upper part and elements of C_K in its lower part. DRS-conditions are properties of discourse referents in U_K, as well as relations between them. They can be atomic formulae or contain embedded DRSs. This form of recursion is occurs if expressions like "not","if","all" are to be interpreted. For example, "not" is represented as an operator $\neg$ that takes a DRS as its argument. The embedded DRS represents the content over which the negation has its scope. For instance, sentence (3.35a) is represented as (3.35b).

(3.35) a. John does not own a bicycle.

b.



In general, a DRS is a logical representation of a sentence which emerges by the recursive application of a set of DRS construction rules to the syntactic analysis of linguistic expressions. DRS construction algorithms have been specified for many contemporary syntactic theories. For illustration, if an input sentence contains an indefinite noun phrase $[_{NP}$ DET N], e.g. "a bicycle", a new discourse referent u is introduced in U_K, and a new condition [N](u), i.e. $bicycle(n)$ is added to C_K, where [N] is an unary predicate corresponding to the meaning of the noun N.

In order to determine the meaning of an utterance, the incrementally built DRS is model-theoretically interpreted by embedding it into a model. A model M for an DRS K is a pair $\langle U_M, I_M \rangle$ with a set of entities U_M and a function I_M which assigns to each n-ary predicate P_n from C_K a set of n-tuples in U_M.

As said above, unlike in Montague semantics, where the meaning of a sentence is identified with the conditions under which the sentence is true, the meaning of a sentence in DRT is identified with its context change potential, which can be seen as a relationship between two contexts: an *input context* and an *output context*.

Model-theoretically, these contexts can be characterized as assignment functions mapping discourse referents to entities in the model. While this is an extensional view, an intensional characterization contains possible worlds in both contexts[10]. Thus, a context is a pair (w, f) consisting of a possible world and an assignment function.

The truth conditions of a DRS K are defined as the conditions under which K transforms an input context (w, f) into an output context (w', g). The meaning of a DRS K, i.e. its

9 There has been an extensive discussion in the literature on the question whether this operation should be commutative or not (cf. *inter alia* van Eijck and Kamp, 1997). For our purposes, it is sufficient to assume a simple merge operation based on set union. See footnote 11 for an alternative to this operator based on dynamic conjunction.

10 I introduce the parameter for possible worlds here, although I do not need it until later when extending these semantics in chapter 5.

context change potential, is written $(w, f)[\![K]\!]_M(w', g)$. A DRS is true if and only if it can be embedded in a model.

Truth of a DRS

$M, w, f \models K$ iff there is a world w' and an assignment g such that $(w, f)[\![K]\!]_M(w', g)$.

The semantics of a DRS is defined as follows:

Semantics of DRSs

a. $(w, f)[\![\langle U, \emptyset\rangle]\!]_M(w', g)$ iff $w = w'$ and the domain of g extends the domain of f by U, i.e. $dom(g) = dom(f) \cup U$.
b. $(w, f)[\![P(x_1, ...x_n)]\!]_M(w', g)$ iff $(w, f) = (w', g)$ and $\langle f(x_1), ..., f(x_n)\rangle \in I_M(P)(w)$.
c. $(w, f)[\![\neg K]\!]_M(w', g)$
iff $(w, f) = (w', g)$ and there is no (w'', h) such that $(w, f)[\![K]\!]_M(w'', h)$.
d. $(w, f)[\![K \Rightarrow K']\!]_M(w', g)$
iff $(w, f) = (w', g)$ and for all (w'', h) with $(w, f)[\![K]\!]_M(w'', h)$ there is a (w''', i) such that $(w'', h)[\![K']\!]_M(w''', i)$.
e. $(w, f)[\![K \oplus \langle \emptyset, \gamma\rangle]\!]_M(w', g)$ iff there is a (w'', h) such that $(w, f)[\![K]\!]_M(w'', h)$ and $(w'', h)[\![\gamma]\!]_M(w', g)$.

The definition (a) introduces a new discourse referent, (b) determines the truth conditions of predicates, (c) and (d) provide adequate interpretations of logical operators, and (e) is a rule for combining a DRS with a new DRS-condition γ by virtue of the merge operator.[11] The context is actually changed only by introducing new discourse referents (a), whereas the other conditions do not actually change the context, rather they can be seen as tests for input contexts. If an input context fulfills these conditions, the assignment is returned as output context, if not, it will be rejected.

11 The notation of the semantics of DRSs follows the representation of Asher and Lascarides (2003, p. 48) which relies on van Eijck and Kamp (1997, pp. 21f.), who in turn provided a relational semantics for DRSs in the spirit of Groenendijk and Stokhof (1991).

In fact, the semantics as defined above corresponds very closely to the semantics of Groenendijk and Stokhof (1991)'s dynamic predicate logic (DPL). In DPL, instead of defining representation structures as a pair of referents and conditions, the semantics of logical operators and quantifiers is changed to the effect that they get a dynamic interpretation themselves, and thus maintaining much of the syntax of first-order predicate logic. The major difference to 'static' predicate logic is that the scopes of dynamic existential quantification and of dynamic conjunction are open to the right, such that new occurences of variables automatically get bound by the quantifier. Thus, in DPL there is no difference in meaning between $\exists x P(x) \wedge Q(x)$ and $\exists x\,[P(x) \wedge Q(x)]$. Dynamic conjunction and existential quantification in DPL are (extensionally) defined as follows (cf. Groenendijk and Stokhof, 1991, pp. 47f.):

(i) $[\![\phi \wedge \psi]\!] = \{\langle f, g\rangle \mid \exists h : \langle f, h\rangle \in [\![\phi]\!] \quad \& \quad \langle h, g\rangle \in [\![\psi]\!]\}$
(ii) $[\![\exists x \phi]\!] = \{\langle f, g\rangle \mid \exists h : dom(h) = dom(f) \cup \{x\} \quad \& \quad \langle h, g\rangle \in [\![\phi]\!]\}$
(iii) $[\![\exists x P(x)]\!] = \{\langle f, g\rangle \mid dom(g) = dom(f) \cup \{x\} \quad \& \quad g(x) \in I(P)\}$

Definition (i) corresponds to our rule (e) modulo possible worlds. Note that $\wedge$ is the dynamic conjunction and & is the static conjunction. Definition (ii) corresponds to a combination of our rules (a) and (e), thus introducing a new referent x and, at the same time, imposing a condition ϕ on it. Definition (iii) specifies (ii) for predicates. So the meaning of text (3.33) can be represented in DPL as $[\![\exists x\, man(x) \wedge enter(x) \wedge smile(x)]\!]$, which can be spelled out as $\{\langle f, g\rangle | dom(g) = dom(f) \cup \{x\} \quad \& \quad g(x) \in I(man) \quad \& \quad g(x) \in I(enter) \quad \& \quad g(x) \in I(smile)\}$.
Since the operator & is the classical commutative conjunction, this formula corresponds to the DRS $\langle x, \{man(x), enter(x), smile(x)\}\rangle$.

The construction algorithm of DRSs was originally modelled as a top-down process (Kamp, 1981; Kamp and Reyle, 1993), where the syntactic structure of a sentence is translated starting from the top node which stands for the whole sentence. Since this conception turned out to imply certain deviations from the compositionality principle (cf. Zeevat, 1989), newer versions of DRT model DRS construction as a bottom-up process (van der Sandt, 1992; Asher, 1993; Muskens, 1996; van Eijck and Kamp, 1997; Kamp et al., 2005), where preliminary semantic representations are assigned to leaves of the semantic tree, which are then composed to complex structures. In the new architecture, presuppositions are treated as a requirement which a sentence imposes on the context in which it is uttered. If the context does not satisfy the presuppositions imposed by a sentence, the presupposition is added to the context, or *accommodated*.

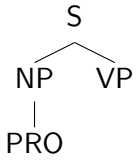
A two-stage procedure is assumed in order to cope with presuppositions. In a first stage, a preliminary DRS is constructed for each sentence. In this representation, all presuppositions which are carried by the sentence are explicitly represented. In a second stage, the already constructed DRS is checked if it satisfies these presuppositions, and they are resolved if possible, or accommodated if necessary. Then, the asserted part of the sentence is merged with the updated context, resulting in a DRS representing both the context and the asserted content of the sentence.

3.4.3.4 Anaphora in DRT

Let us briefly sketch how anaphora are interpreted in the original top-down version of DRT. DRT does not make a distinction between different kinds of anaphora. In this way, all pronominal anaphora are interpreted as variables which have to be bound by an antecedent that is accessible in the discourse structure.

For instance, pronouns are interpreted according to a particular construction rule CR.PRO (Kamp and Reyle, 1993, p. 122). This rule is triggered by a (slightly simplified) syntactic configuration as shown in (3.36).

(3.36)



The rule consists of the following steps: (i) introduce a new discourse referent u into the universe U_K, (ii) introduce a condition in C_K, replace the branch $[_{NP}$ PRO] in the syntactic structure by the referent u, and delete the syntactic structure, and (iii) add a new condition of the form $\alpha = \beta$ where α is the new discourse referent and β is a suitable discourse referent from the DRS' universe.

For example, the first sentence in the short discourse (3.33), repeated below as (3.37), translates into the DRS (3.38). The indefinite noun phrase "a man" introduces a discourse referent x and a DRS-condition $man(x)$, and the verb phrase "entered" is represented by a condition $enter(x)$[12].

(3.37) A man entered. He smiled.

12 For the sake of clearness of the exposition, information about tenses are ignored.

(3.38)

x
$man(x)$ $enter(x)$

The second sentence then is added to this DRS by introducing a referent u for the pronoun "he", together with a DRS-condition $u = x$. In this case, x is the only available discourse referent and thus chosen as the antecedent of the anaphoric condition.

(3.39)

x, u
$man(x)$ $enter(x)$ $smile(u)$ $u = x$

However, things are not always that simple, and in discourses with more than one discourse referent, the right one has to be selected. In order to determine which referent is a suitable antecedent for an anaphor, we first need the notion of *subordination*.

Subordination A DRS K_1 is *immediately subordinated* to another DRS K_2, $K_1 \leq K_2$, if the DRS-conditions of K_2 contain either K_1 or $\neg K_1$.

DRS-Accessibility A discourse referent v is DRS-*accessible* for an anaphoric DRS-condition in K_1 iff v is introduced in K_2 and the following holds:

1. $K_1 \leq K_2$, or
2. $K_2 Q K_3$ and $K_1 \leq K_3$, where Q stands for $\Rightarrow$ or another logical operator.

Thus, in order to determine what is accessible, the following procedure is applied recursively. Starting from the DRS containing the anaphor, accessible discourse referents are either immediately to the "left" of the current DRS, or in the immediately superordinated DRS. If no such DRS can be found, then the anaphor cannot be linked to an antecedent. This is the case in example (3.34), repeated below as (3.40), where "a donkey" is represented in an embedded DRS which is not accessible for the pronoun in the second sentence.

(3.40) Every farmer who owns a donkey$_i$ beats it. *It$_i$ lives in a stable.

In sum, anaphora resolution in DRT is a function from DRSs to DRSs which is constrained by the accessibility condition.

An important modification of DRT was presented by van der Sandt (1992) who first proposed a two-stage bottom-up architecture of DRT (as sketched in the last section), which allows for a uniform treatment of anaphora and (certain kinds of) presupposition. Anaphora, i.e. pronouns and definite noun phrases, are regarded as carrying the presupposition that the context provides a suitable antecedent. This presupposition can be resolved by binding an anaphor to an antecedent via coreference. If such a binding is not possible, the antecedent must be accommodated. Indefinite noun phrases are assumed to be without presuppositions, allowing them to act in a way similar to existential quantifiers.

3.4.3.5 Discussion

Dynamic semantic theories model the meaning of a sentence as a context change operation. This conception stands in contrast to static theories of formal semantics and is straightforward in accounting for some type of anaphoric expressions. However, DRT, as it is originally conceived, excludes most part of pragmatic information from the analysis. But in many cases, reference resolution relies on knowledge not modelled by a purely semantic theory. We will see in chapter 7 how some parts of extralinguistic knowledge can enter a DRS.

There are two other issues which are not treated by DRT. We have seen in section 3.3.2 that not all discourse referents that are present in a discourse model are equally suited for subsequent anaphoric references. Recently introduced referents are more likely to be in the center of attention of discourse participants and are thus more probable candidates for antecedents of anaphora than referents introduced earlier. The universe of a DRS is modelled as a set of discourse referents. A natural extension of DRT would consist in ordering this set according to their salience or activation or familiarity. Again, in chapter 7, I will make a proposal how DRT can be minimally extended in this respect.

Furthermore, and most crucially, natural language discourses are more than just a concatenation of single utterances. They have an internal structure which has an influence on many other anaphoric phenomena. With the DRS merge operation as defined in DRT, discourse structure gets lost in the interpretation process and is not reflected in the final representation any more. In the next chapter, I will discuss the structure of discourses. As we will see then in chapter 5, there are approaches to discourse interpretation which take the structure of discourses into account.

3.4.4 Conclusion

In this chapter, I have introduced the basic notion of a discourse model as a representational level where discourse referents, i.e. entities introduced in a discourse, are stored. Closely connected with these terms are discourse anaphora, which constitute relations between discourse referents. A series of competing notions was proposed in order to capture the fact that not all discourse referents in a discourse model are equally suited for anaphoric references.

On the one hand, starting from this observation, Centering Theory delivers an explicit mechanism for the resolution of anaphoric expressions, especially pronouns. However, this theory does not take the semantics of sentences into account, but rather restricts the linguistic knowledge it uses to surface word order, grammatical roles, or information structure. It does not make explicit how discourse referents enter the discourse model.

On the other hand, dynamic semantic theories were developed in order to capture the incremental process of how a discourse model emerges as a discourse proceeds. DRT explicitly states how representations of discourse entities are obtained from the semantic content of sentences. However, these theories do not make statements about the salience or activation of discourse referents.

An obvious conclusion from these observations suggests a combination of both types of theories. The performance of DRT could probably be raised by taking the relative salience of discourse referents into account. However, none of the presented theories in this chapter accounts for the complex hierarchical structuring of discourses. Centering Theory treats

successive utterances simply as a sequence, and in DRT, the addition of a new sentence to an existing discourse representation consists in a simple concatenation, so that, in spite of the dynamicity of this process, the incremental discourse structure is occulted in the final representation. It thus seems natural that the next chapter is devoted to the complex structuring of discourses.

Chapter 4

Discourse Structure

When we perceive or produce spans of text consisting of various sentences, or when we take part in a conversation, we arrange sentences or utterances in a structured way. When examined closely, it becomes obvious that texts and discourses have some kind of structure. So it is not a surprise that most current theories of discourse interpretation agree to the assumption that discourses are structured. However, theories differ substantially as to what theoretical grounds structures are built upon, which form this structure has, and on what representational and/or interpretational level these structures are conceived.

Most theories of discourse structure assume some kind of hierarchical discourse structure that is based on a notion of discourse coherence (cf. section 1.3.3) involving coherence relations between discourse segments. Different accounts of discourse coherence have proposed different taxonomies of coherence relations on different representational levels. Within these theories, a number of subgroups can be distinguished. Some of them relate discourse relations with the information contained in discourse segments connected by these relations (e.g. Hobbs 1985a; Kehler 2002; Asher and Lascarides 2003). Hobbs (1996) calls this group of theories *informational* accounts. Opposed to them are *intentional* accounts, which take intentions of conversation participants, or authors and addressees of texts, into account (e.g. Grosz and Sidner 1986; Poesio and Traum 1997). A third group is less interested in the meaning of texts, but rather in the form discourse structure can take. Let us call these theories *discourse syntactical* accounts (e.g. Polanyi 1988; Mann and Thompson 1988; Forbes et al. 2001). A further group of theories, which are centered on the notion of *discourse topic* (e.g. Klein and von Stutterheim 1987; van Kuppevelt 1995), derives discourse structures from questions which are assumed to be underlying a text.

The concern of this chapter is to examine the structure of discourses. The theories mentioned above shed light on different aspects of discourse structure. First, in section 4.1, I will discuss some essential characteristics of discourse structure. Then, in section 4.2, I will have a closer look at discourse relations. Finally, in section 4.3, I will close the chapter with a discussion of different conceptions of the notion of discourse topic.

4.1 Characteristics of Discourse Structure

Any theory of discourse structuring has to account for the following characteristics: (i) the basic structural units and its representation, (ii) relations between units, and (iii) the form of the discourse structure. As these notions are used in quite different ways in the existing theories, I will discuss them in the following sections.

4.1.1 Discourse Segments: Basic Structural Units

When perceiving a text we have the intuition that certain spans of the text naturally group together. Generally, discourse segmentation is taken to be a chunking of a text into sequences of related clauses or sentences (cf. Grosz and Sidner, 1986; Webber, 1988). In principle, basic structural discourse units can be linguistic expressions and combinations thereof, such as phrases, sentences, or even paragraphs. I will refer to these basic units of discourse structure as *discourse segments*.

In current theories of discourse structure there is little consensus on how a particular text should be segmented. It is not clear whether discourse segments are mere abstract linguistic or psychologically real entities. Mostly, they are gathered by intuitive decisions of annotators or test subjects, and thus are vague and not unequivocal (Walker, 2000). As a consequence, it seems that there is no uniform definition for a discourse segment, apart from being a nonoverlapping, contiguous span of text. There are different assumptions on what grounds a discourse is to be segmented. Grosz and Sidner (1986) assume intentional units and take a discourse segment to be a chunk of text expressing a common purpose – the discourse segment purpose – with respect to the speaker's intentions. Hirschberg and Nakatani (1996) argue that boundaries between segments can be determined by prosodic features, e.g. pauses and fundamental frequency changes. Hobbs (1985a) takes a discourse segment to be a chunk of text with a common meaning. For him, as for others (Polanyi, 1988; Webber et al., 2003; Asher and Lascarides, 2003), the minimal discourse segment is a clause, i.e. a group of words containing a subject and a verb. It can either be a simple sentence or a part of a complex sentence. Finally, some accounts (Poesio and Traum, 1997) allow even smaller units down to any fraction of an utterance. In this thesis, I will assume that discourse segments correspond to utterances of clauses. In this way, I will use the terms "discourse segment" and "utterance" synonymously.

4.1.2 Connecting Discourse Segments

4.1.2.1 Discourse Markers

A special role in discourse structuring play *discourse markers*. These are expressions that indicate a particular structuring of discourses. Markers can be adverbs, connectives and particles, especially discourse particles. The latter ones are divided in two groups: some of them have a quite transparent meaning (e.g. *in sum*, *for example*, *because*, while others are rather non-transparent (e.g. *anyway*, *though*, *and*). As Zeevat and Karagjosova (2007) illustrate, non-transparent meanings can be a sign of older grammaticalizations.

Various classifications of discourse markers can be found in the literature, e.g. Alonso i Alemany (2005) characterizes a total of 84 markers by their structural properties (subordinating vs. coordinating markers) and their semantics (markers expressing revision, cause, equality, context). The actual inventory of discourse markers, as well as their morphosyntactical properties, may vary from one language to another. For instance, the English marker for a coordinating marker expressing revision is *nevertheless* and its Spanish equivalent *sin embargo*[1].

Some discourse markers are *underspecified* or *ambiguous* with respect to their semantic and syntactic properties. This is due to the fact that often a relation holding between

1 See below (section 4.2.3) for more examples of discourse markers and a discussion concerning their relationship to discourse relations.

discourse segments is just reinforced by the presence of a discourse marker. Examples for underspecified markers are *anyway, moreover* or *even*.

In general, a discourse marker can be taken as indicating that a particular relation holds between two or more discourse segments. For instance, in the following short discourse from Lascarides and Asher (1991), there is a a causal relationship between the two events involved. The marker *because* indicates that the two discourse segments are related by a causal discourse relation.

(4.1) a. Max fell

b. because John pushed him.

However, in many (if not most) cases, a certain relation holds between discourse segments, although it is not explicitly indicated by a discourse marker. Thus, discourse markers are *optional* in many cases, and in order to properly account for the structuring of discourses, a notion is needed that goes beyond explicitly linguistically coded discourse markers.

4.1.2.2 Discourse Relations

In a coherent discourse, utterances or discourse segments stand in certain relationships one to each other, being indicated by a discourse marker or not. Text recipients always try to find relations between parts of a text in order to perceive it as coherent. Consider example (4.2) without a discourse marker:

(4.2) a. Max fell.

b. John pushed him.

In this short discourse, in order to make sense for a listener, the two utterances must be interpreted as being connected in some way. When we take world knowledge into account, we can infer that the second utterance can be an explanation for the first one, because an event of pushing someone can have as a possible consequence that this person falls. Thus, we can assume that coherence is established by assuming an EXPLANATION relation between the two discourse segments (4.2a) and (4.2b).

This kind of relation is known as *coherence relation* (Hobbs, 1979, 1985a), or *rhetorical relation* (Mann and Thompson, 1988), or *discourse relation* (Lascarides and Asher, 1991). Although originally conceived in distinct ways, I will use these terms synonymously.

Crucially, the establishment of discourse relations is context-dependent. Without further contextual specification, utterance (4.2b) is treated as an EXPLANATION for utterance (4.2a), as explained above. However, in a further specified context (4.3), this inference, which may be drawn after the second utterance is made, has to be cancelled after the last utterance is made, and another discourse relation, NARRATION, has to be assumed in order to make the text coherent. Thus, the property of defeasibility underlines the pragmatic nature of inferring coherence relations.

(4.3) a. Max fell.

b. John pushed him.

c. Max rolled over the edge of the cliff.

In section 4.2, I will examine the most influential proposals and address the questions what kinds of rhetorical relations can be distinguished, how they can be marked by a speaker, and how they are recognized by addressees[2].

4.1.3 The Form of Discourse Structure

As we have seen in the last section, most theories of discourse structure agree on the assumption that a structured discourse emerges by virtue of the existence of coherence relations between discourse segments. However, there are competing views on how a complex structure for a bigger discourse is obtained. In the following, I will discuss different proposals made in the literature.

4.1.3.1 Sequences

In the first place, due to the property of texts to unfold in a linear way, one can observe that consecutive discourse segments constitute a **sequence**. Every segment has a set of information which does not depend on the preceding segment. If we take larger chunks of text as discourse segments, e.g. book chapters or sections, we can easily obtain a sequence. An example of a sequential structure is depicted in Fig. 4.1.

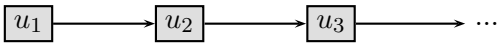


Figure 4.1: A sequence structure

But when we take a closer look at texts we note that there is more to discourse structure than being a mere sequence. For instance, narrative texts are not always a simple mapping of the time course of described events to a sequence of utterances. Look again at example (4.2), here repeated as (4.4). As explained above, the preferred reading for this text takes (b) as an EXPLANATION for (a), presupposing that the pushing event described in (b) took place before the falling event described by (a).

(4.4) a. Max fell.

b. John pushed him.

4.1.3.2 Stacks

To account for this text, we can fall back on a notion that is widely applied in computer science. A **stack** is a data structure based on the principle of "Last In First Out". The underlying metaphor is a stack of plates in a restaurant. Two operations are possible: *push* and *pop*. Push adds an element to the top of the stack, leaving previous elements below and making them invisible. Pop removes the current top element from the stack. In principle, a stack element can consist of any data type, also of stacks themselves. Applied to discourse structure, complex segments containing smaller, more specific segments, can

2 Note that not all theories of discourse structure make use of a concept of rhetorical relations. While for some accounts, rhetorical relations are a basic concept, for others it is just derived or does not have any theoretical status at all. In section 4.3, I will briefly review some alternative ways to structure discourses.

be pushed on the stack, but removed from the stack only when its sub-segments have been popped already. Such a structure is illustrated in Fig. 4.2.

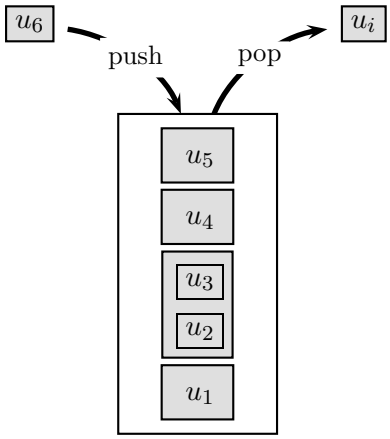


Figure 4.2: A stack structure

Stacks play an important role in the focus theory of Sidner (1981) and in the attentional structure of Grosz and Sidner (1986). Freitas (2005) observes two disadvantages of this type of structure: first, there is no history record of pushed and popped segments which would permit a revision of already made interpretations. Second, this structure is too restricted because the only possible relation between segments is subordination, not allowing for any other relation between segments to be expressed.

4.1.3.3 Trees

A way that accounts better for the hierarchical character of discourses is to assume a **tree** structure for representing discourses. In fact, most accounts of discourse structure assume a tree-like structure for discourses (e.g. Grosz and Sidner, 1986; Mann and Thompson, 1988; Polanyi and Scha, 1984; Polanyi, 1988; Marcu, 2000).

In general, trees keep record of the history of generated discourse structure and permit a structural distinction of coordination and subordination. The main advantage of a tree structure is its inherent assumption of *compositionality* on the discourse level, as stated in the following quotation:

> A clause is a segment of discourse, and when two segments of discourse are discovered to be linked by some coherence relation, the two together thereby constitute a single segment of discourse. By recognizing coherence relations between segments, we can thus build up recursively a structure for the discourse as a whole.
>
> (Hobbs, 1985a, p. 23)

Note that in this definition, the combination of two segments yields again a segment, though on a more abstract level. A simple tree structure is shown in Fig 4.3, where terminal nodes u_i represent discourse segments, and non-terminal nodes R_j represent coherence relations.

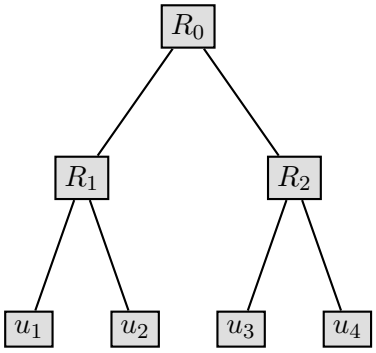


Figure 4.3: A tree structure

In what follows I will point out some more important properties of tree structures and briefly introduce two syntactically oriented theories of discourse structure where trees play a central role, though being conceived in a rather distinct manner.

The Linguistic Discourse Model

A theory that emphasizes the role of trees in discourse structure is the Linguistic Discourse Model (LDM), an account developed by Polanyi and Scha (1984); Polanyi (1988); Polanyi et al. (2003). In this proposal, a discourse tree is constructed by recursively applying a set of discourse construction rules to a sequence of basic discourse units. The choice of a particular rule is determined by information contained in the surface structure of the unit. Units are attached to the existing discourse tree on its right edge.

> A new constituent is attached to the Discourse Parse Tree as the rightmost constituent at a structurally accessible existing level in a Tree.
>
> (Polanyi, 1988, p. 613)

This constraint, also known as *Right Frontier Constraint*[3], is a very important constraint in discourse interpretation and anaphora resolution. The Right Frontier of a tree is illustrated in Fig. 4.4. I will come back to this constraint in section 4.2.4 below, and again in more detail in chapter 5 in section 5.3.2.4.

The main discourse rules in the LDM, besides some additional rules for language specific constructions, are coordination and subordination. In coordinations, all daughter nodes contribute equally to the structure of the constructed node. In subordinations, only the subordinating node contributes to the discourse structure, while the structure of the subordinated node does not play a role in the overall discourse structure.

While in LDM the question of how contextual information of various kinds interacts with discourse trees is not addressed explicitly, Polanyi et al. (2003) concentrate more on the discourse integration of sentences according to their internal information structure[4].

3 The Right Frontier Constraint was formulated in a similar way by Webber (1988, p. 114).

4 Information structure is the partition of a sentence according to the informational status of its constituents. Information structural properties include distinctions between, e.g., new and given information, or focus and background. The exact use of these terms depends on the adopted theory (cf. *inter alia* Vallduví, 1992; Hajičová et al., 1998; Steedman, 2000, see also section 4.3 in this chapter).

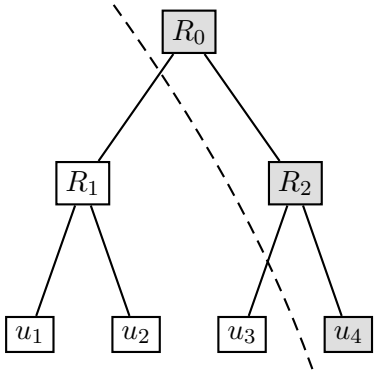


Figure 4.4: The Right Frontier of a tree

D-LTAG

Tree structures for discourses can also be conceived in a rather different way. D-LTAG (Discourse - Lexicalized Tree Adjoining Grammar, Forbes et al., 2001; Webber, 2004) extends a sentence level grammar (Tree Adjoining Grammar, TAG) to the discourse level. TAG (Joshi, 1985) is a syntactic formalism that has been successfully implemented in parsing systems for many languages. Later it has been extended in different ways in order to cover semantic interpretation of sentences (Joshi et al., 2007), one of them known as lexicalized TAG.

An obvious justification for assuming the same mechanisms at the sentence and discourse levels is the fact that discourse relations can hold between two sentences and within a single sentence, as in example (4.5).

(4.5) a. John held out a bone to the dog. She caught it quickly.

b. John held out a bone to the dog who caught it quickly.

(Danlos, 2007)

This approach stresses the compositional character of discourse structure, besides that additional discourse phenomena such as anaphora resolution and other inferential processes are accounted for, as well. A lexicalized TAG is made of two kinds of elementary trees: initial trees encoding predicate-argument structures and auxiliary trees which recursively modify and elaborate elementary trees.

Complex structures are built by two operations: substitution and adjunction. An initial tree is formed by clauses connected by a subordinating conjunction (e.g. *because*). Its compositional semantics is determined by the semantics of the conjunction and the clauses, where the conjunction acts as a predicate and the clauses as its arguments. Auxiliary trees provide additional, both anaphoric and inferential, information. In this way, this approach tells apart the compositional part of discourse structure from additional non-compositional elements. An important feature of discourse structure – connections between discourse segments via discourse relations – is missed by this approach. An interesting variation of an extension of TAG to discourses, D-STAG (Discourse Synchronous Lexicalized Tree Adjoining Grammar), was presented recently by Danlos (2007). This approach, as opposed to D-LTAG (Forbes et al., 2001), takes both marked and unmarked discourse relations into

account. I will not go into further details here; for a recent overview and introduction to TAG, see Joshi et al. (2007).

Nodes in Tree Structures

Another theory based on pure tree structures is RST (Rhetorical Structure Theory, Mann and Thompson, 1988). Here, in contrast to the model discussed in the previous section, a tree is obtained by recursively connecting two ore more adjacent text spans by a rhetorical relation. Due to the nuclearity principle of RST[5], the edges in the tree are directed. While this account sticks closely to the surface structure of a text and permits only nodes corresponding to spans of text, others, e.g. LDM (Polanyi, 1988) or SDRT (Asher and Lascarides, 2003), allow (non-terminal) nodes to be representations of abstract discourse topics which do not correspond to specific text spans.

Some theories allow for deviations from a pure tree structure. Although discourse representations in earlier versions of SDRT (Lascarides and Asher, 1991) appeared to be tree structures, the standard versions of this theory (Asher and Lascarides, 2003) allow for nodes with multiple parents. Hence, more than one rhetorical relation may hold between two discourse segments. More recent work in SDRT does in fact assume a representation that, although for many texts still close to trees, is actually an instance of a more general structure, namely *directed acyclic graphs*[6].

4.1.3.4 Graphs

A further observation made by Hobbs (1985a) is that every discourse segment can be related to any preceding segment in any order. If we allow this, we go beyond the descriptive power of tree structures and need to assume that discourses have a less restricted and more general **graph** structure.

Hobbs (1985a) proposes a graph structure where the basic units, or vertices, represent the propositional content of sentences, and labelled directed edges represent coherence relations between the basic units. In order to build such a structure, each sentence is compared to all sentences interpreted so far, generating a graph with diverse possibilities of connecting the basic unit representing the currently interpreted sentence to the previously interpreted discourse. As a result, the graph establishes a series of links between the current sentence and the previous discourse, permitting the addressee to establish a coherent discourse representation.

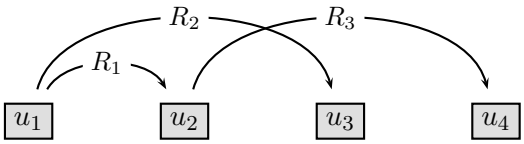


Figure 4.5: A graph structure

A graph representation, illustrated in Fig. 4.5, seems to be more close to reality than trees, but – due to the enormous computational complexity involved in reasoning with

5 RST relations are introduced in section 4.2.3.
6 SDRT graphs are introduced in more detail in section 4.2.4.

unrestricted graphs – most theories assume a more restricted representation. For a recent detailed discussion on data structures for representing discourses, see Wolf and Gibson (2006). These authors argue that trees are not an adequate representation for discourses, mainly because of existing cross-dependencies in naturally occurring texts, and opt for a graph structure. An instance of these rather complex structures is depicted in Fig. 4.6, where discourse segments are represented by nodes u_i, and coherence relations by edges R_j or additional nodes T_k.

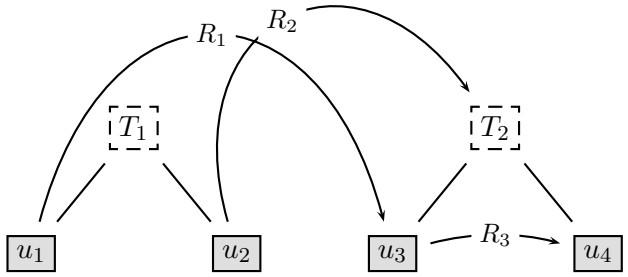


Figure 4.6: Graph structure proposed by Wolf and Gibson (2006)

Wolf and Gibson draw arguments for their model mainly from work on a corpus of manually annotated texts. They find (Wolf and Gibson, 2005, 2006) that there are phenomena that cannot be captured by a purely tree-based approach. In particular, these phenomena are nodes involved in more than one discourse relation and crossed dependencies between nodes. Moreover, the fact that discourse relations affect pronoun resolution (Wolf et al., 2004) is taken to be reflecting the psychological reality of their relations. Finally, the authors show that their graph-based model outperforms tree-based models in automatic text summarization.

To save the treeness of discourse structure – and with that the supplied computational advantages – Egg and Redeker (2008) argue that many apparent crossed dependencies reflect anaphoric relations on the level of discourse cohesion rather than rhetorical relations on the level of discourse coherence. They opt for a pure tree structure for discourse relations, keeping anaphoric relations apart. For instance, Wolf and Gibson cite example (4.6) as evidence for crossed dependencies.

(4.6) u_1 : Susan wanted to buy some tomatoes

u_2 : and she also tried to find some basil

u_3 : because her recipe asked for these ingredients.

u_4 : The basil would probably be quite expensive at this time of the year.

In Wolf and Gibson's model, this discourse would have the structure depicted in Fig. 4.7, whereas Egg and Redeker propose the structure shown in Fig. 4.8 for the same discourse, arguing that the relation between u_4 and u_2 emerges from the anaphor *the basil* in u_4 which refers back to *some basil* in u_2.

While I agree with Egg and Redeker on the necessity of keeping coherence and cohesion apart, there is a danger in arbitrarily postulating additional abstract discourse entities for

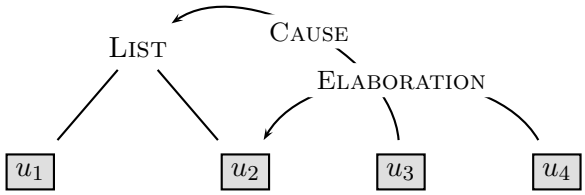


Figure 4.7: Graph structure for (4.6) proposed by Wolf and Gibson (2006)

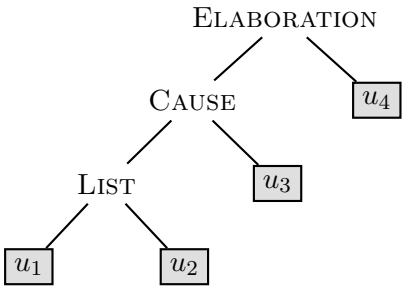


Figure 4.8: Tree structure for (4.6) proposed by Egg and Redeker (2008)

the sake of avoiding otherwise necessary deviations from tree structures. As Danlos (2006) points out, the structure proposed by Egg and Redeker (Fig. 4.8) would remain equal if segment u_4 concerned the price of tomatoes rather than that of basil, thus omitting important details of discourse structure. In any case, as Danlos and Asher and Lascarides propose, we can restrict graphs for discourse structures to be directed and acyclic.

4.1.4 Conclusion

The last example illustrates the current intensive discussion regarding the question of the right data structures for discourses. It is in no wise easy to decide which structure for a given discourse is the most adequate. Different theories provide different structures which serve partly distinct interests. However, an important point is that in order to avoid ad-hoc theoretical constructions, theories of discourse structure should try to avoid additional abstract discourse entities whenever possible. In this thesis, I will comply with Asher and Lascarides (2003); Danlos (2006); Wolf and Gibson (2006) and assume directed acyclic graphs as the descriptively and explanatorily most adequate data type for the form of discourse structures.

Concerning the other characteristics of discourse structure discussed in this section, discourse segments corresponding to utterances of clauses are taken as the basic structural units of discourse organization. Relations between discourse segments are expressed by discourse relations, which I will turn to in more detail in the next section.

4.2 Discourse Relations

A multitude of taxonomies of discourse relation can be found in the literature, with the number of assumed relations ranging from 2 to over 100 (see *inter alia* Hobbs, 1985a;

Mann and Thompson, 1988; Knott, 1996; Marcu, 2000; Kehler, 2002). In this section, I will present the most influential accounts of discourse relations.

In the theory of Grosz and Sidner (1986), coherence relations reflect how the roles regarding the intentions of discourse participants played by successive discourse segments relate to each other. Concerning the composition of complex discourse structures, Grosz and Sidner (1986) assume that discourse relations hold between discourse segment purposes and do not relate the semantic content of segments directly, in contradistinction to informational theories. Two different kinds of discourse relations are assumed: Dominance and Satisfaction-precedence.

Dominance A discourse segment purpose DSP_2 **dominates** another purpose DSP_1 if the satisfaction of DSP_1 may be intended to provide part of the satisfaction of DSP_2.

Satisfaction-precedence A discourse segment purpose DSP_1 **satisfaction-precedes** DSP_2 if DSP_1 must be satisfied before DSP_2 (Grosz and Sidner, 1986, p. 179).

This distinction can be illustrated by two examples taken from Knott (1996, p. 38).

(4.7) Television is bad for children. They grow up on a steady diet of violence and advertising.

(4.8) Try out the gun by firing off a few rounds. First, release the safety catch; then squeeze the trigger gently.

In text (4.7) the first sentence dominates the second. In (4.8), in addition to a Satisfaction-precedence between the purposes of the two last segments (the two clauses in the second sentence), they are both dominated by the purpose of the first segment.

4.2.1 Hobbs' Coherence Relations

In Hobbs (1985a)'s proposal, there are four different classes of coherence relations. Firstly, a discourse can be coherent because "it tells about coherent events in the world". When an event is mentioned in a text, another one can be inferred using a certain amount of background knowledge. Hobbs calls this type of relation the Occasion relation. His example is given in (4.9).

(4.9) a. At 5:00 a train arrived in Chicago.

b. At 6:00 Ronald Reagan held a press conference.

The second class of relations results from the need to relate what has been said to some goal of the speaker. Hobbs labels these relations as Evaluation relations.

(4.10) a. Did your bring your car today?

b. My car is at the garage.

The third group of relations relate a discourse segment to the listener's prior knowledge. Hobbs mentions the relations Background and Explanation. An example for an explanation relation was given in (4.4) above, a background relation is given in (4.11).

(4.11) a. While I was sitting in the Central Station,

b. a young man came up to me.

Finally, the last group is defined in terms of moves between specific and general assertions and their interaction with negation. The most prominent relations in this group are Parallel, Contrast and Elaboration, Exemplification, and Violated Expectation.

The set of coherence relations proposed by Hobbs (1985a) has proven to provide a quite reliable basis. On closer examination, other taxonomies mostly do not propose substantially different relations, but rather arrange them differently in groups.

4.2.2 Kehler's Three Types of Coherence

Kehler (2002) has tried to base the conception of rhetorical relations on philosophical grounds. Hume (1748)'s three categories of ideas are taken to be the basis for Kehler's three-partite categorization of Hobbs' coherence relations. The first group comprises the type of relations expressing some kind of explanations between ideas, *Cause-effect* relations. Another group of relations is based on *Resemblance* of ideas, and the last category, *Contiguity*, deals with the temporal and spatial neighbourhood of ideas. In the following, I will look at each of these groups and give an overview of their most prominent relation types.

4.2.2.1 Coherence Relations: Cause-Effect

The interpretation of certain relations requires the listener to identify a path of implications between the propositions expressed by single utterances. The assumption of a group of causal relations seems convincing, even more if we look at Kehler's subclassification into four types. I will give Kehler's examples and definitions for each of them.

Result Infer P from the assertion of S_1, and Q from the assertion of S_2, where normally holds that $P \rightarrow Q$.

(4.12) George is a politician, and therefore he's dishonest.

In this example, P is expressed by the first clause S_1 ("George is a politician"), and Q by the second clause S_2. We need world knowledge, namely that politicians are normally dishonest, to draw the required inference.

Explanation Infer P from the assertion of S_1, and Q from the assertion of S_2, where normally holds that $Q \rightarrow P$.

(4.13) George is dishonest, (because) he's a politician.

Striking in this definition is that the Explanation relation is just an inversion of the Result relation involving an inverse ordering of segments. Furthermore, when negation is taken into account, we have two more cases in the same logical pattern.

Violated Expectation Infer P from the assertion of S_1, and Q from the assertion of S_2, where normally holds that $P \rightarrow \neg Q$.

(4.14) George is a politician, but he's honest.

Denial of Preventer Infer P from the assertion of S_1, and Q from the assertion of S_2, where normally holds that $Q \rightarrow \neg P$.

(4.15) George is honest, even though he's a politician.

In sum, causal relations rely on the semantic content of clauses, in contrast to the next category.

4.2.2.2 Coherence Relations: Resemblance

These relations are based on Hume's category of resemblance of ideas. Consider example (4.16), where a contrast is expressed by means of parallel syntactic patterns.

(4.16) a. John McCain is extraordinarily energetic for a 71-year-old;

b. Barack Obama is remarkably grounded for a 46-year-old.

(Time Magazine, April 10, 2008)[7]

The establishment of a Resemblance relation involves two steps: First, corresponding parallel entities and eventualities have to be identified, and second, these entities and eventualities are put in a relation of Resemblance. The prototypical Resemblance relation is Parallel:

Parallel Infer $p(a_1, a_2, ...)$ from the assertion of S_1, and $p(b_1, b_2, ...)$ from the assertion of S_2, where for some property vector $\overrightarrow{q}$, $q_i(a_i)$ and $q_i(b_i)$ for all i.

Coherence is reached by identifying a common relation p together with common properties $\overrightarrow{q}$ of parallel entities. In example (4.16), p can be described by $\lambda x \lambda P \lambda y[x$ is P for $y]$, while a_1 is (the discourse referent corresponding to) McCain, a_2 is $\lambda z[extraordinarily - energetic(z)]$, a_3 is a 71-year-old, and b_1 is Obama, etc. The properties talked about are a presidential candidate as q_1, a personal property as q_2, and a description of a certain type of person as q_3.

So far, commonalities between entities and relations in utterances are accounted for. For Contrast relations, negation is taken into account, yielding two definitions. In the first one, relations between parallel entities are contrasted, while in the second case, a property of an entity stands in contrast to a property of a parallel entity. Kehler's corresponding examples are given in (4.17a) and (b) respectively.

Contrast 1 Infer $p(a_1, a_2, ...)$ from the assertion of S_1, and $\neg p(b_1, b_2, ...)$ from the assertion of S_2, where for some property vector $\overrightarrow{q}$, $q_i(a_i)$ and $q_i(b_i)$ for all i.

Contrast 2 Infer $p(a_1, a_2, ...)$ from the assertion of S_1, and $p(b_1, b_2, ...)$ from the assertion of S_2, where for some property vector $\overrightarrow{q}$, $q_i(a_i)$ and $\neg q_i(b_i)$ for all i.

(4.17) a. Gephardt supported Gore, but Armey opposed him.

7 http://www.time.com/time/magazine/article/0,9171,1729712,00.html

b. Gephardt supported Gore, but Armey supported Bush.

Other Resemblance relations involve a membership or subset relation between elements in two or more clauses. Depending on the clause order, we have EXEMPLIFICATION and GENERALIZATION. Again, with a negated predicate we get two EXCEPTION relations.

Exemplification Infer $p(a_1, a_2, ...)$ from the assertion of S_1, and $p(b_1, b_2, ...)$ from the assertion of S_2, where $b_i \in a_i$ or $b_i \subset a_i$ for some i.

(4.18) Young aspiring politicians often support their party's presidential candidate. For instance, Bayh campaigned hard for Gore in 2000.

Generalization Infer $p(a_1, a_2, ...)$ from the assertion of S_1, and $p(b_1, b_2, ...)$ from the assertion of S_2, where $a_i \in b_i$ or $a_i \subset b_i$ for some i.

(4.19) Bayh campaigned hard for Gore in 2000. Young aspiring politicians often support their party's presidential candidate.

Exception 1 Infer $p(a_1, a_2, ...)$ from the assertion of S_1, and $\neg p(b_1, b_2, ...)$ from the assertion of S_2, where $b_i \in a_i$ or $b_i \subset a_i$ for some i.

(4.20) Young aspiring politicians often support their party's candidate. However, Rudy Guiliani supported Mario Cuomo in 1994.

Exception 2 Infer $p(a_1, a_2, ...)$ from the assertion of S_1, and $\neg p(b_1, b_2, ...)$ from the assertion of S_2, where $a_i \in b_i$ or $a_i \subset b_i$ for some i.

(4.21) Rudy Guiliani supported Mario Cuomo in 1994. Nonetheless, Young aspiring politicians often support their party's candidate.

The last coherence relation in this class, ELABORATION, is seen as a special case of PARALLEL, in which the parallel entities are in fact identical. This leads to a very simple definition:

Elaboration Infer $p(a_1, a_2, ...)$ from the assertions of S_1 and S_2.

(4.22) A young aspiring politician was arrested in Texas today. John Smith, 34, was nabbed in a Houston law firm while attempting to embezzle funds for his campaign.

However, this definition is not very useful, since nothing is said about how to infer p. We will see more comprehensive definitions of this relations later on.

4.2.2.3 Coherence Relations: Contiguity

This class comprises, like Hobbs' first category, only one coherence relation: OCCASION. Again, Kehler gives two definitions for each clause ordering. Hobbs' example was given in (4.9) above, Kehler's examples are below.

Occasion 1 Infer a change of state for a system of entities from S_1, inferring the final state for this system from S_2.

Occasion 2 Infer a change of state for a system of entities from S_2, inferring the initial state for this system from S_1.

(4.23) a. George picked up a speech. He began to read.

b. Larry went into a restaurant. The baked salmon sounded good and he ordered it.

Again, little is said about the inference processes involved. In fact, inferring Contiguity relations (in SDRT, the most important relations of this class are called NARRATION and CONTINUATION) turns out to be a very complex process when looked at more closely. Particularly, it seems that some notion of discourse topic may be unavoidable. I will discuss this issue in section 4.3.

4.2.2.4 Linguistic Phenomena Explained by Kehler's Taxonomy

Kehler (2002)'s taxonomy can account nicely for the impact coherence relations have on certain linguistic phenomena, like VP-ellipsis, gapping, and pronoun resolution.

VP-Ellipsis

Concerning VP-ellipsis, i.e. the elusion of verb phrases, competing theories can explain disjoint sets of data. On the one hand, in syntactic approaches (e.g. Lappin, 1996), ellipses are reconstructed by copying a suitable syntactical structure. For instance, (4.24a) is a valid abbreviation of (4.24b), bearing the same denotation.

(4.24) a. Bill likes playing golf, and George does, too.

b. Bill likes playing golf, and George likes playing golf, too.

In many cases, this approach works very well, but often there is no syntactically parallel VP which can be copied. In example (4.25b), the voice in source (passive) and target clause (active) is different, but still it is possible to elide the VP resulting in (4.25a).

(4.25) a. In March, four fireworks manufacturers asked that the decision be reversed, and on Monday the ICC did.

b. In March, four fireworks manufacturers asked that the decision be reversed, and on Monday the ICC reversed the decision.

On the other hand, in semantic approaches (e.g. Dalrymple et al., 1991), the reconstruction of ellipses is seen as a form of anaphora resolution and works at the level of semantic representation. The ellipsis in (4.24a), repeated here, is accounted for by the reference of *does* to an entity or a concept mentioned before.

(4.26) Bill [likes playing golf]$_1$ and George does$_1$ too.

The problem here is that cases of ellipses involving a violation of syntactic binding constraints are not ruled out. A syntactical account would predict (4.27) with an intended coreference of *he* and *Bob* as unacceptable due to a violation of Binding Condition C[8], while purely semantic accounts have no means to exclude this reading.

8 Binding Condition C (roughly): A referring expression must not be bound (cf. Chomsky, 1981; Büring, 2005). The intrasentential property of an expression to be *bound* amounts to having an antecedent in the same sentence, e.g. a pronoun or a trace.

(4.27) *John defended Bob_1, and he_1 did too.

Kehler claims that the resolution of ellipses interacts with the establishment of coherence relations. Accordingly, ellipses are resolved in different ways depending on the type of coherence relation involved. On the one hand, for establishing Contiguity and Cause-effect relations, a semantic representation is reconstructed. Subsequently, in discourse segments connected by these relations, different syntactic constructions in ellipsis and antecedent are possible. For instance in example (4.28), the verb phrase in the first segment is in passive voice while the elided verb phrase has active voice.

(4.28) This problem was to have been looked into, but obviously nobody did. [look into the problem]

(4.29) I expected $Bill_1$ to win even when he_1 didn't. [expect $Bill_1$ to win]

Effects of condition C do not arise. Sentence (4.29) with a coreferential reading of *Bill* and *he* is perfectly acceptable. Syntactic parallelism is not necessary. For the establishment of coherence, only access to semantic content on the sentence level is required.

On the other hand, in order to establish Resemblance relations, syntactic parallelism is required. In these cases, a syntactic representation is reconstructed. As a consequence, active-passive voice alternations are not possible (4.30), and binding conditions must be respected (4.31).

(4.30) #This problem was looked into by John, and Bob did, too. [look into the problem]

(4.31) *Mary introduced $John_1$ to everyone, and he_1 did too. [introduce $John_1$ to everyone]

In order to find parallel arguments of the coherence relation, syntactical knowledge must be considered. VP-ellipsis indicates that parallel elements are shared or contrasted and can be reconstructed.

Gapping

Another phenomenon where Kehler's taxonomy accounts well for the data is Gapping. This type of of elusion is illustrated by the following examples:

(4.32) a. Sue became upset and Nan became downright angry.

b. Sue became upset and Nan $\emptyset$ downright angry.

Discourse (4.32a) has two readings. According to the symmetrical reading, both events are independent of each other. According to the asymmetrical reading, the first event is understood as the cause of the second. Discourse (4.32b) involving a Gapping construction only has a symmetrical reading. To illustrate this point, a context favoring a symmetrical reading is given in (4.33), and a context inviting an asymmetrical reading is shown in (4.34).

(4.33) Sue and Nan had worked long and hard for Carter. When Reagan was declared the winner, Sue became upset and Nan [became / $\emptyset$] downright angry.

(4.34) Susan's histrionics in public have always gotten on Nan's nerves, but it's getting worse. Yesterday, when she couldn't have her daily Egg McMuffin because they were all out, Sue became upset, and Nan [became / #∅] downright angry.

Kehler offers the following solution. On the one hand, establishing a Resemblance relation involves access to the semantics of constituents inside the conjuncts, i.e. shared or contrasted relations and parallel entities. On the other hand, establishing a Cause-effect relation only requires access to semantics on the sentence level, i.e. the propositions P and Q in the definitions of these relations. Gapping is only successful if the elided verb can be reconstructed syntactically. This is the case in Resemblance relations. In the asymmetrical reading of (4.32), a Cause-effect relation is established where a reconstruction of the verb is not possible.

An alternative explanation for this phenomenon is provided by Hendriks (2004). According to her, the Cause-effect reading disappears because the Resemblance relation is getting more salient by the existence of a contrastive topic. The function of a contrastive topic is to indicate that the sentence is a partial answer to an implicit question (Krifka, 1998). There are other possible answers which are alternatives to the uttered sentence. Typically, contrastive topics are realized by a so-called bridge accent (cf. Büring, 1997; Steube, 2003, see also the discussion in section 4.3.3.1).

Discourse (4.33) is about the pair of individuals Sue and Nan, while discourse (4.34) has Sue as topic. The two subjects in (4.33) are interpreted as contrastive topics whereas in (4.34) they are not. Both conjuncts of the second sentence in (4.33) are partial answers to the same implicit question: What happened to Sue and Nan when Reagan was declared the winner?

Thus, an additional device for telling apart coherence relations is the construction of (contrastive) topics. In Cause-effect and Contiguity relations, the second segment builds on the first one and the topic can be moved. In Resemblance relations, there must exist a common topic, but every segment is related independently to the preceding discourse. In section 4.3, I will have a closer look at the notion of topic in discourse.

Pronoun Resolution

A third linguistic phenomenon Kehler explicitly wants to account for is pronoun resolution. In a nutshell, Kehler argues that pronoun resolution and coherence establishment affect each other. On the one hand, preferences for pronoun resolution can change according to the coherence relation involved. On the other hand, in order to establish coherence relations, pronouns must be resolved. This interaction works in a different fashion in each of the three groups of relations. In Resemblance relations, a pronoun refers back to a parallel antecedent which is not necessarily the semantically most plausible referent. In Cause-effect relations, a pronoun refers to a plausible antecedent which enables coherence establishment. Saliency plays some role, as well. In Contiguity relations, a pronoun refers to the aboutness topic of the text passage which typically is in a salient position, e.g. the subject position in English.

4.2.2.5 Problems with Kehler's Theory

Kehler's account of coherence relations is both theoretically very clear and empirically motivated. Nevertheless, there are a few problems I do not want to gloss over. First, Kehler

does not say much about the crucial problem of how the proposed relations are to be inferred. As Kehler (2002, p. 32) notes himself, there is no robust mechanical procedure which would return the appropriate coherence relations for arbitrary examples. Kehler proposes to use intuitive judgments about possible paraphrases or substitutions with explicit discourse connectives. Resemblance relations can be made explicit by inserting the connectors *and [likewise]* or *too* for PARALLEL, *but* for CONTRAST, *for example* for EXEMPLIFICATION, *in general* for GENERALIZATION, and *that is* for ELABORATION. Similarly, explicit connectives for Cause-effect relations include *and [as a result]* or *therefore* for RESULT, *because* for EXPLANATION, *but* for VIOLATED EXPECTATION, and *even though* for DENIAL OF PREVENTER. For example, discourse (4.16) above can be paraphrased as (4.35a) but not as (4.35b), clearly indicating a Resemblance relation. In contrast, (4.36) shows the opposite pattern, permitting only a Cause-effect relation.

(4.35) a. John McCain is extraordinarily energetic for a 71-year-old, and likewise Barack Obama is remarkably grounded for a 46-year-old.

b. John McCain is extraordinarily energetic for a 71-year-old, #and therefore Barack Obama is remarkably grounded for a 46-year-old.

(4.36) a. If John McCain wins Michigan, then he wins Ohio, and therefore he wins the election.

(TIME.com, July 31, 2008)[9]

b. If John McCain wins Michigan, then he wins Ohio, #and likewise he wins the election.

However, results are not always unambiguous. The following newspaper text (4.37) exhibits markers indicating both relation types Resemblance and Cause-effect.

(4.37) The researchers found that the grooves kept the bacteria moving in straight lines, and as a result the cellulose fibers were straight, too.

(The New York Times, October 15, 2002)[10]

To conclude, Kehler provides a useful general distinction of different classes of coherence relations, though he does not provide us with a means to infer them in a systematic way. More about this issue is said in SDRT, to which I will turn shortly.

A second problem in Kehler's taxonomy (and not only there) concerns the discourse relation ELABORATION. The question is whether this relation is unambiguously a Resemblance relation, or maybe rather Contiguity. Kehler's definition of ELABORATION appears to be too simple. The requirement of identity of properties and entities is very restrictive, and many potential cases fall out of this definition. Usually it is assumed that in ELABORATION relations both segments relate to only one eventuality. However, the elaborating phrase (or its topic) often relates only to a part of the eventualities and concepts described in the elaborated segment. Sometimes, only one or some stages or sub-events of an event are specified, as in (4.38).

9 http://swampland.blogs.time.com/2008/07/31/inside_the_new_battleground_po/
10 http://query.nytimes.com/gst/fullpage.html?res=9D02EFDB153AF936A25753C1A9649C8B63

(4.38) Nicholas flew from Austin to Paris. He took off at 6 am. He landed at 2 pm. (Danlos, 2001)

ELABORATION relations differ from other discourse relations in various respects. As Knott (1996) observes and elaborates in Knott et al. (2001), there is no characteristic keyword for the identification of ELABORATION relations. Moreover, these relations depend on both the attentional states of interlocutors and on the information structure of involved utterances. For further elaboration of this point, see Knott (1996); Knott et al. (2001); Danlos (2001); Asher and Lascarides (2003).

Many authors who propose taxonomies of discourse relations, among them Taboada and Mann (2006b), point out that their taxonomies are not fixed, and that there are many overlaps among existing hierarchies. However, a problem with a subjective and open-ended list of discourse relations was pointed at by Knott and Dale (1994). They discuss the possibility that, in principle, even for incoherent texts one can define relations describing them. For example, in order to account for the incoherent discourse (4.39), one can imagine a discourse relation INFORM-ACCIDENT-AND-MENTION-FRUIT.

(4.39) #John broke his leg. I like plums.

This example nicely illustrates the difficulty in establishing a well-founded taxonomy of rhetorical relations. Let us now turn to a more descriptive approach, which has been applied to large amounts of real texts.

4.2.3 Rhetorical Structure Theory

Rhetorical Structure Theory (RST, Mann and Thompson, 1988; Marcu, 1997, 2000; Taboada and Mann, 2006b, a) emerged, in the first place, as a theory of text generation. It is a descriptive theory of how discourse segments or text spans compose a complex text structure by means of rhetorical relations. Two types of basic discourse units are assumed, and RST relations are typically asymmetrical. They express a relationship between a discourse segment carrying the main information (*nucleus*) and a second, subordinated segment (*satellite*). In these nucleus-satellite relations, the nucleus is essential and cannot be omitted while the satellite is optional. The representation of these relations as a tree is exemplified in Fig. 4.9. Besides that, there are also symmetrical relations, which have two or more nuclei, the multinuclear relations. In these cases, all text spans connected by the relation are equally essential. The tree structure is given in Fig. 4.10.

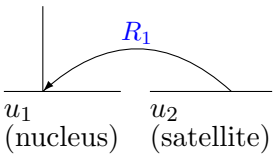


Figure 4.9: A nucleus-satellite relation

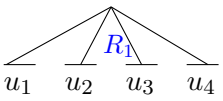


Figure 4.10: A multinuclear relation

Rhetorical relations are defined in terms of constraints on the nucleus and on the satellite or on their combination, and in terms of the intended effect achieved on the text recipient. This intended effect can be the mere recognition of the discourse relation by the addressee (subject matter relations, see below), or an influence on his beliefs, desires, and

intentions (presentational relations, see below). For example, the relations EVIDENCE and ELABORATION have the following semantics (Mann and Thompson, 1988, pp. 251, 273).

Evidence

Constraints on Nucleus: The reader might not believe the nucleus N to a degree satisfactory to the writer

Constraints on Satellite: The reader believes the satellite S or will find it credible

Constraints on the N+S-Combination: The reader's comprehending S increases the reader's belief of N

Effect: The reader's belief of N is increased.

Locus of the effect: N

Elaboration

Constraints on the N+S-Combination: 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 following ways ($\langle N, S\rangle$-pairs): $\langle set, member\rangle$, $\langle abstract, instance\rangle$, $\langle whole, part\rangle$, $\langle process, step\rangle$, $\langle object, attribute\rangle$, $\langle generalization, specific\rangle$

Effect: The reader recognizes the situation presented in S as providing additional detail for N. The reader identifies the element of subject matter for which detail is provided.

Locus of the effect: N and S

For illustration, consider the following example from an annotated corpus of German newspaper texts (Stede, 2004b).

(4.40) a. Die Bundeswehr geht davon aus, dass ihre geplanten Aktivitäten aus der Luft "keine relevanten Auswirkungen auf die Entwicklung der Region im touristischen Bereich" haben werden.

b. Jahrelange Erfahrungen in anderen Gebieten der Bundesrepublik hätten gezeigt, dass militärische Einrichtungen durchaus mit den Interessen des Tourismus in Einklang zu bringen seien.

c. So jedenfalls steht es in der Erläuterung zum Luft/Bodenschießplatz bei Wittstock.

a. The German Federal Armed Forces assume that their planned airdrop activities "will not have any relevant consequences for the development of the region in the field of tourism".

b. Years of experience in other regions of the Federal Republic of Germany had shown that military facilities could indeed be brought in line with the interests of tourism.

c. At least this is what is said in the explanatory report to the air/ground shooting range near Wittstock. (maz9612)

The proposition expressed by the second segment provides EVIDENCE for the proposition expressed by the first segment, and the complex segment containing both segments is elaborated by the third segment, thus constituting an ELABORATION relation. A typical

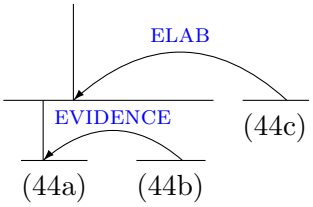


Figure 4.11: RST tree for (4.40)

visualization of the text structure of (4.40) in RST's graphical tree notation is given in Fig. 4.11.

In the original paper of Mann and Thompson (1988), a set of 24 rhetorical relations was defined[11]. They were grouped according to the intended effect into *subject matter relations* like Elaboration, Circumstance, Cause, Restatement and *presentational relations*, e.g. Evidence, Background, Enablement, Concession. This differentiation is related to what other authors (van Dijk and Kintsch, 1983; Sanders et al., 1992; Sanders, 1997) called the distinction between semantic and pragmatic relations. According to Sanders (1997), semantic relations hold between propositions and depend on the locutionary force, whereas pragmatic relations hold between speech acts and therefore rely on the illocutionary force.

This distinction is not the only possible classification of rhetorical relations. A group of authors have tried to replace RST's vague descriptions involving hard to define intentions by a classification based on usage data (Sanders et al., 1992; Knott and Dale, 1994; Knott, 1996; Knott and Sanders, 1998) in terms of basic cognitive concepts. Their approach is not restricted to RST; it rather aims at bringing more light into discourse structure in general. The asssumed underlying cognitive concepts include following primitives: (i) basic operation, (ii) polarity, (iii) source of coherence, and (iv) order of segments. The basic operation can be causal (a relevant causal connection exists between segments) or additive (all other cases). Polarity can be positive (if the basic operation links the content of the two segments as they are) or negative (if the content of one segment is linked with the negation of the content of the other segment). Discourse markers indicating negative polarity are *but*, *whereas*, *nevertheless*. The source of coherence is either semantic or pragmatic (see above). The order of segments can be basic (e.g. cause - consequence) or non-basic (e.g. consequence - cause). All these distinctions combined will yield a well-structured set of coherence relations and discourse markers. Despite this, the actual set of of explicit connectives can differ across languages, since not for all coherence relations there is always an explicit discourse connective. For instance, there is no direct English equivalent for German *denn* marking a relation which is causal and positive, has a semantic source and involves non-basic order. As noted earlier, some relations are never (e.g. Elaboration, Evaluation, Enablement) or rarely marked (e.g. Background).

Another differentiation made in the RST set of rhetorical relations is the distinction of volitional and non-volitional causal relations. For instance, Volitional Cause requires the satellite being a cause for the nucleus and additionally, the nucleus presenting

11 In more recent versions of the theory the number of relations is further reduced, so, although the RST discourse corpus (see below) uses 78 relations, they can be grouped to form a set of 16 more general relations.

a volitional action. For example, the German discourse marker *dadurch dass* indicates a NON-VOLITIONAL CAUSE while the marker *weil* (like English *because*) is not specified for volition. As a consequence, utterance (4.41a) is odd because staying at home is a volitional action. In contrast, both variants in (4.42) are acceptable since they involve only non-volitional circumstances[12].

(4.41) a. *Dadurch dass es regnet, bleibe ich zuhause.
because-of-the-fact that it rain$_{PRS.3SG}$ stay$_{PRS.1SG}$ I at-home
'Because of the fact that it is raining, I'll stay at home.'

b. Weil es regnet, bleibe ich zuhause.
since it rain$_{PRS.3SG}$ stay$_{PRS.1SG}$ I at-home
'Because it is raining, I'll stay at home.'

(4.42) a. Dadurch dass es regnet, wird die Straße naß.
because-of-the-fact that it rain$_{PRS.3SG}$ will$_{PRS.3SG}$ the street wet
'Because of the fact that it is raining, the street is getting wet.'

b. Weil es regnet, wird die Straße naß.
since it rain$_{PRS.3SG}$ will$_{PRS.3SG}$ the street wet
'Because it is raining, the street is getting wet.'

Although intentions play an important role in the definition of RST relations, Taboada and Mann (2006b, p. 10) point out that while the reason for connecting text spans is to create an effect on the reader which may well be considered an intention, intentions can also be satisfied by a single utterance. Seen in this way, intentions are unary functions that apply to a single proposition while rhetorical relations apply to at least two arguments, i.e. discourse segments or propositions expressed by them.

RST structures are trees, and the most important principle is that of *nuclearity*. It can be defined in the following way: "if a rhetorical relation holds between two textual spans of the tree structure of a text, that rhetorical relation also holds between the most important units of the constituent spans" (Marcu, 1997, p. 31). In other words, if a rhetorical relation links two segments, the relation eventually holds for the nuclei of the two segments. If the nuclei are complex themselves, this principle can be applied recursively.

A more complex example (4.43), from the same corpus as (4.40), together with its graphical representation in Fig. 4.12, shall illustrate how a text is composed by the recursive connection of rhetorical relations.

(4.43) a. Angriff ist die beste Verteidigung.
b. Wer wüsste das besser als die Strategen von der Bundeswehr.
c. In Sachen Bombodrom orientieren sich die Militärs jedenfalls ungeniert am Vorbild des alten "Marschalls Vorwärts" namens Blücher.
d. Denn eines der Hauptargumente der Schießplatz-Gegner – die Gefahr für den Tourismus – wollen sie mit einem Frontalangriff aus dem Weg räumen.
e. So heißt es in einem Bundeswehr-Papier knapp , dass das Bombodrom "keine relevanten Auswirkungen" auf den Tourismus haben wird.
f. Problem erledigt.
g. – Das ist schon dreist.

12 These examples are due to Markus Egg, from a talk held at the University of Leipzig in 2008.

(4.43') a. Offence is the best defense.
b. Who knew it better than the strategists of the Federal Armed Forces.
c. At least regarding the "Bombodrom" (air/ground shooting range), the militars uninhibitedly took the old "Marshal Vorwärts" called Blücher as an example.
d. Because one of the main arguments of the shooting range's opponents – the threat for tourism – they want to get away with it by means of a frontal attack.
e. So a paper of the armed forces briefly says that the shooting range would not have "any relevant consequences" on tourism.
f. Problem solved.
g. – This is rather bold. (maz16250)

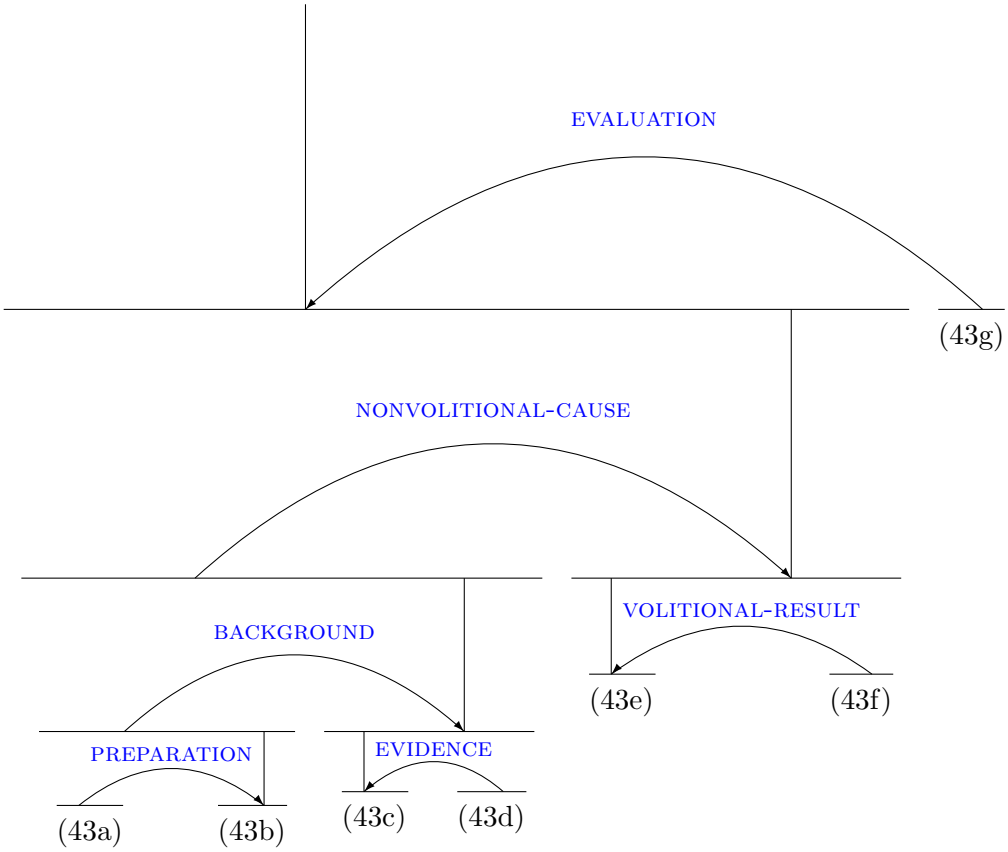


Figure 4.12: RST tree for (4.43)

RST is now widely used in various areas beyond text generation, mostly in computational linguistics, among them text summarization, parsing, and machine translation. For a recent overview, see Taboada and Mann (2006a) and the RST web site[13]. A main advantage of RST is the availability of rather clear instructions and useful tools for annotating texts with RST relations. It has been applied to many different text sorts, and some large annotated corpora in several languages are available, e.g. the RST Discourse Treebank [14] and the Potsdam Commentary Corpus (Stede, 2004b).

13 http://www.sfu.ca/rst
14 non-free; available from the Linguistic Data Consortium

4.2.4 Rhetorical Relations in SDRT

Whereas RST mainly aims at *describing* text structures, a theory with a rather different objective has been developed by Asher and Lascarides (Lascarides and Asher, 1991; Asher, 1993; Asher and Lascarides, 2003): Segmented Discourse Representation Theory (SDRT). The business of SDRT is explaining, apart from describing, discourse structures. Thus, one of its main concerns is the recognition of unmarked discourse relations. As I have already pointed out in section 4.1.2, there are many cases in which a discourse relation holds between clauses or utterances without being explicitly indicated by a discourse marker. In order to arrive at this goal, various different knowledge sources have to be taken into account. Among these are the formally analyzed linguistic input and prior context, contextual knowledge specific to the discourse situation, information about discourse participants and their intentions and goals, and general knowledge about the world. In a special nonmonotonic logic, the *glue logic*, all these knowledge sources are connected, and discourse relations are defeasibly inferred.

SDRT builds upon a dynamic semantic theory such as DRT (cf. section 3.4.3 on page 85). The DRS construction algorithm creates an underspecified semantic representation for individual sentences. These are input to the computation of complex discourse structures which consist of elementary discourse units, i.e. discourse segments represented as DRSs, and connections via rhetorical relations between them.

In this section, I will direct the attention to the inventory of SDRT relations and their classification, as well as outstanding properties that are possibly different from the taxonomies discussed so far. The detailed architecture and the logics used in SDRT will be discussed in section 5.3.

SDRT draws a basic distinction between two types of rhetorical relations, *coordinating* and *subordinating* relations, which allows the construction of complex structures for discourses. Basically, a simple SDRS (Segmented Discourse Representation Structure) is a labelled logical form for a sentence, i.e. a DRS. Complex graph structures for discourses, i.e. complex SDRSs, are constructed according to the following construction rules:

Complex Graph Structures

(i) Every discourse segment is a node in the graph,

(ii) every subordinating relation draws a vertical edge between two nodes that is directed downwards,

(iii) every coordinating relation draws a horizontal edge between two nodes.

These rules impose some constraints on what are possible graph structures. First, two edges cannot be subordinating and coordinating at the same time. Second, more than one relation of the same type can hold between two segments (in contrast to RST). Thus, many SDRSs can be represented as trees but some cannot. Another consequence, to which I will turn in more detail in the last part of this thesis, is the insight that anaphora resolution and discourse update depend on the structure of the graph. In particular, the Right Frontier Constraint (RFC), introduced in section 4.1.3.3 above, can be spelled out as follows: A coordinating relation pushes the right frontier to the right, closing off its current attachment point, and a subordinating relation extends the right frontier downwards, leaving open its attachment point. In SDRT, an antecedent for an anaphoric expression must be DRS-

accessible on the right frontier (Asher and Lascarides, 2003)[15].

A typical example discourse which shows the main characteristics of SDRT is (4.44). This discourse can be represented as a directed acyclic graph as shown in Fig. 4.13. Note that, apart from nodes π_i representing labels for discourse segments (e.g. π_1 is a label for segment (4.44a)), the graph also features abstract topic nodes[16] π' and π'', which emerge as a consequence of CDP, to which I will turn shortly. In the graph structure it becomes clear that π_5 (4.44e) can be attached to π_2 (4.44b), which lies on the right frontier of the graph.

(4.44) a. John had a great evening last night. (π_1)

b. He had a great meal. (π_2)

c. He ate salmon. (π_3)

d. He devoured lots of cheese. (π_4)

e. He then won a dancing competition. (π_5) (Asher and Lascarides, 2003)

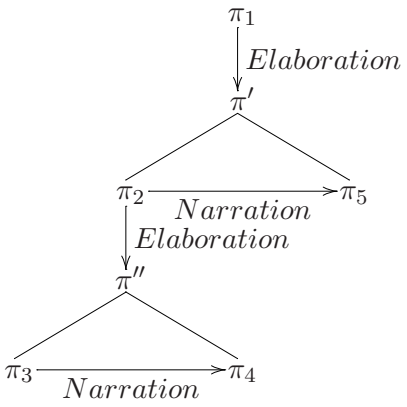


Figure 4.13: SDRT graph for (4.44)

In order to illustrate the different structures of SDRT and RST, consider again example (4.40), repeated below in its English translation. The corresponding SDRT graph is depicted in Fig. 4.14, where π_1 labels the representation of (4.40a), π_2 corresponds to (4.40b), and π_3 stands for (4.40c).

(4.40) a. The German Federal Armed Forces assume that their planned airdrop activities "will not have any relevant consequences for the development of the region in the field of tourism".

b. Years of experience in other regions of the Federal Republic of Germany had shown that military facilities could indeed be brought in line with the interests of tourism.

c. At least this is what is said in the explanatory report to the air/ground shooting range near Wittstock. (maz9612)

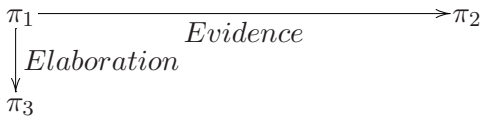


Figure 4.14: SDRT graph for (4.40)

The distinction of coordinating and subordinating discourse relations is supported by a row of characteristics of texts. The type of a relation affects, among other things, the temporal order of narrative texts, communicative intentions, and topicality. Coordinations typically indicate a temporal progression of events, whereas subordinations break this progression. With respect to communicative intentions, coordination amounts to Grosz and Sidner (1986)'s SATISFACTION-PRECEDENCE, while subordination corresponds to DOMINANCE. I will look more closely on the notion of topicality in discourses in section 4.3.

In order to decide whether a discourse relation is coordinating or subordinating, Gómez Txurruka (2003) proposed that the conjunction "and" is a marker of coordination. Every relation holding between two segments connected by "and" is coordinating. Although this is a very clear and easily applicable test, it does not explain which type of relation is to be assumed in cases not marked by "and".

Asher and Vieu (2005) proposed four tests for checking whether a relation is coordinating or subordinating. The starting point are two prototypical relations. NARRATION is the most typical example for a coordinating relation, and ELABORATION is the prototypical case of a subordination. I will not go into details of the tests here, and refer the reader to Asher and Vieu (2005) instead. In sum, the tests give an implicit definition of the two types of relations. Most importantly, Asher and Vieu found out that being subordinating or coordinating is not an intrinsic property of discourse relations but is rather only a default that can be overridden in certain contexts. Some relations, e.g. RESULT or CONSEQUENCE, that are normally coordinating, can behave like subordinating relations in particular instances, e.g. in (4.45).

(4.45) a. Lea bought a new car. b. As a result, she'll be able to go to Mexico this Christmas, c. and she will get to work quickly. d. It's a Subaru. (Asher and Vieu, 2005, p. 606)

There is a RESULT relation between (a) and (b), as indicated by the marker "as a result". The marker "and" indicates a CONTINUATION between (b) and (c). The pronoun in (d), which is clearly coreferential with the car introduced in (a), can only be resolved if the RESULT relation is assumed as being subordinating. If it were coordinating, then segment (a) would not be on the right frontier of the discourse, and hence the car would not be an available antecedent for the pronoun in (d). Chapters 5 and 6 of this thesis contain more details on how availability of antecedents for anaphora is constrained in SDRT.

An important structural constraint on SDRT graphs is known as *Continuing Discourse Patterns*. It can be stated as follows (for the first part, cf. Asher and Vieu, 2005, p. 595):

Continuing Discourse Patterns (CDP)

15 More details on the RFC in SDRT will be given in chapter 5 on page 158.
16 Discourse topics are subject of section 4.3.

1. If $R_1(\alpha, \beta)$ and $R_2(\beta, \gamma)$ and $Subord(R_1)$ and $Coord(R_2)$ then $R_1(\alpha, \gamma)$ and $Continuation(\beta, \gamma)$ (in addition to $R_2(\beta, \gamma)$).
2. If $R_1(\alpha, \beta)$ and $R_1(\alpha, \gamma)$ and $R_2(\beta, \gamma)$ and $Subord(R_1)$ and $Coord(R_2)$ then $\exists \pi'$ such that $R_1(\alpha, \pi')$ and π' is a label for $R_2(\beta, \gamma)$.

The first clause of this principle says that coordinated constituents must behave homogeneously with respect to a third superordinated constituent. CONTINUATION, a coordinating relation, is like NARRATION but lacks the spatio-temporal consequences of the latter. The second clause rewrites graph structures involving two occurrences of the same subordinating relation R_1 holding between α and both β and γ by introducing an intermediate implicit topic node π', which is made of β and γ linked by at least a CONTINUATION relation. The working of CDP is shown graphically in Fig. 4.15, where CDP turns a discourse structure (a) into a structure (b), which, in turn, is always represented as (c).

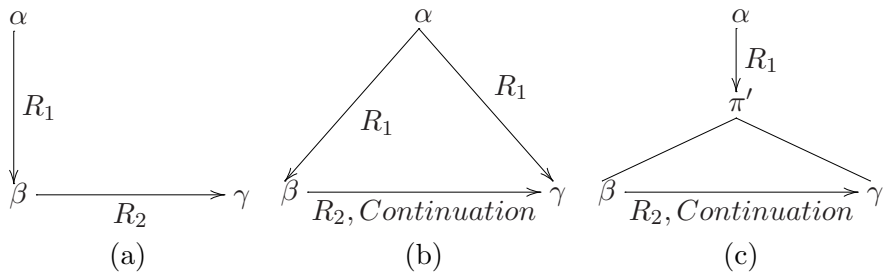


Figure 4.15: Continuing Discourse Patterns (CDP)

SDRT draws another distinction among discourse relations. *Veridical* relations semantically entail the contents of both discourse segments involved, while *non-veridical* relations do not entail the content of at least one of their arguments. In chapter 5, we will see that this property is crucial for defining the semantics of rhetorical relations. A recent annotation scheme for annotating texts with SDRT relations (Reese et al., 2007) proposes an inventory of 14 discourse relations, which are listed in Table 4.1.

	Coordinating relations	Subordinating relations
veridical	Continuation, Narration, Result, Contrast, Parallel, Precondition	Background, Elaboration, Explanation, Commentary, Source
non-veridical	Consequence, Alternation	Attribution

Table 4.1: Inventory of SDRT relations (Reese et al., 2007)

4.2.5 Conclusion

In this section, we have seen some different conceptions of discourse relations. The taxonomy of relations varies considerably across theories, as well as their classification and their properties. However, there seems to be a core inventory of discourse relations that different theories largely agree upon. For the purposes of this thesis, I will take the SDRT inventory

(cf. Table 4.1) as a basis, and with it, the distinction of coordinating and subordinating relations.

4.3 Discourse Topic

The notion *discourse topic* has been extensively discussed in the literature for the last 30 years. There are many proposals but an exact formulation what it is and how it has to be conceived was not made so far. In this section, I will review the main proposals and point out their differences. As Stede (2004a) summarizes, a discourse topic can be understood

1. as an **entity** that is salient or prominent and of which can be said that the discourse is "about" it,

2. as a **question** that is either explicit or implicit and that is answered by the discourse,

3. as a **proposition** that the hearer has to construct interpreting a discourse and which has consequences for the interpretation of the ongoing discourse.

These are quite distinct conceptions, though related to each other. I will discuss each of these alternatives in the following sections, but before, we should take a look at the use of the term "topic" within a sentence.

The terms "topic" and, related to it, "comment", on the sentence level have received many different conceptions across the linguistic literature (cf. *inter alia* Reinhart, 1982; Vallduví, 1992; Hajičová et al., 1998; Büring, 1997). In particular, as these terms have to do with the information structure of texts[17], they interact with other phenomena at this level such as the partition of a sentence into "focus" and "background", or "given" and "new", or "contextually bound" and "informative". I do not want to discuss all the different proposals made so far concerning the interaction of the two pairs of terms (see Jacobs, 2001), but some words are needed in order to clarify the terms used here.

Already in the second half of the 19th century, Georg von Gabelentz and Hermann Paul observed that a sentence can be divided into a psychological subject and a psychological predicate, which do not necessarily correspond to the grammatically defined subject and predicate in a sentence. This distinction has received other terms as "theme" and "rheme", or also, topic and comment. The topic is commonly understood as that what the sentence is about, i.e. the psychological subject of the sentence. This notion is also known as "aboutness topic", following Reinhart (1982). The corresponding comment is what is said about the topic, i.e. the psychological predicate. I will refer to this concept as *sentence topic*, in contrast to *discourse topic* which refers to the subject dealt with by a bigger span of text, i.e. a discourse segment, or a discourse as a whole.

4.3.1 Discourse Topic as Entity

Intuitively, it seems clear that, in general, a discourse deals with one or more particular entities. In many cases, recipients would choose one of them as particularly prominent and

17 The term *information structure* refers to the partition of utterances reflecting the speaker's view on the hearer's information state at the time of the utterance. The term *information packaging* refers to the way this structuring is realized by syntactic, morphological, and prosodic means (cf. Vallduví, 1992).

say that the discourse is about that entity. Although this idea seems very appealing, it is difficult to make out precisely which entity is the topic of a particular text.

A difficulty that many authors gloss over, as Averintseva-Klisch (2008, p. 97) points out, is the question whether "entity" is understood as a real world entity or an entity on the discourse level, i.e. a discourse referent. Actually, there is a three-way distinction between a linguistic expression which refers to an entity, the entity in the discourse model or discourse referent, and the entity in the real world. The confusion in the literature and the lack of a uniform definition is for a great part due to this distinction. I agree with Averintseva-Klisch (2008) in that the notion of discourse topic should be settled at the level of discourse representation, although I will discuss the question if a concept of topic is central to discourse or just emerges as an epiphenomenon.

In this spirit, a way of conceiving sentence topics is to think of an address in the discourse model of the recipient under which the information given by the sentence is stored. This approach is taken by Vallduví (1992) who calls the topic the *link* and the comment the *tail* of a sentence. The recipient has to choose the address among contextually available alternatives.

This conception of aboutness can be raised easily to the discourse level. The discourse topic is the mental address where the information conveyed by the discourse is stored, and subordinated entities are linked to this address.

However, although intuitively clear, there is no formal approach which unambiguously singles out a unique referent an arbitrary discourse is about. Hence, other proposals have been made which do not rely on entities with an unclear ontological status, but rather on propositions and questions.

4.3.2 Discourse Topic as Proposition

A great deal of the difficulties that arise when we have to give a principled way of finding the right entity which is the topic for a given discourse can be overcome by assuming that a discourse topic is a proposition. In an assertion, a speaker provides new information about this proposition, and, likewise, in a question she requests new information. This approach is taken by Keenan and Schieffelin (1976b). They derive a complex discourse topic structure from the propositions expressed by individual sentences.

This view on discours topics is adopted by various scholars who see the discourse topic as a proposition or a set of propositions. In general, approaches to discourse topics as propositions have to be careful with not confusing two opposite views on topics. On the one hand, discourse topics are seen as (a set of) propositions about something, and on the other hand, propositions are seen as that what discourses are about (cf. Averintseva-Klisch, 2008).

Among the supporters of the propositional view on topics is Asher (1993), albeit in (Asher, 2004a) he poses serious questions on the usefulness of an all-purpose notion of discourse topic, which I will discuss in the following. A discourse topic can be seen as "a proposition that summarizes the content of a constituent in an SDRS, and it bears a particular structural relation to that constituent" (Asher, 1993, p. 267). (Asher, 2004a) discusses several reasons why SDRT assumes an abstract notion of discourse topic in the first place.

First, discourse topics are needed for explaining coherence in discourses containing relations like Narration and Continuation. In the seminal versions of SDRT (Lascarides

and Asher, 1991; Asher and Lascarides, 2003), these relations are regarded as the default discourse relations to be assumed in case that there are no clues indicating another relation. In order to avoid overgeneration, a guarding criterion is used to restrict this default: a new utterance must share a common discourse topic with the one to which it is to be attached via Narration or Continuation.

As we have seen in section 4.2.4, a consequence of the principle CDP (Continuing Discourse Patterns) is that a Continuation relation emerges, together with an implicit discourse topic, if two coordinated segments are subordinated to a third one. As for coherence in Narration, consider the following example:

(4.46) a. #My car broke down. Then the sun set.

b. My car broke down. Then the sun set and I knew that I was in trouble.

(Asher, 2004a, p. 181)

This example is Asher's key example for motivating a discourse topic for Narration. The question is how to explain that (4.46a) is less coherent than (4.46b). Asher argues that in (4.46b), an implicit topic for the discourse can easily be constructed, while in (4.46a) it is difficult to find a common topic for both sentences, and one had to assume a topic which would be too general to have any descriptive or explanatory value.

This need for a propositional discourse topic has been questioned by various authors. Zeevat (2004) argues that although a common discourse topic may be a heuristic for indicating coordination, it is neither a necessary nor a sufficient condition for a coordinated relation to hold. Going further in this direction, Oberlander (2004) poses the question whether an abstract notion of discourse topic can be avoided in general, or at least be reduced to what is minimally required. He argues that the phenomena for which Asher needs a notion of an implicit, constructed discourse topic can be explained without topics. Oberlander points out that in example (4.46), the assumption of a discourse topic is not necessary. Instead, the difference in coherence between (4.46a) and (4.46b) can be explained by the fact that there are causal relations in the event structure of the described circumstances. The first two events constitute causes for the speaker to realize that she was in trouble.

Asher argues that a similar notion of discourse topic is needed for Resemblance relations like Parallel, Contrast, and Alternation. As noted above in the discussion of Kehler's Resemblance relations (section 4.2.2.2), these relations exhibit parallel or contrasted predicates in the involved discourse segments. A way to infer these predicates is assuming a common discourse topic, which provides a relevant set of alternatives in the sense of Rooth (1985) or Krifka (1992). Such an account, however, shares many properties with the conception of discourse topic as an implicit question, to which I will turn in section 4.3.3 below.

In SDRT, discourse topics also play an important role in the Elaboration discourse relation. In difference to the before-mentioned relations, discourse topics are not needed to be constructed independently of this relation, rather they are built into the semantics of the relation itself, thus constituting "part and parcel" (Asher, 2004b, p. 257) of this relation. If an Elaboration relation between two segments can be inferred, then the superordinated segment that is elaborated on is the discourse topic of the elaborating segments. Since this is a logical consequence of a discourse relation, this kind of discourse

topic can be seen as an epiphenomenon of establishing discourse coherence, in line with the point of view taken by Kehler (2004).

Asher mentions another reason for assuming discourse topics. When connecting an utterance to the preceding discourse and when searching for antecedents of anaphora, discourse topics can be used for explaining how some parts of discourse are closed off for attachment. Asher assumes that discourse topics are sometimes needed as attachment points in discourse structures where new utterances can be connected to the preceding discourse. Consider the following example:

(4.47) a. I ate a lovely dinner.

b. I had quenelles de brochet.

c. I had salmon.

d. I had duck.

e. I had a nice wine.

f. I then went for a walk around the old city.

g. I slept well.

(Asher, 1993, p. 279)

According to Asher, a discourse topic is needed to explain the discourse pop in (4.47f), i.e. the fact that the discourse structure demands (f) be attached to (a) rather than to (e). Here, the role of discourse topic is twofold: The negative role is that the walk cannot be part of the (explicit) topic *the dinner*. The positive role is that the walk can be part of the (implicit) topic *things I did last night*.

Again, Oberlander (2004) questions this view and claims that for both points there is no need to fall back on the notion of topic. For the negative role accounts the fact that world knowledge enables the hearer to infer that the walk is not a sub-part of the dinner. For the positive role, the attachment can be explained again in terms of a causal relation: (a-e) together with (f) cause (g), and (g) is the Result of (a) and (f). Thus, Oberlander concludes that there is no need for an implicit, constructed, propositional discourse topic. He suggests an entity-based view of coherence, thinking of discourse topic in terms of a collection of sentence topics. Entity-based views of coherence see coherence in terms of what others call cohesion of a text. Recent approaches include Centering Theory (Grosz et al., 1995; cf. section section 3.4.2.2 above), Veins Theory (Cristea et al., 1998), or the theory sketched in Knott et al. (2001).

It should be noted here that, although the particular example (4.47) can be explained without discourse topics, it remains to be seen how this works for arbitrary texts. Sometimes it might be far-fetched, if not dangerous, to construct a suitable causal relationship. The same critique holds also for Oberlander's replacement of discourse topics for Narration. Thus, Asher is right in asking what then is required for Narration. Some sort of a thematic continuity must exist – something tricky to capture, but possibly well indicated by information structural clues, to which I will turn now.

4.3.3 Discourse Topic as Question

4.3.3.1 Contrastive Sentence Topics

As seen so far, although formally appealing, conceiving the discourse topic as a proposition seems to bring along a series of problems. Without fully giving up the formal advantages of a propositional view, we can think of the topic as a question that a discourse gives an answer to. Formally, a question can be thought of as an open proposition, or a set of propositions, and the answer singles out one member of this set. This idea is exploited both in Krifka's Structured Meanings account of questions (Krifka, 1992) and in Alternative Semantics (Rooth, 1985).

The central idea in Alternative Semantics is that a focussed constituent of a sentence evokes a set of alternatives which have the same semantic type and are contextually available. Any expression is assigned two semantic values: its ordinary meaning, and, additionally, a *focus value* which is a set of ordinary meanings. The two are distinct only for focussed expressions. This is illustrated by the following example[18].

(4.48) [GREG]$_F$ broke the kitchen window.

Here, the ordinary meaning of *Greg* is just Greg, while the focus value is the set consisting of contextually given alternatives. In this example, we can imagine a discourse situation where a group of children are playing football in the backyard, and the alternative set would be, say, $\{Max, Greg, Tom, Liv\}$.

This approach is taken by Büring (1997), who applied it to the semantics of topics at the sentence level, although in this important work the term sentence topic is used in the most part for a phenomenon also known as contrastive topic. As Asher (2004a) notes, caution is needed with the particular use of this term in order to avoid possible confusions. Contrastive topics arise, among other uses, when an utterance answers a question only partially. They can be grammatically marked, e.g. in Hungarian, or intonationally marked, e.g. in Germanic languages. Büring gives the following example with a specific *bridge contour*, or hat contour, intonation pattern (cf. also Steube, 2003). This pattern (L+H* for German; slightly different for English) is characterized by two intonational peaks, one rising and one falling, and its graphical representation resembles a bridge or a hat. A typical example is given in (4.49).

(4.49) a. A: What did the pop stars wear?

b. B: The [/FEmale]$_T$ pop stars wore [\ CAFtans]$_F$.

(Büring, 1997, p. 56)

The question in (4.49a) denotes a set of propositions $\{x \mid$ the pop stars wore $x\}$. The answer (4.49b) has a focus value $\{x \mid$ the female pop stars wore $x\}$ that does not match the question. Now Büring (1997) assumes that the contrastive topic induces alternatives to the focus value, and defines the *topic value* as a set of focus values, or a set of a set of propositions. The topic value for (4.49b) can be characterized by $\{y \mid \{x \mid$ the y popstars wore $x\}\}$. In this way, Büring can account for the special intonational pattern needed for (4.49b) to be a felicitous answer to (4.49a).

18 In these examples, the topic is marked by $[]_T$, the focus by $[]_F$, syllables bearing an accent are marked by UPpercase, rising intonation by '/', and falling by '\', respectively.

This conception of contrastive sentence topics can be useful for the construction of discourse topics. In the remainder of this section, some accounts which take the discourse topic as a question are discussed.

4.3.3.2 Topic-Comment Structures for Discourses

Van Kuppevelt (1995) presents an account in which topicality is the basic organizing principle of discourse structure. The central claim is that discourse segmentation is determined by a hierarchy of topic-comment structures. This structure emerges by the assumption of topic-forming questions. Often, these questions are not posed explicitly but remain implicit. The notion of topic is defined as follows: "Every contextually induced explicit or implicit (sub)question Q_p that is answered in discourse constitutes a (sub)topic T_p" (van Kuppevelt, 1995, p. 814). The topic is that which is being questioned and is identified with the set of possible answers to the underlying question. One member of this set is selected by the answer A_p. The comment C_p then is that what is asked for, and is provided by the answer A_p. If the answer is assumed to be satisfactory, the topic is closed off, otherwise it gives rise to further subquestions. For illustration, look again at example (4.48), repeated here.

(4.50) a. Greg broke the kitchen window.
b. A girl broke the kitchen window.

Sentence (4.50a) can be understood as an answer to the question *Who broke the kitchen window?* As mentioned above, the set of possible answers in the specific utterance situation would be the set $\{Max, Greg, Tom, Liv\}$. Note that the answer does not necessarily have to pick out a unique member of this set, since (4.50b) is another possible answer to the same question. In this case, there would be reason for further subquestions.

In this approach, discourse topics are closely related to sentence topics, in the sense that a discourse topic is seen as a topic of higher-order consisting of a set of sentence topics whose actuality lasts as long as there are subtopics arising from subquestions. It is defined as "the set of all topics T_p that have arisen as the result of indeterminacies provided by one and the same feeder F_i". A feeder is either a topicless discourse unit or a unit whose topic is not salient at utterance time. A simple example for the question-answer structure of dialogue (4.51) is given in Fig. 4.16.

(4.51) F_1 A: Last Tuesday our company got a new president.
A_1 A: It is the former manager of a successful software house.
Q_2 B: Why did your company choose a new president?
A_2 A: The former president failed to solve the increasing financial problems of the company.

(van Kuppevelt, 1995, p. 815)

Both questions Q_1 and Q_2 are induced by the feeder F_1. In the graphical representation, a horizontal line connects a question to the part of discourse to which it is directed, and a vertical line connects an answer to a corresponding question.

As an example for a structure of a monologic text, let us take again text (4.43) from the Potsdam Commentary Corpus and extend it by implicit topic-constituting questions (marked by <>).

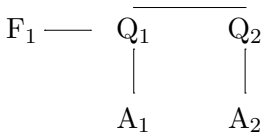


Figure 4.16: Question-answer structure for (4.51)

(4.52) F_1 Offence is the best defense.

Q_1 Who knew it better than the strategists of the Federal Armed Forces?

A_1 $< \emptyset >$

Q_2 <Why did nobody know it better than the militars?>

A_2 At least regarding the "Bombodrom" (air/ground shooting range), the militars uninhibitedly took the old "Marshal Vorwärts" called Blücher as an example.

Q_3 <What evidence is there for that?>

A_3 Because one of the main arguments of the shooting range's opponents – the threat for tourism – they want to get away by means of a frontal attack.

(maz16250)

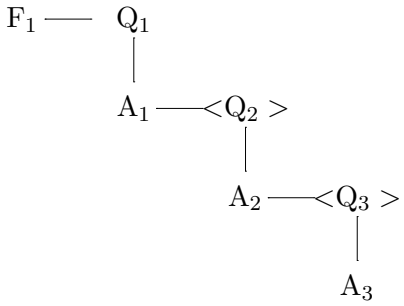


Figure 4.17: Question-answer structure for (4.52/4.43)

In this example, Q_1 is a topic-constituting question whereas the other (implicit) questions are subquestions that arise because Q_1 is not directly answered. Q_3 is subordinated to subquestion Q_2 because it can be interpreted as a question asked with the purpose of completing the answer to Q_2. As can be observed, the structures built by RST (Fig. 4.12) and by van Kuppevelt's theory represent quite different aspects of the same text. There is no isomorphic mapping from a question-answer structure to an RST tree. Although some information is encoded in both theories, e.g. subordination of A_3 to A_2, the more abstracted layers of representation differ substantially. In terms of discourse relations, the structures in van Kuppevelt's theory reflect only two different types of relations, that of dominance and precedence, very much as assumed in Grosz and Sidner (1986)'s theory.

4.3.3.3 Quaestio Theory

Another proposal on discourse topics as questions does not take sentence topics as the basis for constructing complex discourse topics, it rather takes the discourse level as a starting

point. In the approach taken by Klein and von Stutterheim (1987); von Stutterheim and Klein (2002), a discourse topic is understood as an explicit or implicit question answered by the corresponding discourse, or, as they call it, the *quaestio*. A basic assumption is that any coherent discourse can be understood as an answer to an introductory question. Consider the example given by von Stutterheim and Klein (2002, p. 69).

(4.53) a. A: Which car came from the left?

b. B: A BMW 730 came from the left.

b' B: From my position, I could not see it very well. Everything went so fast. But it was a big car, a limousine; blue, dark blue. One of my neighbours had such a car. I guess it must be very expensive, one of these old-fashioned dinosaurs.

Here, the speaker asks for the specification of a certain entity, and evokes with her question a set of alternatives from which the recipient has to select one, in this case the car that came from the left at the moment under discussion. The question is answered exhaustively by (4.53b), but as more usual in naturally occurring discourses, even a question for a particular entity often is answered not straightforwardly but rather presenting additional peripherical information that can, nevertheless, deliver very important facts and insights about the communicative intentions. Thus, in (4.53b'), only parts of the utterances are directly related to the question. This part is called *main structure* of the text, and the part giving additional material its *side structure*. Side structures often are essential for revealing the speaker's intentions and can consist of comments, evaluations, or background information.

Central to Quaestio Theory is the assumption that the underlying question determines not only the partition into main and side structures but also the structure of discourses in general. Other constraints on discourse structure defined by von Stutterheim and Klein (2002) concern the main structure and include the meaning assignment to topic and focus in an utterance, as well as constraints on anaphoric reference in terms of domains of reference and the referential movement within these domains. These constraints are defeasible; the speaker can disobey them, leading to side structures, or, in the spirit of Paul Grice, to particular rhetorical effects.

Often, in most non-dialogic texts, the quaestio is not stated explicitly, and it is by no means clear how to single out one underlying question. Both utterances (4.53b) and (4.53b') could have been answers to any of the questions *What came from the left?, What was the cause for the accident? What happened?* All these questions raise a set of alternatives, and an element of this set is specified by the answer. In a slightly different fashion from, though still close to, Büring (1997), the sentence topic is identified with this alternative set, and the element of this set which is actually chosen by the answer is called the focus. Hence, in this approach the information structure of an utterance depends on the quaestio.

The quaestio also determines the domains of reference in the text, not only of entities and events, but also the time frame, the spatial orientation and the modalities the speaker wants to express. These domains can be quite distinct for a specific question, e.g. (4.53a), and for a more general quaestio, e.g. *What happened?*.

Although it is intuitively clear that for any discourse a quaestio can be constructed, the approach does not explicitly state how an implicit quaestio is constructed from a particular

text. For many texts it seems difficult to single out a unique question the text answers. Furthermore, as Zeevat (2004) asks, it is not clear whether the speaker knows in advance the question she is going to address, or whether the question can be specified only with respect to the Common Ground of speaker and hearer. Thus, we remain with the suspicion that in general, a discourse topic defined as a question the discourse answers can only be constructed after the discourse is finished.

4.3.3.4 Questions Under Discussion

More light on this issue is shed by defining more precisely the information flow between conversational participants and the Common Ground, as discussed in chapter 2. There is an approach which takes questions to be the key notion for discourse structuring, but takes a starting point that is different from the proposals discussed so far. Instead of starting from single sentences or utterances, or texts as a sequence of sentences, one can start from the more general and most frequent type of occurrences of natural language, i.e. dialogues.

The notion of *Questions Under Discussion* was developed independently, though almost at the same time, by Roberts (1996) and Ginzburg (1996). The question under discussion is, in general, "What are we talking about?".

Questions under discussion are a partially ordered set of questions (QUD) reflecting the way in which they are related to each other, e.g. a subquestion can serve as elaborating a more general superquestion. This order is seen as underlying the partition of discourses in segments. The attachment of an utterance to the preceding discourse depends on which question an utterance is intended to address. A discourse segment is closed when the question is answered, if necessary by means of a sequence of acknowledgments. As for anaphoric accessibility, only entities in the answer and entities in the question are accessible.

As a result, a hierarchy of questions provides a shallow discourse structuring, in contrast to the more complex discourse structures built by other theories, such as the tree structures of RST, or the even less constrained graph structures for discourses assumed in SDRT. Although questions under discussion are sometimes identified with discourse topics, it can be stated more explicitly that the topmost entity in the set of questions under discussion corresponds to the discourse topic.

The questions in these accounts are closely related to the goals and intentions of the discourse participants. In fact, Roberts (2003) uses the terms "discourse goal" and "question under discussion" as synonyms. Thus, QUD pertains to the intentional structure from Grosz and Sidner (1986), and the notion developed by Roberts and Ginzburg takes us away from the conception of discourse topics in its aboutness sense. Moreover, as soon as we take intentions into account, we loose the formal transparency of a conception of questions as the set of answers.

4.3.4 Conclusion

In this section, we have seen that there is no uniform conception of discourse topic. Rather, there are approaches assuming topics to be of diverse ontological types stretching from propositions over questions to entities.

While propositions have a clear ontological status and formalizations can rely on a lot of work done in formal semantics, the main drawbacks of propositional accounts are twofold. First, it seems rather counterintuitive to think of propositions as that what discourses are

about. Second, a proposition is an abstract notion, and Asher (2004a) raises the discussion whether we need this additional abstract layer at all, or if it can be avoided. Asher sees discourse topics as an integral component of the semantics of certain discourse relations like NARRATION and PARALLEL, and points out that the establishment of these relations demands different kinds of discourse topics. In response to Asher's paper, Kehler (2004); Oberlander (2004); Stede (2004a); Zeevat (2004) take other points of view. Kehler (2004) sees discourse topics as an epiphenomenon of establishing discourse coherence, without the need for constructing them in an explicit step of the interpretation process. Facing all the difficulties in defining a proper notion of discourse topic, Oberlander (2004) suggests that the linguistic phenomena where a discourse topic seems to be needed for interpretation can be explained by other means. Zeevat (2004) argues that a discourse topic in form of a proposition or a question is something that can be constructed only after the discourse has taken place, thus is not essential neither for structuring nor for interpreting discourses.

These points also apply to the conception of discourse topics as questions. In cases where an explicit question is part of the discourse, there is no doubt that it must show up in the discourse representation. However, if an implicit question (or proposition) is to be constructed, it is not clear how to single out in a constrained way the unique question a discourse answers.

While the existence of an abstract proposition or question as discourse topic can be doubted, there is always some entity that is talked about. Hence, Oberlander (2004), as well as Stede (2004a) and Zeevat (2004) opt for an entity-based view on discourse topics. Zeevat (2004) makes an interesting proposal. He suggests the assumption of one or more *protagonists* that would give a usable notion of discourse topic. He draws from evidence in languages where protagonists are grammatically marked, especially Chinese and Spanish pro-drop phenomena. In Spanish, pro-drop is a very common phenomenon, but it only indicates that the antecedent was already topic. In Chinese, subject pro-drop can occur both within a sentence and across longer distances. The subject of a subordinate clauses may be dropped if it is identical to the main clause subject. If occurring between sentences, it is restricted to the hero of the story, or the protagonist. Zeevat argues that if protagonists are needed anyway in a story, then they could be available as anaphoric antecedents after stories about them.

I conclude in agreement with Averintseva-Klisch (2008) and define the discourse topic as an entity on the level of discourse representation, i.e. in the discourse model. A discourse topic is the discourse referent a discourse segment is about. It is the most salient discourse referent in a segment. Averintseva further claims that every coherent discourse segment has exactly one discourse topic. To maintain this assumption, for segments with equally salient referents, abstract plural entities must be assumed, which, nevertheless, still are discourse referents. In fact, plural antecedents are one of the reasons why Asher (2004a) assumes the notion of discourse topic. As he notes later (Asher, 2004b), the fact that one often does not know what the discourse topic is until the discourse is over can be accounted for by allowing discourse topics to remain underspecified. If we assume this point of view, we can easily agree with the assumption that every discourse segment has exactly one discourse topic. Regarding the abstract topic of ELABORATION in SDRT, the fact that discourse relations in SDRT introduce a label, and therefore a discourse referent, allows us to treat proposition-like topics as abstract entities on the level of discourse representation.

I will leave the matter here at this point for further research. A typological investigation on the marking of discourse topics across languages would certainly be insightful.

Chapter 5

Discourse Interpretation

As outlined in previous chapters, I start from the assumption that semantic forms are extracted from the linguistic information supplied by an utterance. These semantic representations are underspecified and largely context-independent, especially not influenced by intentions of conversation participants. Further contextual enrichment of these representations by means of pragmatic inferences yields the content of a discourse.

The sections in this chapter are devoted to theories of discourse interpretation. All of them use one of the formalizations of defeasible reasoning presented in chapter 1 (section 1.2) in order to account for inferences in discourse interpretation.

Section 5.1 discusses Interpretation as Abduction, a general framework of discourse interpretation in which pragmatic inferences are modelled in terms of abductive reasoning. (cf. section 1.2.4).

Section 5.2 deals with Model Generation, a framework that is suitable for the search of preferred models for discourses. It can be seen from an abductive perspective or as a form of circumscriptive default reasoning (cf. section 1.2.2).

Section 5.3 describes Segmented Discourse Representation Theory, a theory of discourse interpretation that uses Commonsense Entailment for representing defaults in the process of enrichment of underspecified semantic forms (cf. section 1.2.3).

5.1 Discourse Interpretation as Abduction

In the influential paper of Hobbs et al. (1993), the following steps are taken in order to interpret a sentence:

> Prove the logical form of the sentence,
> together with the constraints that predicates impose on their arguments,
> allowing for coercions,
> Merging redundancies where possible,
> Making assumptions where necessary. (Hobbs et al., 1993, p. 70)

The proof of the logical form, which emerges by syntactical analysis and semantic interpretation of the sentence, must be derived from the facts and rules in the presupposed knowledge base. The knowledge base contains premises in form of axioms in predicate calculus and is fed from different information sources. Linguistic knowledge from syntax and lexical semantics, as well as information about the cognitive state of the speaker, pragmatic maxims, and world knowledge, have an influence on the interpretation.

Hobbs and colleagues base their theory on the following general picture: in a discourse, both speaker and hearer have their own belief sets, and there is an overlapping set of beliefs

shared by both. In general, an utterance consists of the speaker's private knowledge and shared knowledge, and it aims at extending the shared knowledge by some beliefs of the speaker.

5.1.1 Flat Logical Forms

The concept of "logical form" used by Hobbs (1985b); Hobbs et al. (1993), which corresponds to the level of semantic form (cf. chapter 1), makes use of a non-intensional first-order predicate logic. Syntax is kept as simple as possible in order to be able to easily define inference processes. The logic is as "flat" as possible: modal and intensional operators, disjunctions, negations, and nested quantifications over predicates are avoided. Any semantic content is conceived by predicates, and constants are only place-holders. A motivation for this is the fact that in natural language, entities are not communicated directly, instead the hearer has to identify them by means of the uttered properties. Many linguists and philosophers try to keep the underlying ontology as "chaste" as possible by assuming as few different kinds of entities as possible. Often, one searches for the most simple conceptual schema explaining the observed phenomena. However, simplicity of a theory can be achieved not only by minimizing the number of entities, but also by reducing the complexity of the rules in the assumed system. In this way, the number of assumed entities is increased by allowing as entities everything what can be referred to by a noun phrase. Hobbs (1985b) calls this method "ontological promiscuity", moving away from a minimalistic ontology.

(5.1) a. A boy builds a boat.

b. $\exists x, y[build(x, y) \wedge boy(x) \wedge boat(y)]$

To illustrate this account, the simple sentence (5.1a) is assigned a logical form (5.1b). This representation does not differ much from a standard formal semantic notation. Differences show up when more complex linguistic phenomena like tense, intensional contexts and adverbials are taken into account. Instead of enhancing the complexity of the logic, the number of assumed predicates is increased.

In a Davidsonian style of event representation (Davidson, 1967), the predicates corresponding to verbs get an additional event argument which stands for the eventuality denoted by the verb. For example, sentence (5.2a) gets a logical form (5.2b) in which e_1 denotes a past event which expresses the desire e_2 of a boy x towards the quickness of an event e_3 which in turn is a building of a boat y by x.

(5.2) a. A boy wanted to build a boat quickly.

b. $\exists e_1, e_2, e_3, x, y[Past(e_1) \wedge want'(e_1, x, e_2) \wedge quick'(e_2, e_3) \wedge build'(e_3, x, y) \wedge boy(x) \wedge boat(y)]$

In short, the logical form of a sentence is a conjunction of atomic predicates in which all variables are existentially quantified with a scope as wide as possible.

5.1.2 Weighted Abduction

Abductive inference is the conclusion to the best explanation of an observation. Via abduction, the premises, or goal expressions, needed for the interpretation of a sentence are inferred. Additional assumptions are made in case the premises available in the knowledge base are not sufficient for proving the logical form. A subset of the expressions to be proven must be assumed.

Of course, for an observation there are many possible explanations. Criteria for choosing the best explanation are needed in order to restrict the search space of hypotheses which otherwise would be tremendous. The method used by Hobbs et al. (1993) imposes a metrics on the quality of explanations. This mechanism is called *weighted abduction* (Stickel, 1991). To select the best of various assumable goal expressions, numerical costs for their assumption are assigned to them. The optimal proof of a logical form is the one with the least total costs for assumed expressions. Consider the simple sentence (5.3a) with the corresponding logical form (5.3b).

(5.3) a. The car is red.

b. $car(x) \wedge red(x)$

In the course of the selection of the best explanation, often the assumption of a particular predicate is preferred. The information structure of sentence (5.3) suggests that $car(x)$ is already known or given information, and $red(x)$ is the new information conveyed. To express this preference, the assumption costs could be assigned as follows: for $car(x)$ we assume a cost of \$10, and for $red(x)$ \$1. The costs of an abductive proof is the sum of the costs of used axioms in the knowledge base plus costs for assumed additional expressions. The preferred interpretation of (5.3) in which $red(x)$ is assumed has the least total costs of \$1, while alternative interpretations are not totally excluded but penalized by higher assumption costs.

The inference system of Stickel (1991) is implemented in Prolog. The logical form of a sentence is expressed by a conjunction of positive literals as shown in (5.4). Each conjunct Q_i is assigned an assumption cost c_i.

(5.4) $Q_1^{c_1} \wedge Q_2^{c_2} \wedge ... \wedge Q_n^{c_n}$

The knowledge base consists of axioms in form of Horn clauses (5.5) in which the literals P_i are assigned weights w_i.

(5.5) $P_1^{w_1} \wedge P_2^{w_2} \wedge ... \wedge P_n^{w_n} \supset Q$

By assigning costs to the axioms themselves, the relevance of a rule or a fact in a particular utterance situation can be expressed. In this way, axiom costs can reflect the prominence of particular facts. Entities already mentioned in the preceding discourse can be assigned lower costs.

Literals can be marked as proven, assumed, or unproven. In the original formula (5.4), all literals are either unproven or assumed. Unproven literals can be either assumed directly or must be deduced from the knowledge base, possibly by means of assumptions made during the proof. Once proven or assumed literals stay marked in subsequent proof steps. The formula resulting at the end of a proof must consist only of proven or assumed literals. Abductive proofs are produced by four inference rules:

Resolution with a fact If the formula to be proven contains a literal which is unifiable with a fact in the knowledge base, it will be marked as proven. Axiom costs for the fact are added to the total costs.

Resolution with a rule If a literal Q_i in the formula to be proven and a consequence Q of an axiom (5.5) are unifiable with the most general unificator σ, then the formula (5.6) can be inferred.

(5.6) $Q_1^{c_1}\sigma \wedge ... \wedge Q_{i-1}^{c_{i-1}}\sigma \wedge P_1^{cw_1}\sigma \wedge ... \wedge P_m^{cw_m}\sigma \wedge Q_{i+1}^{c_{i+1}}\sigma \wedge ... \wedge Q_n^{c_n}\sigma$

If c is the price of assuming Q, then cw_i is the cost of assuming P_i. The cost of the axiom (5.5) is added to the total costs.

Making assumptions Each unproven literal of a goal expression can be marked as assumed.

Factorization If a literal Q occurs more than once in a proof, every time with different costs, the occurrences of Q can be unified. For the resulting expression, the least costs of the original expressions are assumed.

If in an axiom $P_1^{w_1} \wedge P_2^{w_2} \supset Q$ the relation $w_1 + w_2 < 1$ holds, then it is a case of *most specific abduction*. Against the assumption of Q speaks the fact that it is cheaper to assume P_1 and P_2. In the case of $w_1 + w_2 > 1$ the less specific explanation is preferred. There is no reason to assume P_1 and P_2 if it is cheaper to assume Q. For instance, in the inference schema (5.7), assuming $Q_1^{10} \wedge Q_2^{10}$ has a price of \$20, whereas assuming $P_1^6 \wedge P_2^6 \wedge P_3^6$ has a cost of \$18 provided the two occurrences of predicate P_2 in the two axioms are unified. In general, the mechanism prefers *least specific abduction*, although redundancies can be used for obtaining more specific interpretations.

$$(5.7)\quad \frac{\begin{array}{l} P_1^6 \wedge P_2^6 \supset Q_1 \\ P_2^6 \wedge P_3^6 \supset Q_2 \end{array}}{Q_1^{10} \wedge Q_2^{10}}$$

An abductive proof is complete if all literals are either proven or assumed. The cost of a proof is the sum of all axiom costs plus the costs of the assumed literals.

5.1.3 Local Pragmatic Interpretation

The standard example cited by Hobbs et al. (1993) is the following:

(5.8) The Boston office called.

I will explain the mechanism of abductive interpretation by means of this example. It involves three pragmatic phenomena: the reference of the definite description *the Boston office* must be resolved, the metonymy must be recognized and expanded to *[Some person at] the Boston office called.*, and the implicit relation between *Boston* and *office* must be determined. The logical form of the sentence to be proven by abduction is (5.9).

(5.9) $\exists x, y, z, e[call'(e, x) \wedge person(x) \wedge rel(x, y) \wedge office(y) \wedge boston(z) \wedge nn(z, y)]$

The knowledge base – shared knowledge of speaker and hearer – consists of the following facts:

$$(5.10)\ \left\{ \begin{array}{lll} boston(B_1), & office(O_1) \wedge in(O_1, B_1), & \forall y, z[in(y, z) \supset nn(z, y)], \\ person(J_1), & work - for(J_1, O_1), & \forall x, y[work - for(x, y) \supset rel(x, y)] \end{array} \right\}$$

The first four propositions encode the domain knowledge that a person J_1 works for an office O_1 which is situated in Boston. The two universally quantified formulae express that *in* can be a possible implicit relation between two nouns in a compound noun, and that $work - for(x, y)$ is a relation between x and y.

The proof of (5.9) follows directly from these facts, except the conjunct $call'(e, x)$. Thus, it is simply assumed. It is the new information conveyed by the sentence. Deductive parts of the proof represent old information, abductive parts represent new information. The mentioned local pragmatic problems are solved as byproducts: *the Boston office* is resolved to O_1, the implicit relation in the compound noun is *in*, and the metonymy is expanded to *John, who works for the Boston office, called*. The abductive interpretation of this sentence is illustrated in Fig. 5.1. Note that boxed predicates cannot be proven from the knowledge base and thus stand for additional assumptions made. Arrows indicate deductive derivations, and the abductive interpretation process searches explanations in the opposite direction.

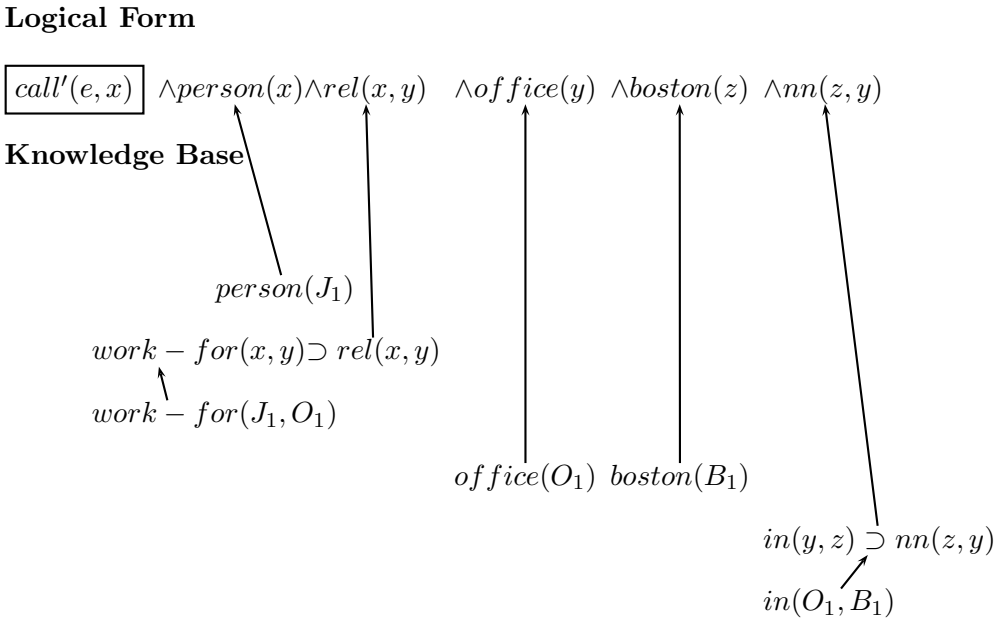


Figure 5.1: Abductive interpretation of (5.8) (Hobbs et al., 1993)

5.1.4 Abduction in Structured Discourses

The theory of Hobbs et al. (1993) is designed not only for interpreting single sentences but also for more complex discourses. As already discussed in chapter 4, interpreting a discourse involves the specification of implicit discourse relations between discourse segments. The abductive interpretation process can be used to infer these relations. In order to arrive at this goal, the knowledge base is extended by axioms on coherence relations.

The interpretation of a discourse u consists in the proof that u is a coherent discourse segment. In Hobbs' notation, this amounts to proving the proposition $\exists e[Segment(u, e)]$, stating that u is a coherent segment and e is the main eventuality of that segment. Unfortunately, Hobbs et al. (1993, pp. 104ff) do not make a distinction between the interpretation of a sentence (i.e. a discourse segment) and the eventuality expressed by it[1]. As a consequence, their coherence relations, although supposed to hold between discourse segments, actually hold between eventualities. One way to keep this distinction is to extend the proposition to be proven by an additional parameter for the logical form ϕ expressed by it. In the following, I will adapt Hobbs' formulae in this respect.

(5.11) $\exists \phi, e[Segment(u, \phi, e)]$

A sentence is a discourse segment, and the combination of two coherent segments via a coherence relation yields again a coherent segment, and a tree structure for the whole discourse emerges (cf. Hobbs, 1985a). This is captured by the following two axioms.

(5.12) $\forall u, \phi, e[s(u, \phi, e) \supset Segment(u, \phi, e)]$

(5.13) $\forall u_1, u_2, u, \phi_1, \phi_2, e_1, e_2[Segment(u_1, \phi_1, e_1) \wedge Segment(u_2, \phi_2, e_2)$
$\wedge CoherenceRel(u_1, u_2, u) \supset Segment(u, \phi_1 \wedge \phi_2, e_1 \uplus e_2)]$

Axiom (5.12) simply states that sentences are discourse segments[2], with $s(u, \phi, e)$ meaning that u is an interpretable sentence with the logical form ϕ and the main eventuality e. Here again, I try to keep the distinction between the logical form of a sentence, i.e. a proposition, and the eventuality expressed by it, i.e. an entity.

Axiom (5.13) is the underlying composition rule for subsequent discourse segments u_1 and u_2. The predicate $CoherenceRel(u_1, u_2, u)$ expresses that there is a coherence relation between u_1 and u_2 and their combination yields u. Note that in this approach, the logical form of a discourse consists in the logical conjunction of the logical forms of its constituent sentences. The operator $\uplus$ merges two events to a complex eventuality. The ontological status of this operator is, however, questionable, and a more sophisticated treatment of the relationship between segments and eventualities would be desirable. Anyway, later chapters of this thesis will deal with alternative approaches to event semantics. Let us illustrate the working of the abductive mechanism with an example from Hobbs et al. 1993:

(5.14) The police prohibited the women from demonstrating. They feared violence.

Let u be the discourse (5.14), then, according to (5.11) and (5.13), its interpretation is the proof that each uttered sentence is a segment and that a coherence relation holds between them. Since, according to (5.12), individual sentences are discourse segments, we must abductively infer their logical forms shown in (5.15).

(5.15) a. $\exists e_1, p, d, w[police(p) \wedge women(w) \wedge prohibit'(e_1, p, d) \wedge demonstrate'(d, w)]$

1 Actually, Hobbs et al. (1993) write first that e refers to the eventuality expressed by u, and later that e is the assertion or the topic of u.

2 Note that this notation presupposes Hobbs' abductive approach to syntax, which is outlined in Hobbs et al. (1993, ch. 6) and elaborated in Hobbs (2001).

b. $\exists e_2, X, v[fear'(e_2, X, v) \wedge violence(v)]$

A possible coherence relation is EXPLANATION, which in turn can be derived if there is a causal connection between the involved events. These two assumptions can be expressed by the following two axioms, which, unlike the axioms of Hobbs et al. (1993), do reflect the distinction between eventualities and propositions.

(5.16) a. $\forall u_1, u_2[Explanation(u_1, u_2) \supset CoherenceRel(u_1, u_2, u_1)]$

b. $\forall e_1, e_2, \phi_1, \phi_2, u_1, u_2[cause(e_2, e_1) \wedge Segment(u_1, \phi_1, e_1) \wedge Segment(u_2, \phi_2, e_2) \supset Explanation(u_1, u_2)]$

In order to make use of these axioms for the interpretation of example (5.14), the specific knowledge base must contain an axiom stating that an event e_1 of prohibiting can be caused by an event e_2 of fearing. An additional assumption that has to be made in order

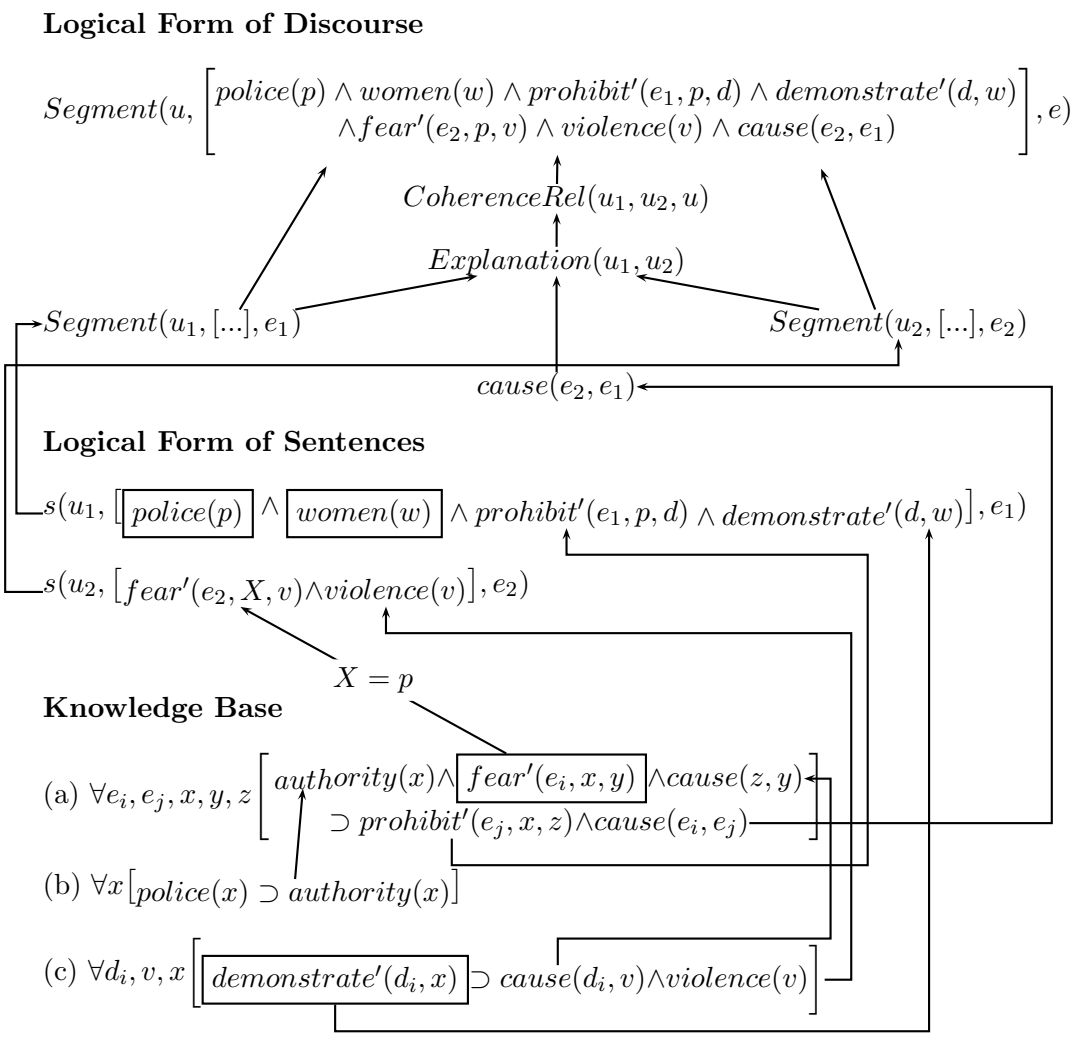


Figure 5.2: Abductive interpretation of (5.14)

to arrive at this causal connection is that the pronoun "they" is coreferential with "the police". Fig. 5.2 shows the interpretation of this example using abductive reasoning on the basis of a specific knowledge base. As before, arrows indicate deductive proofs of abductive inferences, and not provable assumptions are boxed.

Starting from the bottom of the picture, the deductive justification of the abductive inference process follows the arrows. The axioms in the knowledge base can informally be described as follows: (a) if an authority x fears some y that is caused by a circumstance z, then the authority prohibits z, (b) the police is an authority, and (c) demonstrations cause violence. Not that it is necessary to assume (c) as an axiom in order to infer an EXPLANATION relation such that, in course of the interpretation process, the underspecified parameter X standing for the pronoun "they" can be unified with the referent p representing the police. However, assuming a different set of axioms, the pronoun could refer to the demonstrators, as well. In that case, a CONSEQUENCE relation could be inferred, with an inverted causal connection.

Once the logical forms of the individual sentences are proven, the proof of the logical form for the discourse is straightforward. According to (5.12), both u_1 and u_2 are discourse segments. A causal connection between the involved events follows from the knowledge base and enables the assumption of an EXPLANATION relation between the segments according to (5.16). This relation, in turn, makes the discourse u coherent (5.13), bearing a logical form that consists of the conjunction of the logical forms of the individual sentences.

5.1.5 Conclusion

The abductive approach to discourse interpretation is intuitively very appealing. It seems plausible that abduction is the human way of reasoning in general. In this way, a unique inference mechanism for the whole interpretation process is assumed.

The work of Hobbs et al. (1993) has had a considerable influence on research in formal semantics and pragmatics in the last decade. In a multi-layer theory of utterance interpretation (Bierwisch, 1983; Dölling, 2005, see also chapter 1), the level of semantic form contains underspecified parameters, which are resolved in the course of contextual interpretation. In example (5.8) above, the predicates nn and rel in the logical form can be regarded as underspecified parameters which are specified as in and $work-for$ during the abductive interpretation process. Dölling (1997) shows in detail how abduction can be employed in the fixation of underspecified parameters for systematic meaning variations within a multi-layer theory of meaning. Maienborn (2001) shows how underspecifications introduced by locative modifiers can be interpreted in such a framework.

However, the abductive discourse interpretation presented by Hobbs et al. (1993) has some fundamental problems. For instance, as Asher and Lascarides (2003) point out, the coherence relation holding between two discourse segments can change depending on the context. Even though discourse (5.17a) features an EXPLANATION relation in a neutral context, a NARRATION relation can hold in a more specific context (5.17b).

(5.17) a. Max fell. John pushed him.

b. John hit Max on the back of his neck. Max fell. John pushed him. Max rolled over the edge of the cliff.

Discourse structures are built as part of the discourse interpretation process, but they are not used to dynamically predict which coherence relations are to be assumed in subsequent segments. Similarly, preferences for the resolution of anaphora and for the specification of other underspecified parameters can change according to the context. This cannot be adequately modelled in Hobbs' original account, since weights on axioms in Stickel (1991)'s weighted abduction are defined as static and context-independent. Stone and Thomason (2002) present an abductive approach with a differentiated representation of context. Weights are dynamically assigned to axioms in the knowledge base in form of functions of contexts.

As noted in the previous section, a problem regarding the abductive interpretation of complex discourses according to Hobbs et al. (1993) is the missing distinction between (the informational content of) discourse segments and eventualities expressed by them. The interpretation of a discourse segment (i.e. a sentence) – which is traditionally regarded to be a proposition – is identified with its main eventuality – which is an entity.

Another problem of weighted abduction is that the weights are responsible for deciding what knowledge sources are to be preferred over others in order to arrive at intuitively preferred readings. This problem was hinted at by Asher and Lascarides (2003) in connection with examples involving *bridging* such as (5.18).

(5.18) a. John moved from Luton to central London.

b. The rent was cheaper.

Here, the intuitively preferred reading is the one in which the rent in (b) refers to the rent in London, even though it violates general world knowledge about current rent rates in England. If we additionally assume that the cheaper rent caused John to move to central London, an EXPLANATION relation can be inferred.

On the other hand, in accordance with world knowledge, the most plausible reading for discourse (5.18) is the one which regards the rent in (b) as the rent in Luton. In this case, a discourse relation BACKGROUND can be inferred. This interpretation should be cheaper in the weighted abduction schema than the former because no assumption is necessary that involves a violation of world knowledge.

In order to make sure that the intuitively preferred reading is actually assigned cheaper derivation costs, additional rules would have to be added to the logic. It seems very difficult to model the interaction of world knowledge and discourse structural preferences in terms of numerical values. A solution might be to apply techniques of Machine Learning to large knowledge corpora, which, however, still have to be constructed.

5.2 Minimal Model Generation

A paradigm that was payed increasing attention to in the recent literature on computational semantics is the technique of Model Generation (cf. *inter alia* Konrad, 2000; Gardent and Konrad, 2000; Kohlhase, 2000; Blackburn et al., 1999). Approaches to discourse interpretation using this method are based on some notion of *minimality* that could be described as follows: an ambiguous contribution in a discourse should be interpreted in a way that a reading in which one continues to speak about the same entities and concepts is favoured over a reading in which one has to assume new entities.

In this section, I will first introduce and elaborate on the notions of *Model* and *Herbrand model*. Then, I will sketch how Model Generation can be used in discourse interpretation before discussing different conceptions of minimal models.

5.2.1 Herbrand Models for First-Order Languages

In general, a model is a mathematical structure that describes how the expressions of a language are interpreted. A model $\mathcal{M} = \langle \mathcal{D}, \mathcal{I} \rangle$ consists of a domain $\mathcal{D}$ and an interpretation function $\mathcal{I}$. The domain $\mathcal{D}$ is a set of individuals. It is also called *discourse universe*. The interpretation function $\mathcal{I}$ maps constant symbols to elements of $\mathcal{D}$ and, similarly, relation symbols to sets of n-tuples of elements in $\mathcal{D}$. It can be seen as a set of *assertions* about individuals in $\mathcal{D}$.

For illustration, consider a simple example. Be $\mathcal{L}$ a first-order language. Its vocabulary includes constants $mary, john, jack$, one-place predicates $man, woman$, and two-place predicates $love$. Imagine a situation in which Mary loves both John and Jack, and John loves Mary. This situation can be described by the following model:

$$
\begin{array}{ll}
\mathcal{D} = \{d_1, d_2, d_3\} & \\
\mathcal{I}(mary) = d_1 & \mathcal{I}(man) = \{d_2, d_3\} \\
\mathcal{I}(john) = d_2 & \mathcal{I}(woman) = \{d_1\} \\
\mathcal{I}(jack) = d_3 & \mathcal{I}(love) = \{(d_1, d_2), (d_1, d_3), (d_2, d_1)\}
\end{array} \tag{5.19}
$$

A *Herbrand Model* (cf. Konrad, 2000; Kohlhase, 2000) is a special class of models in which constants and other terms are interpreted as themselves, i.e. $\mathcal{I}(t) = t$ for all ground terms t. The domain $\mathcal{D}$ of $\mathcal{M}$ consists of the terms that occur in the set of all ground terms that are satisfied by $\mathcal{M}$. In a Herbrand model, only the interpretation of predicate symbols must be specified. For example, the situation above can be expressed by the following Herbrand model.

$$
\begin{array}{ll}
\mathcal{D} = \{mary, john, jack\} & \\
\mathcal{I}(mary) = mary & \mathcal{I}(man) = \{john, jack\} \\
\mathcal{I}(john) = john & \mathcal{I}(woman) = \{mary\} \\
\mathcal{I}(jack) = jack & \\
\mathcal{I}(love) = \{(mary, john), (mary, jack), (john, mary)\} &
\end{array} \tag{5.20}
$$

As Baumgartner and Kühn (2000) point out, Herbrand models are particularly well suited for discourse representation. In Herbrand interpretations, different terms denote different objects, a property also known as *unique name assumption*. Moreover, in a domain only those entities exist that are talked about. This property is known as *closed world assumption*. This is a desired property for anaphora resolution, but in order to interpret indefinite descriptions and deictic references, the introduction of new elements into the discourse model must be allowed. We will see later how this can be done. First, some formal definitions are needed to define the notion of a Herbrand model. Some specific details are left out for the sake of a clear exposition; a more comprehensive formal introduction is provided by Konrad (2000).

Literal A *literal* is either an atomic formula $P(t_1, ...t_n)$ or its negation $\neg P(t_1, ...t_n)$.

Ground Term A literal or clause is said to be *ground* if it does not contain variables.

Herbrand Universe Let $\mathcal{L}$ be a first order language. The set $\mathcal{H}_\mathcal{L}$ of all ground terms constructed from functors and constants in $\mathcal{L}$ is called the *Herbrand Universe* of $\mathcal{L}$.

Herbrand Model Let $\mathcal{M} = \langle \mathcal{D}, \mathcal{I} \rangle$ be a first order model. $\mathcal{M}$ is a *Herbrand Model* iff $\mathcal{D} \subseteq \mathcal{H}_\mathcal{L}$ and $\mathcal{I}(t) = t \in \mathcal{D}$ for all ground terms $t \in \mathcal{H}_\mathcal{L}$.

Herbrand Base The set $\mathcal{B}_\mathcal{M}$ of all ground literals that are satisfied by $\mathcal{M}$ is called the *Herbrand Base* of $\mathcal{M}$.

A Herbrand model $\mathcal{M}$ can be uniquely represented by its Herbrand base $\mathcal{B}_\mathcal{M}$. Accordingly, a model for our example would be $\mathcal{M} = \{woman(mary), man(john), man(jack), love(mary, john), love(mary, jack), love(john, mary)\}$.

Herbrand Theorem If a first order theory is satisfiable at all, then it must have at least one Herbrand model.

This theorem is very useful regarding consistence of models. In order to show that a theory is inconsistent, it suffices to show that it cannot have a Herbrand model.

5.2.2 Generation of Discourse Models

In the following sections, I will use the notions *model* and *discourse model* (cf. section 3.2) in the rather strict sense of a logical model for a set of propositions representing the discourse to be interpreted. A Herbrand model can be seen as a set of basic assumptions under which a given logical specification in form of a set of propositions is true. In particular, for a logical specification Φ, a model generator proves that Φ is satisfiable by generating some of its models. Model generation computes all assumptions that are necessary for a model to satisfy the specification and which are consistent with the context. Satisfying models explain how a logical specification representing a discourse can be made true; in this sense they represent the meaning of a discourse. Note that model generation can be seen as a form of abductive reasoning (cf. section 1.2.4): abduction is the inference from an observation to an explanation for this observation. Models offer an *explanation* for the truth of a discourse, the *observation*. From Φ and $\mathcal{M} \models \Phi$, the model $\mathcal{M}$ is derived.

To illustrate the close relationship between model generation and abduction, let us reconsider example (5.8) from section 5.1, repeated here together with its semantic representation (5.22):

(5.21) The Boston office called.

(5.22) $\exists x, y, z, e[call'(e, x) \wedge person(x) \wedge rel(x, y) \wedge office(y) \wedge boston(z) \wedge nn(z, y)]$

Available background knowledge includes the following information.

(5.23) $\left\{ \begin{array}{lll} boston(b), & office(o) \wedge in(o, b), & \forall y, z[in(y, z) \supset nn(z, y)], \\ person(j), & work - for(j, o), & \forall x, y[work - for(x, y) \supset rel(x, y)] \end{array} \right\}$

A model $\mathcal{M}$ satisfying the specification Φ, which consists in the union of (5.22) and (5.23), would be the following:

(5.24) $\mathcal{D} = \{j, b, o\}$
$\mathcal{I}(boston) = \{b\}$, $\mathcal{I}(office) = \{o\}$, $\mathcal{I}(person) = \{j\}$,
$\mathcal{I}(in) = \{(o, b)\}$, $\mathcal{I}(work - for) = \{(j, o)\}$, $\mathcal{I}(nn) = \{(b, o)\}$,
$\mathcal{I}(rel) = \{(j, o)\}$, $\mathcal{I}(call) = \{j\}$

Thus, model generation and abduction are indeed very similar. A difference (pointed out by Blackburn and Bos, 2003) consists in its computational properties. In weighted abduction, once found proofs are stored in the knowledge base and cannot be reconsidered. In model generation, alternative proofs are available unless they are inconsistent. The question remains, however, how to obtain a model, and which models are suitable as a discourse representation.

Suitable models should have the following properties: (i) the entailed assumptions are either evident or accommodated as required, (ii) accommodated assumptions should be consistent with what we know about the situation, (iii) the assumptions should explain completely why a certain sentence can be true in a given context, and (iv) the model should not state more information than required. For illustration, look at the following example (adapted from Gardent and Webber, 2001). The sentence in (5.25a) can be given a simple semantic representation (5.25b).

(5.25) a. John greeted his friend.

b. $\exists x[friend(x) \wedge\ of(x, john) \wedge\ greet(john, x)]$

I will explain the way models are constructed in section 5.2.4. For now, I will illustrate what it means for a model to satisfy a given logical specification, i.e. a semantic representation. Consider the following models:

- $\mathcal{M}_1 = \{friend(c_1), of(c_1, john), greet(john, c_1)\}$
- $\mathcal{M}_2 = \{friend(john), of(john, john), greet(john, john)\}$
- $\mathcal{M}_3 = \{friend(c_1), of(c_1, john), greet(john, c_1), married(john)\}$

As expected, $\mathcal{M}_1$ is a Herbrand model satisfying (5.25b) and a valid interpretation of (5.25a). Its domain is $\mathcal{D} = \{john, c_1\}$. However, it is not the only possible model. Also $\mathcal{M}_2$ is a Herbrand model satisfying (5.25b), but it does not capture the meaning of the sentence. Similarly, more complex models containing more assertions can be satisfying as well. In this way, $\mathcal{M}_3$ is a Herbrand model satisfying (5.25b), but it is not appropriate because the assertion $married(john)$ is not necessary to account for the meaning of (5.25a). Thus, we can observe that some models are "too small" while others are "too big" to capture the preferred meaning of a discourse.

5.2.3 Minimality of Models

To eliminate models that are too big, we can look for models that are in some sense *minimal*. Various notions of minimality have been proposed in the literature. The following definitions are adapted from Gardent and Webber (2001) and Konrad (2000).

If we look at the basic structure of a model, there are two straightforward options to reach minimality: we minimize either $\mathcal{D}$ or $\mathcal{I}$. On the one hand, we can select models with the smallest discourse universe.

Domain Minimality Let Φ be a set of first order formulae, and S be the set of Herbrand models of Φ. Then a model $\langle \mathcal{D}, \mathcal{I} \rangle \in S$ is *domain minimal* iff there is no other model $\langle \mathcal{D}', \mathcal{I}' \rangle \in S$ such that the cardinality of $\mathcal{D}'$ is smaller than the cardinality of $\mathcal{D}$, i.e. $|\mathcal{D}'| < |\mathcal{D}|$.

On the other hand, we can select models with the smallest number of facts, or assertions, or instances of predicates.

Subset Minimality Let Φ be a set of first order formulae and S be the set of Herbrand models of Φ. Then a model $\langle \mathcal{D}, \mathcal{I} \rangle \in S$ is *subset minimal* iff there is no other model $\langle \mathcal{D}', \mathcal{I}' \rangle \in S$ such that $\mathcal{I}' \subseteq \mathcal{I}$.

Another approach to minimality also seeks to minimize the extension of the interpretation function, but, unlike subset minimality, it selects models with the smallest number not of all predicates, but of instances of a particular predicate.

Predicate-Specific Minimality Let p be a certain predicate symbol, and $p(\mathcal{M})$ be the subset of atoms in the Herbrand base of $\mathcal{M}$ whose head predicate symbol is p. A model $\mathcal{M}$ is *p-minimal* if there is no model $\mathcal{M}'$ such that $|p(\mathcal{M}')| < |p(\mathcal{M})|$.

This form of minimality is used in the formalism of Circumscription (see section 1.2.2). There, models are chosen which minimize the occurrences of an abnormality predicate ab. An assertion like "Birds normally fly" is represented as $\forall x[bird(x) \wedge \neg ab(x) \rightarrow flies(x)]$. For instance, if Tweety is a bird, and as long as $ab(Tweety)$ cannot be proven, it can be assumed that Tweety flies.

Finally, we can combine domain and subset minimality by selecting models that have a minimal domain $\mathcal{D}$ *and* that are subset minimal with respect to all other domain minimal models.

Local Minimality Let Φ be a set of first order formulae and S be the set of Herbrand models of Φ that use some finite domain $\mathcal{D}$ whose size is minimal. Then a model $\langle \mathcal{D}, \mathcal{I} \rangle \in S$ is *locally minimal* iff there is no other model $\langle \mathcal{D}', \mathcal{I}' \rangle \in S$ such that $\mathcal{I}' \subseteq \mathcal{I}$.

Locally minimal models satisfy a specification without referring to more individuals than needed and without making unnecessary assumptions with respect to all other models that can be found in the smallest domain. These models are the simplest models in the sense of Occam's Razor[3]. Locally minimal models minimize accommodated individuals and assumptions. Looking again at the models of example (5.25), model $\mathcal{M}_1$ is locally minimal, and model $\mathcal{M}_3$ is correctly ruled out. Still, we cannot rule out models that are too small: model $\mathcal{M}_2$ is, like $\mathcal{M}_1$, locally minimal. But it is dispreferred because it involves a reflexive interpretation of "to greet". Additional information from general world knowledge is necessary for discarding models like this one. In this specific case, if the fact that one cannot greet oneself were integrated in the specification, model $\mathcal{M}_2$ would be ruled out (or not even generated) because of inconsistency.

To summarize so far: a constraint of minimality, specifically local minimality, can rule out some models corresponding to dispreferred readings. However, this constraint is not always sufficient; to rule out other non-suited models, additional knowledge from different sources such as contextual and world knowledge must be taken into account.

3 Occam's Razor is the principle of parsimony in science, attributed to William of Ockham (1285–1349), stating that the explanation of a phenomenon should involve as few assumptions as possible. It is expressed in Latin as "entia non sunt multiplicanda praeter necessitatem" ("entities must not be multiplied beyond necessity").

5.2.4 Minimal Models and Discourse Anaphora

In this section, I will explain in some detail how minimal model generation can be used to resolve discourse anaphora. Section 5.2.4.1 exemplifies the resolution of direct anaphora involving pronouns. A simple model for a very short discourse will be constructed in detail. Section 5.2.4.2 then shows how the mechanism can be extended in order to overcome some problems with finding minimal models for pronouns.

5.2.4.1 Resolving Pronouns by Model Generation

(5.26) A baby is sleeping. It smiles.

The small text in (5.26) has two possible interpretations. In the preferred reading, "it" refers to the baby in the first sentence. But unless additional information is available, there is always an alternative deictic reading for pronouns. In the deictic reading, "it" does not have a linguistic antecedent and must be accommodated.

In the following, I will examine how a discourse model for this short discourse is constructed. Technical realizations of model construction make use of one or another form of a tableau calculus. A variety of these calculi have been developed and extended in the fields of Artificial Intelligence and Automated Reasoning, not all of them equally suited for model construction. I will present a simplified static calculus for first order logic in order to illustrate the basic ideas. More sophisticated methods are discussed in greater detail in Konrad (2000) and a dynamic calculus for constructing models for DRT (Kamp and Reyle, 1993) is presented in Kohlhase (2000).

A tableau $\mathcal{T}$ is a set of expansion rules (see Fig. 5.3). I will use the following notation convention: ϕ^T indicates that ϕ is true (a positive literal), and ϕ^F means that ϕ is a negative literal, i.e. that $\neg\phi$ holds. Using the equivalencies of $\exists x[\phi]$ with $\neg\forall x[\neg\phi]$ and of $\phi \vee \psi$ with $\neg(\neg\phi \wedge \neg\psi)$ we only need rules involving $\neg$, $\forall$ and $\wedge$. By recursively applying expansion rules to a given proposition, a tableau proof in form of a tree structure is obtained. Each branch in such a tableau proof corresponds to a particular model. A branch (or the set of propositions it contains) is saturated if no application of an expansion rule can add new elements to this set.

$$\frac{\neg\phi^T}{\phi^F}\mathcal{T}(\neg) \qquad \frac{\neg\phi^F}{\phi^T}\mathcal{T}(\neg) \qquad \frac{\begin{array}{c}\phi^T\\ \phi^F\end{array}}{\bot}\mathcal{T}(\bot)$$

$$\frac{(\phi \wedge \psi)^T}{\begin{array}{l}\phi^T\\ \psi^T\end{array}}\mathcal{T}(\wedge) \qquad \frac{(\phi \wedge \psi)^F}{\phi^F \mid \psi^F}\mathcal{T}(\vee) \qquad \frac{(\forall x[\phi])^T}{[a_1/x]\phi^T}\mathcal{T}(\forall)$$

$$\frac{(\forall x[\phi])^F}{[a_1/x]\phi^F \mid \ldots \mid [a_n/x]\phi^F \mid [c_{new}/x]\phi^F}\mathcal{T}(\exists)$$

Figure 5.3: Tableau calculus for model generation (adapted from Kohlhase, 2000)

The rule $\mathcal{T}(\forall)$ is applied exhaustively to all elements of the Herbrand universe of the current branch $(a_1, ..., a_n)$. The rule $\mathcal{T}(\exists)$ either makes use of constants occurring in the

current branch or introduces a new constant c_{new}. Note that if the Herbrand universe is extended by a new constant, all applications of $\mathcal{T}(\forall)$ must be re-instantiated with respect to this constant. To illustrate the calculus, let us assume the following simplified semantic representation of the linguistic content of (5.26).

(5.27) $\exists u[baby(u) \wedge sleep(u)] \wedge \exists v[smile(v)]$

Model construction for (5.27) yields the the following tableau proof:

$$
\begin{array}{c}
\exists u[baby(u) \wedge sleep(u)]^T \\
{[c_1/u]} \\
baby(c_1)^T \\
sleep(c_1)^T \\
\exists v[smile(v)]^T \\
\begin{array}{c|c} {[c_1/v]} & {[c_2/v]} \\ smile(c_1)^T & smile(c_2)^T \end{array}
\end{array}
\tag{5.28}
$$

Applying rule $\mathcal{T}(\exists)$ to the representation of the first sentence replaces u by the new constant c_1. By rule $\mathcal{T}(\wedge)$ we get $baby(c_1)$ and $sleep(c_1)$. In the second sentence, rule $\mathcal{T}(\exists)$ is applicable again, but it creates now two branches, yielding two possible models:

- $\mathcal{M}_1 = \{baby(c_1), sleep(c_1), smile(c_1)\}$
- $\mathcal{M}_2 = \{baby(c_1), sleep(c_1), smile(c_2)\}$.

These models correspond to the two readings of the anaphoric pronoun "it" in the second sentence. In the first model, the constant c_1 corresponding to the referent of the pronoun is equal to the constant corresponding to the referent of the baby in the first sentence c_1. In the second model, two distinct constants c_1 and c_2 are assumed.

This is the point where minimality of models comes into play. To account for the empirical preference for anaphoric binding over a deictic reading, local minimal models are preferred. For a deictic interpretation, more new constants and hence distinct discourse referents are needed. The corresponding model $\mathcal{M}_2$ is not locally minimal. Thus, we correctly remain with the intuitively preferred reading.

As already emphasized at various points in this work, the resolution of anaphora is a fundamentally nonmonotonic process: preferences for anaphora resolution can change as the discourse progresses. New information showing up later can override default preferences and make a previously disregarded reading preferent. To give another example, look at (5.29). After the second sentence is uttered, the preferred interpretation of the pronoun "he" would be one in which it is coreferential with the Vice-President. This preference changes, however, after the utterance of the last sentence. At that point it becomes clear that "he" in both (b) and (c) is coreferential with the President.

(5.29) a. The Vice-President entered the President's office.

b. He was nervous and clutching his briefcase.

c. After all, he couldn't fire the Vice-President without making trouble for himself with the chairman of the board.

A similar case is example (5.30), which I will use to illustrate the problem in more detail.

(5.30) I had gone to see John before I visited Bill and Mary. He doesn't want to speak with her. Cohen (2006)

The pronoun "her" in the second sentence refers unambiguously (disregarding a deictic reading) to Mary. "He" is ambiguous between John and Bill, although most readers would choose John as the preferred antecedent. Now consider the following continuation, where another reading is preferred: "he" in the second sentence refers to Bill.

(5.31) I had gone to see John before I visited Bill and Mary. He doesn't want to speak with her. John knows everything about their relationship.

Minimality of models alone is not sufficient for an antecedent to be chosen. A model with $\mathcal{D} = \{i, j, m, b\}$ and $\mathcal{I}(want\_to\_speak\_with) = \varnothing$ satisfies the discourse (5.30) and is, trivially, minimal. In the following section, we will see a solution to this problem.

5.2.4.2 Equality by Default

In DRT, resolving ambiguities means establishing an equivalency relationship between discourse referents. The meaning of the sentences in (5.30) can be represented as shown in (5.32). Note that the DRS for the second sentence is underspecified: the parameters u and v, which stand for the referents of the pronouns, have to be specified in course of the interpretation. This representation has the advantage that it expresses the ambiguity caused by "he": the parameter u can be specified as coreferential with j (for John) or with b (for Bill), according to the desired reading.

(5.32)

i, j, b, m
$I(i), john(j), bill(b), mary(m)$ $go\_to\_see(i, j), visit(i, b), visit(i, m)$

u, v
$male(u), female(v)$ $\neg want\_to\_speak\_with(u, v)$ $u =?, v =?$

In order to resolve this ambiguity, we can resort to an additional assumption proposed by Cohen (2006): *Equality by Default*.

Equality by Default Two discourse referents are assumed to be equal, unless it can be proven that they are distinct.

This idea can be represented as a simple default rule (5.33), here formulated as a Default Logic rule.

(5.33) $$\frac{: x = y}{x = y}$$

A default theory for (5.32) has two extensions (provided that $j \neq b$ holds): one contains $u = j$, and the other one contains $u = b$. The equality $v = m$ holds in all extensions.

Now, depending on the kind of default reasoning we perform, we can account for the different readings of (5.30). On the one hand, assuming *Credulous Reasoning*[4], if one antecedent is preferred over the other, because of salience or plausibility, one extension can be chosen and the referent of the pronoun is assumed to be equal to the respective antecedent.

On the other hand, assuming *Skeptical Reasoning*, if there is no preference for a particular reading, we accept only that which holds in all extensions, and the anaphor remains ambiguous. In this case, a deictic reading is possible.

It is important to note that the default rule (5.33) is a default of low priority or a "last resort" principle. In examples (5.34), the application of this rule must be blocked by higher ranked defaults.

(5.34) a. John saw Bill. He greeted him.

b. John hates him.

c. John doesn't have a car. It is red.

d. A man came into the ice cream parlour. She was upset.

In example (5.34a), the rule can be blocked by another default rule with a higher priority: prefer an antecedent in the same syntactic position as the anaphor. This rule applies where a PARALLEL discourse relation holds between the discourse segments involved. I will come back to this point in section 5.3.2.4 below.

A coreferential reading of "John" and "him" in (5.34b) is ruled out by syntactic binding constraints. The pronoun "himself" would have to be used instead of "he" for the reflexive reading to hold (cf. section 3.1.1 above).

Assuming the referents of "a car" and "it" in (5.34c) to be equal is not possible because of the accessibility constraint of DRT (cf. section 3.4.3.4).

Coreference between the referents of "a man" and "she" in (5.34d) is blocked by a violation of gender agreement. Here, a coreferential reading would have to be strongly indicated by specific knowledge about the person in question.

However, rule (5.33) is expressed in a very general fashion. It is not restricted to referents of anaphoric expressions. How do we avoid that everything is assumed to be equal to anything? Let us look at some potentially problematic cases.

(5.35) John talked to Bill.

If "John" and "Bill" denote distinct individuals, their discourse referents cannot be equal. This insight is captured by the *Unique Names Assumption*: distinct names denote distinct individuals. However, Equality by Default seems to be rather contrary to the Unique Names Assumption. A remedy is the assumption of a higher ranked default, formalized in (5.36):

(5.36) $$\frac{named(x, P) : \neg named(x, Q)}{\neg named(x, Q)}$$

4 cf. section 1.2.1 above

With that, we can conclude by default: Bill is not named John. The property of being named John is true for John and false for Bill. Hence, John and Bill are distinct individuals.

(5.37) John is meeting a woman tonight. His mother told me so.

Here, we have to avoid that "a woman" is interpreted as being equal to John's mother. This can be accounted for by the following conversational implicature: if the speaker had intended to convey that John meets his mother, then she should have said so. The fact that she did not say so is reason to assume that "a woman" does not refer to John's mother.

(5.38) John went to the clinic. The doctor has a busy day.

In this example, there are two readings, which lead to two extensions of the corresponding default theory: (i) John is the doctor, and (ii) John is a patient of the doctor that is implicitly evoked by "the clinic". We will see in chapter 7 how to account for these kinds of examples.

(5.39) An officer talked to a gentleman.

As widely assumed in dynamic semantics, indefinite noun phrases introduce new discourse referents into the discourse model (cf. e.g., Heim, 1982). New referents are, by definition, not equal to already introduced discourse referents. However, this assumption does not work in all cases. Danlos (2001) mentions an example showing the opposite:

(5.40) Fred damaged a garment. He stained a shirt.

The standard representation of the indefinite noun phrases "a garment" and "a shirt" is as follows: new discourse referents x and y are introduced, together with DRS-conditions $garment(x)$ and $shirt(y)$, respectively. Danlos argues that the second utterance is a PARTICULARIZATION of the first one. The discourse relation PARTICULARIZATION is a special case of ELABORATION which implies coreference of events and, with that, coreference of arguments in the same thematic roles. In this way, a coreference relation $x = y$ can be established.

Equality by Default allows us to establish coreference relations in cases where it is difficult to come at otherwise. It is a means to minimize the domain of models. However, remember that in order to avoid overgeneralizations, rule (5.33) should only be applied if nothing contradicts it.

As sketched so far, the resolution of pronouns can accounted for in the presented framework of model generation with Equality by Default. The approach will be applied to bridging anaphora in section 6.3.2.1.

5.2.5 Conclusion

In a nutshell, model generation can perform the following tasks in natural language interpretation. From a given logical specification Φ, models are constructed which satisfy the specification Φ. In case that Φ is an underspecified semantic representation of a discourse, each model represents a possible pragmatic enrichment. Thus, model generation offers a technique of incrementally constructing discourse models by taking into account (some

part of) contextual and world knowledge. Ambiguities can be represented easily by the existence of various minimal models. The principle of minimality helps to select preferred readings. Typically, suitable models are minimal, i.e. they contain no more individuals and assertions than necessary. A way of ensuring domain minimality is given by assuming Equality by Default.

However, model generation for discourse interpretation as presented here does not account for the interpretation of complex discourses involving rhetorical relations.

5.3 Segmented Discourse Representation Theory

A theory that uses the concept of discourse relations to build complex structures for coherent discourses is Segmented Discourse Representation Theory (SDRT), developed by Nicholas Asher and Alex Lascarides. In difference to abductive interpretation, SDRT does not assume a unique logic but distinct logics for each knowledge source that contributes to discourse interpretation (compositional and lexical semantics, domain knowledge, cognitive states etc.). To bring them together, SDRT basically differentiates between two logics.

Firstly, the *logic of information content* is a device for *representing* the logical forms of discourse, with a dynamic semantic interpretation (à la Kamp, 1981, à la Heim, 1982, or à la Groenendijk and Stokhof, 1991), extended by including semantics for rhetorical relations.

Secondly, the *glue logic* is a device for *constructing* logical forms. It involves commonsense reasoning with both linguistic and non-linguistic information. The glue logic has only restricted access to this information. So it has access only to *descriptions* of formulae in the logic of information content, not to the logic itself. The glue logic together with discourse update and revision constitutes the *logic of information packaging*.

In the following, a short synopsis of the logics used in SDRT is given, adapted mainly from Lascarides and Asher (2007). Special emphasis lies on what part of the interpretation process each logic represents, and what inferences it performs.

5.3.1 Representing Discourse Structures

Discourse structures in SDRT are represented in the language of information content. It consists of the language of DRT (or any other dynamic semantic theory)[5], extended by two new expressions: (1) speech act discourse referents: they label content of text segments and keep track of token utterances, and (2) rhetorical relations: they relate speech act discourse referents.

An SDRS-formula represents the content of an utterance or a rhetorical relation between labels. The content of an utterance is represented as a DRS. The set of well-formed SDRS-formulae is defined as follows:

SDRS-formulae

1. if K is a DRS, then K is an SDRS-formula.
2. if R is rhetorical relation, and π_1 and π_2 are labels, then $R(\pi_1, \pi_2)$ is an SDRS-formula.

5 A short introduction to DRT was given in section 3.4.3.

The resulting structures are segmented discourse representation structures (SDRSs).

SDRS A discourse structure, or SDRS, is a triple $\langle A, \mathcal{F}, LAST \rangle$ with

- a set A of labels or speech act discourse referents,
- a label $LAST \in A$ corresponding to the content of the last clause that was added to the discourse structure,
- a function $\mathcal{F}$ which maps each label A to an SDRS-formula.

The labels in an SDRS can be used to form a hierarchical structure for a discourse. If π_0 is a label for a rhetorical relation $R(\pi_1, \pi_2)$, then π_0 *outscopes* both π_1 and π_2.

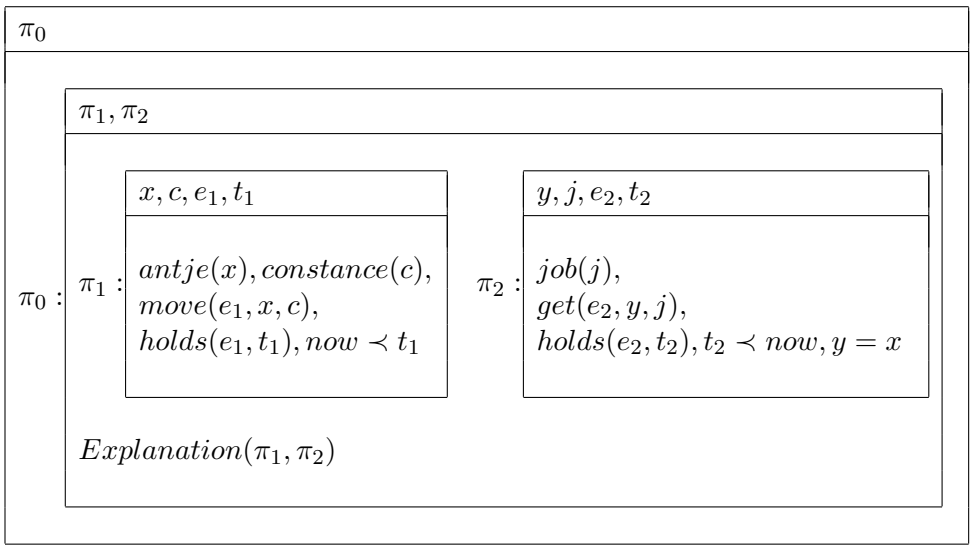
For a simple example, consider (5.41). Its discourse structure is given in (5.42). An SDRS can be represented in the box notation known from DRT, given in (5.43). By notational convention, I write $\pi : \phi$ for $\mathcal{F}(\pi) = \phi$ and $\pi : \boxed{\phi}$ for $\phi \in \mathcal{F}(\pi)$.

(5.41) a. Antje will move to Constance.

b. She got a job.

(5.42) $\langle A, \mathcal{F}, LAST \rangle$, where
$A = \{\pi_0, \pi_1, \pi_2\}$
$\mathcal{F}(\pi_0) = Explanation(\pi_1, \pi_2)$
$\mathcal{F}(\pi_1) = K_{\pi_1}$
$\mathcal{F}(\pi_2) = K_{\pi_2}$
$LAST = \pi_2$

(5.43)



The language of SDRSs is dynamic and undecidable. The meaning of discourses is, like in DRT, a relation between input and output contexts. The semantics of SDRSs extend the semantics of DRSs (see the definition in section 3.4.3 on page 89) by assigning meaning to rhetorical relations. While DRS-conditions in form of predicates are interpreted as imposing a test on the input context, rhetorical relations define a real transition from an input to an output context. They represent speech acts that, like other actions, can change the context. Recall from chapter 4 that veridical rhetorical relations (cf. Fig. 4.1 on page

121) are relations that entail the contents of both discourse segments. These relations, including, among others, CONTINUATION, NARRATION, RESULT, CONTRAST, PARALLEL, BACKGROUND, ELABORATION, EXPLANATION, satisfy the following schema:

Veridical Discourse Relations [6]

$(w, f)[\![R(\pi_1, \pi_2)]\!]_M(w', g)$ iff $(w, f)[\![K_{\pi_1} \oplus K_{\pi_2} \oplus \phi_{R(\pi_1,\pi_2)}]\!]_M(w', g)$

(Asher and Lascarides, 2003, p. 459)

Roughly speaking, $R(\pi_1, \pi_2)$ is true iff both K_{π_1} and K_{π_2} are true and the additional condition ϕ is met. This condition must be spelled out for particular rhetorical relations. Specific axioms in form of meaning postulates determine their interpretation. For example, NARRATION entails a temporal ordering of the main events described in the segments. In contrast, EXPLANATION entails that the second event cannot occur before the first one.

Temporal Consequence of Narration

$\phi_{Narration(\pi_1,\pi_2)} \Rightarrow \pi_1 : \boxed{e_1} \prec \pi_2 : \boxed{e_2}$

Temporal Consequence of Explanation

$\phi_{Explanation(\pi_1,\pi_2)} \Rightarrow \pi_1 : \boxed{e_1} \not\prec \pi_2 : \boxed{e_2}$

The semantic effect of ELABORATION on the discourse content can be described as follows: An ELABORATION relation entails the existence of a *part-of* relation between the eventualities expressed by two utterances.

Consequence of Elaboration

$\phi_{Elaboration(\pi_1,\pi_2)} \Rightarrow part\text{-}of(\pi_2 : \boxed{e_2}, \pi_1 : \boxed{e_1})$

Non-veridical discourse relations, including ALTERNATION and CONSEQUENCE, do not entail the content of at least one of their arguments. However, their dynamic semantics is straightforward. They do not change the world-assignment pair representing the context, i.e. $(w, f) = (w', g)$.

Alternation

$(w, f)[\![Alternation(\pi_1, \pi_2)]\!]_M(w', g)$ iff $(w, f)[\![K_{\pi_1} \vee K_{\pi_2}]\!]_M(w', g)$

(Asher and Lascarides, 2003, p. 460)

Consequence

$(w, f)[\![Consequence(\pi_1, \pi_2)]\!]_M(w', g)$ iff $(w, f)[\![K_{\pi_1} \Rightarrow K_{\pi_2}]\!]_M(w', g)$

(Asher and Lascarides, 2003, p. 460)

So far I have sketched how structured discourse meaning is represented in SDRT. I will turn now to the question of how these representations are constructed.

6 The merge operator $\oplus$ was defined on page 89.

5.3.2 Constructing Discourse Structures

5.3.2.1 The Logic of Underspecified Information Content

The Language $\mathcal{L}_{ulf}$ of underspecified information content describes the form of SDRSs. It allows underspecifications. All it knows about is the *form* of determinate logical forms, but not their content. Take, for instance, example (5.41b) above. From the linguistic input, we get an underspecified SDRS as shown in (5.44b), where the reference of the pronoun "she" still has to be resolved.

(5.44) a. She got a job.

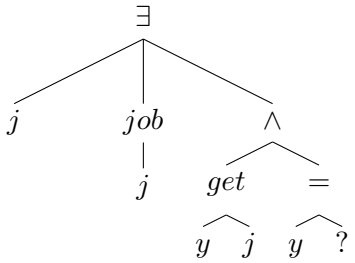
b. π_2 :

y, j, e_2, t_2
$job(j)$, $get(e_2, y, j)$, $holds(e_2, t_2), t_2 \prec now, y =?$

This SDRS, slightly simplified (without information about events and tenses), corresponds to the first-order formula (5.45a), which, in turn, can be represented as a tree (5.45b).

(5.45) a. $\exists j[job(j) \wedge get(y, j) \wedge y =?]$

b.



Every tree corresponds to a model of $\mathcal{L}_{ulf}$, so that $M \models_{\mathcal{L}_{ulf}} \phi$ means: the underspecified logical form ϕ (partially) describes the unique determinate logical form that corresponds to M.

The vocabulary of $\mathcal{L}_{ulf}$ consists of labels which pick out nodes in a tree representing the logical form. All constructors in the SDRS-vocabulary become predicate symbols over labels in $\mathcal{L}_{ulf}$. Accordingly, a discourse referent in an SDRS becomes a one-place predicate in $\mathcal{L}_{ulf}$, its argument being the label that tags its position in the tree. In this way, the compositional semantics of a pronoun involves not knowing the value of a predicate in $\mathcal{L}_{ulf}$, and is represented by a higher-order variable. The gloss $x =?$ stands for $\exists \mathfrak{P}[R_{=}(l_x, l_y, l) \wedge R_x(l_x) \wedge \mathfrak{P}(l_y)]$.

The satisfaction relation $\models_{\mathcal{L}_{ulf}}$ is defined relative to finite first order models (higher-order variables are interpreted substitutionally). $\models_{\mathcal{L}_{ulf}}$ is monotonic, extensional, static and decidable, in contrast to the logic of the SDRSs themselves. It does reasoning about the form of SDRSs, not their dynamic interpretation. $\models_{\mathcal{L}_{ulf}}$ relates an underspecified logical form to all possible ways of resolving its underspecifications[7].

7 There are other logics designed for representing underspecifications, including CLLS (Constraint Language on Lambda Structures, Egg, 2005), Hole Semantics (Bos, 2001), and Minimal Recursion Semantics

5.3.2.2 The Glue Logic

SDRT has a special logic for inferring discourse structures. In the *glue logic* is computed how the logical form for a new discourse segment is to be integrated with the representation of the discourse evolved so far. Like $\mathcal{L}_{ulf}$, the glue logic has only limited access to the logic of SDRSs: it knows about their form but not about their interpretation. The reason is to ensure that constructing semantic form (or computing what is said) is computable.

In this logic, the *pragmatically preferred* interpretations of underspecified logical forms are computed and the following inferences are performed: (i) Infer the (pragmatically preferred) values of underspecified conditions generated by the grammar, (ii) infer what is rhetorically connected to what (equivalent to text segmentation), and (iii) infer the values of the rhetorical relations. This information is computed on the basis of default inferences within the glue logic.

Defaults in the glue logic are expressed by the nonmonotonic conditional operator '>', which was introduced in section 1.2.3 (page 30). Recall that $*$ was defined as a function from worlds and sets of possible worlds to sets of possible worlds. Here, this operator is dynamized by replacing sets of possible worlds by dynamic propositions, i.e. relations between world-assignment pairs. Thus, the relational semantics of this operator in the SDRS language is defined as follows:

Nonmonotonic Consequence

$(w, f)[\![K > K']\!]_M(w', g)$ iff $(w, f) = (w', g)$ and for all w'', h
with $(w, f)\,[*(w, [\![K]\!]_M)]\,(w'', h)$ there are w''', k such that $(w'', h)[\![K']\!]_M(w''', k)$.

Glue logic axioms have the form $A > B$ with the meaning: *if A then normally B*. The glue logic defines a nonmonotonic consequence relation $\mid\!\sim_g$ over $\mathcal{L}_{ulf}$. In the following, I will give some examples of glue logic axioms (cf. Asher and Lascarides, 2003, p. 199ff and Lascarides and Asher, 2007, p. 23ff). For a detailed inventory of glue logic axioms, the reader is referred to the appendix in Asher and Lascarides (2003).

Note that, in the glue logic, rhetorical relations are represented as three-place relations. $R(\alpha, \beta, \lambda)$ means that β is to be attached to α with a rhetorical relation R, and the resulting structure is labelled λ. Thus, the glue logic term $R(\alpha, \beta, \lambda)$ amounts to $\lambda : R(\alpha, \beta)$ in the SDRS language.

Narration can be inferred if an event provides an "occasion"[8] for another event to occur.

Narration : $(?(\alpha, \beta, \lambda) \wedge occasion(e_\alpha, e_\beta)) > Narration(\alpha, \beta, \lambda)$

A relation *occasion* normally holds between two events if there are two event types ϕ and ψ that are related somehow in terms of stereotypical script knowledge. Asher and Lascarides (2003, p. 201) formalize this in a somewhat imprecise fashion as follows.

- **Scripts for Occasion**: $(?(\alpha, \beta, \lambda) \wedge \phi(\alpha) \wedge \psi(\beta)) > occasion(e_\alpha, e_\beta)$

(Copestake et al., 2005). I suppose that any of the mentioned formalisms can be applied for representing underspecified semantic forms in a similar manner.

8 In Hobbs (1985a)'s account (cf. section 4.2.1 on page 105), Occasion is modelled as a proper discourse relation, which corresponds, in fact, very closely to SDRT's Narration relation.

There are attempts to infer NARRATION on the basis of a common discourse topic of two utterances. However, I will not step into the rather complicated details here. For some hints, see Asher (2004a) and the discussion of discourse topic in section 4.3.

Intuitively and formally more satisfying are the axioms for inferring ELABORATION and EXPLANATION. Consider first EXPLANATION.

Explanation : $(?(\alpha, \beta, \lambda) \wedge cause_D(\beta, \alpha)) > Explanation(\alpha, \beta, \lambda)$

(Asher and Lascarides, 2003, p. 206, slightly simplified)

In order to be able to apply this axiom, we need clues to infer the relation between events $cause_D$. We get this by virtue of the following rule[9]:

- **Causation and Change**:
  $(change(e_\alpha, x) \wedge cause - change - force(e_\beta, x)) \rightarrow cause_D(\beta, \alpha)$

This rule reflects the fact that if an object y is changed (in location) by an event e_α and if e_β is an event describing a force that can cause a change in location of y, then there is evidence in the discourse that e_β might be a cause of e_α, i.e. $(cause_D(\beta, \alpha))$.

With these bits of world knowledge, we can infer the discourse relation holding in our example (5.41b). To move is a change of location, thus $change(e_{\pi_1}, x)$ holds, and getting a new job is reason to move, yielding $cause - change - force(e_{\pi_2}, x)$. We thus can infer an EXPLANATION relation between the two utterances in the example.

ELABORATION can be nonmonotonically inferred by the following glue logic axiom[10]:

Elaboration : $(?(\alpha, \beta, \lambda) \wedge subtype_D(\beta, \alpha)) > Elaboration(\alpha, \beta, \lambda)$

(Lascarides and Asher, 2007, p. 26, slightly simplified)

The relation $subtype_D$ holds between two clauses α and β if the following preconditions are met: (i) there are two discourse referents x and y, x is participant of the eventuality e_α and has the theta-role θ_i, (ii) y has the same theta-role in e_β, (iii) the sorts[11] assigned to x and y in the lexicon stand in a mereological subtype relation $\sqsubseteq$[12], and (iv) the sort of the second eventuality is a subtype of the sort of the first one:

- **Subtype**: $(\theta_i(x, e_\alpha) \wedge \theta_i(y, e_\beta) \wedge y \sqsubseteq x \wedge e_\beta \sqsubseteq e_\alpha) \rightarrow subtype_D(\beta, \alpha)$

(Asher and Lascarides, 2003, p. 283)

To illustrate this rule, suppose that x is a discourse referent introduced by the utterance of the lexical unit "meal", and y similarly stands for "meat". From world knowledge we can infer that "meal" is lexically specified to be of sort **food** (cf. Asher and Lascarides, 2003, p. 282). Furthermore, "meat" is also of sort **food**, and, more specifically, it is food

9 Note that the notation e_α is an abbreviation for the fact that e_α is the main eventuality expressed in utterance α. In the DRS-like box-notation of the logic of information content, this is represented by $\alpha : \boxed{e_\alpha}$, i.e. $e_\alpha \in \mathcal{F}(\alpha)$.

10 There are other ways of inferring ELABORATION, see Bras (2007) for more details.

11 Asher and Lascarides (2003) use the term "lexical type" for what is usually called "sort". To avoid confusions, I will use the term "sort" in this thesis; see also footnote 5 on page 193.

12 The subtype relation $\sqsubseteq$ defines a partial ordering on the set of sorts. It is a reflexive, antisymmetric, and transitive relation. See Asher and Pustejovsky (2005) for a formal theory of sorts (viz. lexical types) using this relation.

made of animals. Thus, we can write $y \sqsubseteq x$. Similarly, if we have an eventuality e_α and a sub-eventuality e_β, e.g. cooking and preparing meat or cleaning the kitchen and washing the dishes, then $e_\beta \sqsubseteq e_\alpha$ holds.

As a last important glue logic axiom, I will give SDRT's default rule for inferring a BACKGROUND relation (cf. Asher and Lascarides, 2003, p. 207, Vieu and Prévot, 2004, p. 486), which can be defined as follows:

Background : $(?(\alpha, \beta, \lambda) \wedge event(e_\alpha) \wedge state(e_\beta)) > Background(\alpha, \beta, \lambda)$

That is, if e_α is an event (in the terminology of Vendler, 1957) and e_β is a state, then, unless there is information to the contrary, a BACKGROUND relation holds between the corresponding utterances.

5.3.2.3 Discourse Update

So far, the glue logic allows inferences about which rhetorical relation holds between two discourse segments. For updating a possibly complex discourse representation with the content of a new discourse segment, the following steps have to be taken:

1. the site of attachment in the SDRS representing the preceding discourse has to be found,
2. the most plausible discourse relations for attaching the new information has to be computed, and
3. the new information, together with the discourse relation, has to be added to the discourse structure.

The old information, that is the discourse structure constructed so far, is a set σ consisting of discourse structures which satisfy the constraints on interpretation processed so far, i.e. without the new information. The new information is represented by an underspecified logical form $\mathcal{K}_\beta$. From a procedural perspective, discourse update consists of three steps: (1) first pick a set of attachment sites out of the available labels in the structures in σ. (2) Then for each site α of this set, assume $?(\alpha, \beta, \lambda)$ and compute a rhetorical relation between α and β from glue logic axioms and information transferred from σ and $\mathcal{K}_\beta$ into the glue logic. (3) Finally, eliminate all SDRSs from σ which result in inconsistencies when updated with $\mathcal{K}_\beta$ or which do not maximize discourse coherence. Step (2) is performed by the glue logic, step (1) is subject of section 5.3.2.4, and step (3) will be dealt with in section 5.3.2.5[13].

13 In fact, discourse update in SDRT is defined declaratively. It is defined as a sequence of simple update operations $+$, where $+$ is defined in terms of the glue logic consequence relation $\mid\sim_g$.

The Simple Update $+$ (cf. Lascarides and Asher, 2007)

Let σ be a set of (fully specified) discourse structures, and let ψ be
(a) either an underspecified logical form $\mathcal{K}_\beta$ or (b) a formula $?(\alpha, \beta, \lambda)$ about attachment where $Th(\sigma) \models_{\mathcal{L}_{ulf}} \mathcal{K}_\beta$. Then $\sigma + \psi$ is a set of SDRSs defined as follows:

1. $\sigma + \psi = \{\tau : \text{if } Th(\sigma), \psi \mid\sim_g \phi \text{ then } \tau \models_{\mathcal{L}_{ulf}} \phi\}$, provided the result is not $\emptyset$;
2. $\sigma + \psi = \sigma$ otherwise.

The result of $+$ is a set of discourse structures which is a subset of the old information σ (it satisfies the old and the new information). Moreover, it ensures that any $\mid\sim_g$-consequences of the old information and the new are satisfied, too. For more formal details, see Asher and Lascarides (2003) or Lascarides and Asher (2007).

In essence, SDRT update is a two-place function. It takes (i) a set σ of discourse structures satisfying (a) the underspecified logical forms and (b) the constraints following from these underspecified logical forms in the glue logic and (ii) an underspecified logical form $\mathcal{K}_\beta$ representing the new information. SDRS update does not return a unique updated SDRS, rather its output is a set σ' of updated SDRSs. Thus, even though the update process may resolve underspecifications where possible, some underspecifications can still remain, yielding an ambiguous discourse.

5.3.2.4 Constraining Attachment

In the last section, I have taken for granted that we know at which point in the existing discourse structure the new information conveyed by an utterance is to be attached. A closely related question is how attachment affects the update of context in general and the resolution of anaphora in particular. This point will be important in the last part of this thesis.

In SDRT, as in other theories, it is assumed that attachment obeys the Right Frontier Constraint (RFC) which was introduced in chapter 4 on page 100. Basically, the RFC says that the last attached discourse segment and segments that dominate it are available for attachment. Remember from section 4.2.4 that SDRT makes a basic distinction between coordinating and subordinating discourse relations. This allows to state the RFC in the following way: a coordinating relation pushes the right frontier to the right, closing off its current attachment point, and a subordinating relation extends the right frontier downwards, leaving open its current attachment point. More formally,

Available Nodes for Attachment in a discourse structure are[14]

1. the label $\alpha = LAST$
2. any label $\gamma \geq_D^* \alpha$, where $\geq_D^*$ is defined recursively:
   a) $R(\gamma, \alpha)$ is a conjunct in $\mathcal{F}(l)$ for some label l, where R is a subordinating discourse relation;
   b) $R(\gamma, \delta)$ is a conjunct in $\mathcal{F}(l)$ for some label l, where R is a subordinating discourse relation and $\mathcal{F}(\delta)$ contains as a conjunct $R'(\delta', \alpha)$ or $R'(\alpha, \delta')$ for some R' and δ'; or
   c) $R(\gamma, \delta)$ is a conjunct in $\mathcal{F}(l)$ for some label l, where R is a subordinating discourse relation and $\delta \geq_D^* \alpha$.

With this at hand, the resolution of anaphora can be constrained as follows. Suppose that β is a label for an SDRS $\mathcal{K}_\beta$ containing an anaphoric condition ϕ. Then, available antecedents for ϕ are

Available Antecedents for Anaphora

1. in $\mathcal{K}_\beta$ and DRS-accessible[15] to ϕ,
2. in $\mathcal{K}_\alpha$, DRS-accessible to any condition in $\mathcal{K}_\alpha$, and there is a condition $R(\alpha, \gamma)$ in the SDRS such that $\gamma = \beta$ or $\gamma \geq_D^* \beta$ (where R is not structural[16]).

14 Definitions are from Asher (2008, p. 32).

15 DRS-accessibility was defined in chapter 3 on page 91. Informally, DRS-accessible discourse referents are either immediately to the "left" of the current DRS, or in the immediately superordinated DRS.

16 Structural discourse relations are Parallel or Contrast. In Kehler (2002)'s terminology, these are Resemblance relations (cf. chapter 4 on page 107).

In short, an antecedent for an anaphoric expression must be DRS-accessible on the right frontier, unless the connecting discourse relation is not PARALLEL or CONTRAST. Note that anaphoric availability in SDRT is not a strictly structural constraint, as it depends on the semantics of the involved rhetorical relations. In PARALLEL and CONTRAST relations, anaphoric availability is a byproduct of constructing a maximal partial isomorphism between the related constituents in the way Kehler (2002) proposes. There are other cases where the Right Frontier Constraint seems to be violated, as in certain cases of NARRATION and in the resolution of definite descriptions and plural anaphora. For recent discussions of these topics and their relation to the RFC, see Asher (2008) and Vieu and Prévot (2008).

5.3.2.5 Maximize Discourse Coherence

As we have seen above, SDRT update does not yield a unique update discourse structure, but a set of preferred structures. This is the point where a basic assumption of SDRT comes into play, the principle *Maximize Discourse Coherence* (MDC). It consists of four parts, which can be informally described as follows:[17]

Maximize Discourse Coherence (MDC)

(i) All else being equal, an interpretation with a smaller number of labels is to be preferred.

(ii) A consistent interpretation is preferred over an inconsistent one.

(iii) An interpretation that maximizes the number and the quality of its rhetorical relations is more coherent than one that does not.

(iv) The more ambiguities (including anaphoric conditions) can be resolved, the higher is the quality of the discourse interpretation.

Recall that SDRSs are models of $\mathcal{L}_{ulf}$. The *ranking* of the models in the set of updated discourse structures σ' is done via the MDC. The content of a discourse at a given point will be those things that follow from the highest ranked SDRSs in the update. According to the definition of MDC, these are those with

1. the minimum number of labels,
2. no inconsistencies,
3. the maximum number and the highest quality of rhetorical connections, and
4. the fewest unresolved semantic ambiguities (including anaphoric conditions).

The MDC governs the selection of a preferred discourse update at a given point in a discourse. Note that this task can be seen as a form of Model Generation. However, a tricky point is the question of how the minimality of models is to be defined in order to minimize SDRSs as models of $\mathcal{L}_{ulf}$. The MDC seems to be the "magic" part of SDRT: it solves all the problems which cannot be solved otherwise. I will propose later in this thesis (chapter 7) how some parts of the burden can be taken from the MDC by assuming a small number of general cognitive constraints in discourse interpretation.

17 See Asher and Lascarides (2003, pp. 233f) for a comprehensive definition of MDC.

5.3.3 Conclusion

SDRT is the formal theory of discourse interpretation that is currently worked out in most detail. At first sight, the theory looks very complicated, but a closer look reveals that the complexity of SDRT is a reflection of the complex processes in discourse interpretation.

The representation of complex discourses as graph structures overcomes some potential problems with tree structure representations (see our discussion in section 4.1.3) in other theories. The definition of the semantics of SDRSs as an extension of DRT is straightforward. Some problems of DRT concerning anaphoric binding between sentences can be overcome by assuming rhetorical connections between discourse segments.

The logical form for discourses in SDRT is obtained via a complex process involving nonmonotonic reasoning on information from different information sources, including lexical and compositional semantics, domain knowledge, and information about cognitive states. These information sources can give conflicting clues for different interpretations: alternative interpretations can arise in cases where more than one default is applicable. The specificity principle is an important constraint for resolving these conflicts. However, since it is not always applicable, additional criteria for choosing adequate discourse updates must be found. One of the basic principles is Maximize Discourse Coherence.

The architecture of SDRT, especially its separation of distinct logics for different concerns, ensures that the process of computing pragmatically preferred logical forms for discourses is decidable, even though pragmatic inferences involve nonmonotonic reasoning. Lexical knowledge and domain knowledge is kept apart in logical discourse representation, in contradistinction to Hobbs' abductive approach.

In this thesis, I assume a model of discourse interpretation based on SDRT in its standard version. However, some characteristics of the other theories presented in this chapter will be integrated in the adopted extended theory. In particular, I will propose to perform the ranking of the models in the SDRT update according to general cognitive constraints as a form of minimal model generation.

Chapter 6

Bridging Inferences

6.1 Bridging Anaphora

As we have seen in chapter 3, discourse anaphora express, in general, certain relationships between referents in a discourse model. Various types of discourse anaphora can be distinguished. Direct anaphora bear a coreference relation between two discourse referents, e.g. the relation between a pronoun and a previously mentioned entity. Indirect anaphora connect two discourse referents by means of a relationship different from coreference. In his influential paper, Clark (1975) called this type of anaphora *bridging anaphora*. Other terms used in the literature are "associative anaphora" (Hawkins, 1978), "inferrables" (Prince, 1981), or "indirect anaphora" (Chafe, 1976; Schwarz, 2000).

Generally, in an indirect or bridging anaphor, an entity introduced in a discourse, the *anaphor*, *stands in a particular relation*, which is different from coreference, to some previously mentioned discourse entity, the *anchor*. This relation – the *bridging relation* – is not explicitly stated by linguistic means. Yet it is an essential part of the discourse content because the knowledge of these relations is necessary for successfully interpreting a discourse. A characterizing property of indirect anaphora is the absence of an explicitly linguistically expressed *antecedent*.

There has been a considerable amount of work on anaphoric reference in general. Studies relevant to bridging anaphora can be found in various disciplines:

- psycholinguistic studies (Haviland and Clark, 1974; Clark, 1975; Singer, 1979; Garrod and Sanford, 1982; Garrod and Terras, 2000; Burkhardt, 2006),
- corpus-based studies on the use of definite expressions (Hawkins, 1978; Fraurud, 1990; Poesio and Vieira, 1998; Gardent et al., 2003),
- pragmatic and cognitive accounts (Erkü and Gundel, 1987; Wilson and Matsui, 1998; Matsui, 2000; Schwarz, 2000),
- computational accounts (Sidner, 1981; Nissim, 2001; Freitas, 2005), and
- formal accounts (Bos et al., 1995; Asher and Lascarides, 1998a; Piwek and Krahmer, 2000).

First, I will give a preliminary classification of bridging anaphora in section 6.1.1. Then, I will review evidence based on corpus-based investigations (section 6.1.2) and psycholinguistic studies (section 6.1.3), before presenting a refined classification in section 6.1.4.

Bearing on the proposed classification, I will discuss different types of bridging relations in section 6.2. Finally, pragmatic, computational, and formal accounts of resolving bridging anaphora will be subject of section 6.3.

6.1.1 A Preliminary Classification

Clark (1975, 1977) differentiated various kinds of bridging anaphora. His original definition of bridging anaphora subsumes both direct and indirect anaphora. So he distinguished between "direct reference" and "indirect reference". Direct reference means that the antecedent is an entity just mentioned, as in (6.1).

(6.1) I met a man yesterday. He told me a story.

Among bridging anaphora involving indirect reference, various subtypes can be made out. The most prominent and most studied type is what Clark called "indirect reference by association", where the antecedent is not directly mentioned, but *closely associated* with an entity mentioned before. Putting their emphasis on this subtype of indirect anaphora, some authors use "associative anaphora" as a synonym for indirect anaphora. The associated pieces of information vary in their predictability: they can be necessary parts (6.2), probable parts (6.3) or inducible parts (6.4).

(6.2) I looked into the room. The ceiling was very high.

(6.3) I walked into the room. The windows looked out to the bay.

(6.4) I walked into the room. The chandeliers sparkled brightly.

Another type of bridging anaphora consists of "indirect reference by characterization", where the bridging relation *characterizes a role that something implicitly plays* in an event or circumstance mentioned before. Roles can be agents, objects, or instruments. Again, they can be necessary roles (6.5) or optional roles (6.6).

(6.5) John was murdered yesterday. The murderer got away.

(6.6) John was murdered yesterday. The knife lay nearby.

A third type of bridging anaphora mentioned by Clark concerns reasons, causes, consequences, concurrences, and other relations between events. I will not use the term bridging for these relations because they are more closely related to what I have called coherence relations before. For instance, Clark's example for consequences is (6.7). I rather assume that the two utterances are connected by a coherence relation Result.

(6.7) John fell. What he did was break his arm.

The classification of Clark has been very influential for subsequent work on bridging. In the next sections, I will review empirical work made in the last decades based both on corpus-studies and on psycholinguistic experiments.

6.1.2 Corpus Studies on Anaphoric Expressions

Various empirical investigations on bridging anaphora have been made in the last decades. Most of them started from the observation that bridging relations are triggered mainly by definite noun phrases. Nevertheless, there are also cases of referring, nonquantificational indefinite noun phrases which convey a bridging relation. For example, in (6.8), "a knife" clearly refers to the probable instrument of murdering, almost identically as in example (6.6).

(6.8) John was murdered yesterday. A knife lay nearby.

We will have to keep this in mind because most literature on indirect anaphora focuses on definite descriptions, largely ignoring indefinite noun phrases.

As observed by Russell (1905), definite descriptions are characterized by two properties: *uniqueness* and *existence*. Uniqueness (Russell, 1905; Kadmon, 1990; Roberts, 2003) means that the referent denoted by the definite description must be the only referent that satisfies the given description. Uniqueness is always relative to some restricted context, e.g. a quantification domain or the discourse domain.

Russell's condition of the existence of a referent is subsumed by the more recent notion of *familiarity* (Heim, 1982; Prince, 1981). This notion and related ideas were discussed in chapter 3 on pages 73ff.

An important corpus study on the use of referring expressions was made by Fraurud (1990). She classified both definite and indefinite noun phrases based on a Swedish corpus consisting of written texts of various sorts. She found that 61% of all definite descriptions were "first-mentioned"[1], i.e. without a coreferring noun phrase antecedent, whereas 39% of the cases were "subsequent-mentioned", with a noun phrase antecedent. Regarding indefinite NPs, almost all cases (85%) where first-mentioned. Evidently, subsequent-mentioned NPs are direct anaphora, while indirect anaphora are subsumed by the class of first-mentioned NPs. What is perhaps the most interesting result of her study, is the fact that the case of "first-mentioned" is by no means clearly restricted to indefinite NPs. Thus, the case of indirect anaphora triggered by definite descriptions is not a marginal special case, contrarily to the prevailing previous view according to which indefinite NPs introduce new discourse referents and definite NPs refer back to previously introduced ones. The two classes used by Fraurud correspond to what Roberts (2003) called "strongly familiar" and "weakly familiar", respectively.

	indefinite NPs	definite NPs
first-mentioned	85.2%	60.9%
subsequent-mentioned	8.3%	36.1%
others	6.5%	3.0%

Table 6.1: Distribution of indefinite and definite NPs according to Fraurud (1990)

Poesio and Vieira (1998) investigated the use of definite NPs and proposed a classification based on native-speaker annotations following a predefined classification scheme based on a corpus of English newspaper articles. Their classes are given in Table 6.2, together with a comparison to the classes proposed by Prince, Fraurud, and Roberts. We can identify their class I with Fraurud's "subsequent-mentioned" class, and the rest as "first-mentioned", yielding very similar results. What is becoming more clearly in their study is the fact that only a quarter of all first-mentioned definite NPs is really unfamiliar or brand-new (class IV), while the NPs in class II and III are neither really "old" (in Prince's sense) nor entirely "new". This status of indirect anaphora is emphasized by Schwarz-Friesel (2007).

1 In fact, the "first-mentioned" class compromises cases of initial mentions in a coreferential chain and uniquely occurring cases of isolated mentions.

Class	Poesio & Vieira	Frequency	Prince	Fraurud	Roberts
I	coreferential NPs	43-45%	discourse old	subsequent-mentioned (39%)	strongly familiar
II	bridging	6-11%	discourse new, but related to a discourse-old entity	first-mentioned (61%)	weakly familiar
III	larger situation	20-25%	discourse new and hearer-old		
IV	unfamiliar	18-26%	discourse-new and hearer-new and not related to any discourse-old entity		

Table 6.2: Classifications of definite descriptions

In another corpus annotation study, Gardent et al. (2003) found a different number of cases of bridging definite descriptions. These authors distinguished between "first mention" (78.4%), "coreferential" (16.8%) and "bridging" cases (4.7%) in a corpus of French newspaper articles. Number differences in general can have a number of reasons, among them differences between spoken and written language, or between different text sorts underlying the used corpora, or as a result of different postulations of classes. In this particular case, the difference may be a result of the exclusion of event and discourse deictic anaphora in the latter study.

Nevertheless, Gardent et al. (2003) further differentiate between 5 classes of bridging relations (see Table 6.3): set membership, thematic, definitional, co-participants, and non-lexical. The set membership class compromises 5.8% of their data. The thematic class, which covers 5.3% of their corpus, corresponds to Clark's indirect reference by characterization. The majority of cases (83%) is attributed to the definitional and co-participants classes where the semantic relation is given by lexicographic definitions in form of dictionary entries of either antecedent or anaphor.

Class	Occurrences	Source
set membership	5.8%	hyponymy
thematic	5.3%	event thematic grid
definitional	80.8%	lexicographic definition
co-participants	2.2%	lexicographic definitions
non-lexical	7.8%	discourse structure, world knowledge

Table 6.3: Annotation of bridging definite descriptions (Gardent et al., 2003)

A problem with this classification is that it conflates two orthogonal distinctions. For instance, it is not clear what differentiates cases of "set membership" from "definitional" meronymic relations like "collection/member" or "whole/part". Another question not so easily answered is to what extent a relation is "lexical" or "non-lexical". I will come back to this question in the next section. A more serious problem for being a basic classification is that it requires a well-sorted information source to be available for the hearer which could provide the lexicographic definitions needed in the definitional and co-participants classes.

However, Gardent et al. (2003) make two interesting suggestions: First, the meronymic relations (52%) could be processed using **WordNet** (Fellbaum, 1998), although only 38 of their 187 meronymic cases were actually present in WordNet. Second, the thematic relations (5.3%) could be processed by **FrameNet** (Baker et al., 1998). In 14 of their 19 thematic cases, a frame containing a target and an anchor as frame element were found. I will exploit this suggestion in chapter 7.

To conclude, a large part of all occurrences of definite NPs, as well as cases of indefinite NPs, do involve indirect anaphora, thus a view that makes of direct anaphora the paradigmatic case and of indirect anaphora a marked case turns out not to be empirically adequate. At the moment, I can state that uniqueness is not a necessary condition for the occurrence of bridging anaphora, while some kind of familiarity certainly does play an important role in bridging inferences.

6.1.3 Psycholinguistic Investigations

In this section, I will report psycholinguistic research on bridging anaphora. In his numerous studies on the subject, Clark repeatedly emphasized the role of the interaction of communication participants and the Common Ground in the resolution of bridging anaphora (Haviland and Clark, 1974; Clark, 1975, 1977, 1996). What he called the "Given-New contract", is a mutual agreement of speaker and hearer that the content of an utterance is composed of given and new information. This contract can be seen as part of Grice's principle of cooperativeness (cf. section 1.1.2.1).

> "Given-New Contract: The speaker agrees to try to construct the Given and New information of each utterance in context (a) so that the listener is able to compute from memory the unique Antecedent that was intended for the Given information, and (b) so that he will not already have the New information attached to the Antecedent."
>
> (Clark, 1975, p. 175)

The interpretation of anaphora then is supposed to consist of searching memory for a matching antecedent to the given information, and on finding it, attaching the new information to the antecedent. If no antecedent is found, then either a bridging structure must be built or a recomputation of given and new information must take place. This model predicts that direct anaphora are easier to process than indirect anaphora. Indeed, Haviland and Clark (1974) found experimental evidence that comprehension time of sentence pairs containing a direct anaphor like (6.9) was faster than in sentence pairs with an indirect anaphor like (6.10).

(6.9) We got some beer out of the trunk. The beer was warm.

(6.10) We checked the picnic supplies. The beer was warm.

The results of Haviland and Clark (1974) were replicated for bridging anaphora involving events (Clark's indirect reference by characterization) in a series of experiments carried out by Singer (1979). He compared reading times for sentences such as (6.12) below following a sentence that either explicitly mentioned the instrument *shovel* (6.11b) or only presupposed it (6.11a).

(6.11) a. The boy cleared the snow from the stairs.

b. The boy cleared the snow with a shovel.

(6.12) The shovel was heavy.

Singer found that the reading time of (6.12) was longer in the context of (6.11a) as compared to (6.11b), and suggested that the assignment of a thematic role to a lexical expression is not an automatic process.

In contrast to the previously mentioned studies, Garrod and Sanford (1982) found no difference in reading times of (6.14) when following (6.13a) or (6.13b), and (6.16) when following (6.15a) or (6.15b).

(6.13) a. Mary put the clothes on the baby.

b. Mary dressed the baby.

(6.14) The clothes were made of pink wool.

(6.15) a. Keith drove to London yesterday.

b. Keith took his car to London yesterday.

(6.16) The car kept breaking down.

Garrod and Sanford propose a scenario-based account of text comprehension. A "scenario", or "frame", is a particular part of world knowledge that is activated in the course of interpretation. The information stored in a stereotypical scenario is then used for understanding indirect anaphora. I will outline a formal account based on this idea in chapter 7.

Another possible explanation for the different findings in Singer (1979) and Garrod and Sanford (1982) is that the semantic association between *clear* and *shovel* is weaker than that of *drive* and *car*[2]. It thus may be possible that bridging relations are only readily available when there is a strong semantic relationship between bridging anchor and anaphor. It could be argued that this relation is part of lexical knowledge associated with bridging anchors, especially when the anchor is a verb.

Garrod and Terras (2000) carried out an eye-tracking experiment in order to address the question whether the bridging relation between verb and anaphor (role-filler) is mediated by the context in which the verb occurs ("contextual account"), or rather results from a purely lexical association between the verb and its role-filler ("lexical account"). They compared the processing of target sentences (6.18) when following one of the sentences in (6.17).

(6.17) a. The teacher was busy writing a letter of complaint [with a pen / to a parent].

b. The teacher was busy writing an exercise on the blackboard [with chalk / ∅].

2 It should be noted that if *clear the snow* is taken to be a complex lexical expression, then its relation to *shovel* is certainly stronger.

(6.18) However, she was disturbed by a loud scream from the back of the class and the [pen/chalk] dropped on the floor.

Two main findings resulted from Garrod & Terras' experiment. First, it was observed that first-pass reading times of the target sentence are slower when following an implicit introduction of the thematic argument only for *nondominant* verb-role pairs (e.g. "chalk" in the context of "write"). No difference was found between explicit and implicit antecedents for *dominant* verb-role pairs ("pen" in the context of "write"). The second finding was the observation of an early context effect for the dominant but not the nondominant verb-role pairs. Reading times were faster following appropriate contexts ("write a letter" for "pen") than following inappropriate contexts ("write on the blackboard" for "pen"). No such effect was observed with nondominant verb-role pairs. However, a strong context effect emerged later, in the second-pass reading times on the noun, both for dominant and nondominant targets.

Garrod and Terras (2000) suggest that their results indicate a two-stage process of discourse role resolution (cf. also Sanford et al. 1993). In an initial *bonding* stage, a link between the referring expression and a previous verb is established. This is a low-level automatic process not influenced by the particular context in which the verb occurred. A second stage of *resolution* is a process which evaluates the link made in the bonding process with respect to the overall discourse context, recomputes it if necessary, and integrates it into the discourse model.

Most of the studies mentioned before indicated that given information is integrated more easily in the discourse model than new information. In addition, Haviland and Clark (1974) and Singer (1979) found that the establishment of indirect bridging relations involve more processing costs than direct identity relations. These studies were based on reading or comprehension time measures that can only be taken offline, i.e. after the processing has taken place.

In order to overcome potential error sources arising in offline experiments, Burkhardt (2006) carried out an online experiment measuring event-related brain potentials (ERPs) in order to investigate at what point during real-time sentence processing inferential knowledge influences the interpretation of referring expressions. It is commonly assumed that in ERP measurements, a negative deflection with a delay of around 400 ms ("N400") occurs in connection with semantic implausibilities and contextual incoherence. A reduced N400 emerges with the discourse integration of given noun phrases. Furthermore, a positive deflection with a latency of around 600 ms ("P600") can be analyzed as indicating increased processing costs during the integration of new information. Burkhardt (2006) started from the hypothesis that bridging anaphora share properties with both new and given information. ERP measures were taken on German target sentences like (6.20) following one of the contexts in (6.19) which favours the referent of the underlined expression to indicate either a case of bridging (a), or given (b), or new information (c).

(6.19) a. Tobias besuchte ein Konzert in Berlin.
'Tobias visited a concert in Berlin.'

b. Tobias besuchte einen Dirigenten in Berlin.
'Tobias visited a conductor in Berlin.'

c. Tobias unterhielt sich mit Nina.
'Tobias talked to Nina.'

(6.20) Er erzählte, dass der Dirigent sehr beeindruckend war.
'He said that the conductor was very impressive.'

Two findings resulted from this experiment. First, new NPs registered the most negative deflection on the N400, followed by bridging NPs and given NPs. Second, both new NPs and bridging NPs show a positive P600 effect, while given NPs do not exhibit any positivity. Critically, bridging NPs can be grouped together first (between 350 and 550 ms) with the given NPs, and later (from 600 to 900 ms) with the new NPs.

Thus, the interpretation of bridging NPs show matching patterns with the interpretation of both given and new information. Moreover, it seems that the processing cost observed with the integration of bridging NPs, which is reflected by the P600 effect, is more likely to be attributed to the need to establish a new discourse referent, rather than from complex inferences drawn in order to find a bridging relation.

6.1.4 A Refined Classification

A more recent classification of indirect anaphora with a cognitive emphasis was made by Schwarz (2000); Schwarz-Friesel (2007). She summarized the main characteristics of indirect anaphora as follows: There is no explicit antecedent to which the anaphor refers back. Instead, there is some kind of trigger or anchor in relation to which the anaphor is interpreted. The relation between the anchor and the anaphor is not based on coreference. There are restrictions for coding indirect anaphora with pronouns or demonstratives, as evidenced by Cornish (1999); Koenig and Mauner (1999). In addition, she notes that indirect anaphora require for their full interpretation a cognitive process involving the activation of knowledge structures. She emphasizes that indirect anaphora are cases of entities which are given and new entities at the same time, thus showing characteristics of both activation and reactivation processes.

With regard to the relationship between indirect anaphora and definite descriptions, Schwarz-Friesel emphasizes, referring to Fraurud (1990), that indirect anaphora are quite common and normal uses of definite reference. Although a definite article indicates accessibility of its referent, no direct referent can be found in the discourse structure. Implicit referents have to be accessible in the mental or discourse model, or in general conceptual space. But, as we said before, indirect anaphora are not restricted to definite descriptions only.

Schwarz-Friesel proposes the following classification of indirect anaphora. Basically, there is a distinction between semantic and conceptual types of indirect anaphora. The interpretation of semantic indirect anaphora depends on the activation of knowledge in the mental lexicon, while conceptual indirect anaphora involve the processing of more general world knowledge. Four basic types of indirect anaphora are distinguished:

Mereological semantic indirect anaphora involve a form of part-whole relationships between anchor and anaphor. These relations are part of common semantic knowledge. In (6.21), such a relation holds between the dead man and his temples.

(6.21) Er bemerkte jedoch im gleichen Augenblick, dass der Mann tot war. Die Schläfen waren durchschossen.

'At that particular moment he noticed that the man was dead. His temples were shot through.' ([from a novel of Friedrich Dürrenmatt (1950)]) (Schwarz-Friesel, 2007, p. 9)

Lexical/thematic semantic indirect anaphora are based on lexical knowledge or on thematic roles. For instance, (6.22) bears an *instrument* relation between unlocking the door and the key.

(6.22) Ich wollte rasch die Haustür aufschließen, weil ich das Telefon klingeln hörte. Der Schlüssel war aber tief unten im Einkaufswagen vergraben.

'I wanted to unlock the door quickly, because I could hear the telephone ringing. The key, however, was buried deeply in the trolley.' (Schwarz-Friesel, 2007, p. 9)

Frame-/script-based conceptual indirect anaphora base their interpretation on the activation of frame or script knowledge. (6.23) exhibits an implicit relation between the restaurant and the waiter.

(6.23) Ich kenne ein schönes Restaurant in Refrath. Das Essen ist köstlich, und der Kellner ganz besonders nett.

'I know a lovely restaurant in Refrath. The food is excellent and the waiter is an extremely nice guy.' (Schwarz-Friesel, 2007, p. 9)

Inference-based conceptual indirect anaphora require complex inferencing. For example, in (6.24), establishing a relation between the alleged assault and the rake is certainly not part of easily accessible grammatical or lexical knowledge.

(6.24) One night a man rushes into the police station and tells the policemen that he has just been knocked down in his garden. One policeman is asked to go and look for traces at the place of the assault. After a short time he returns with a huge swelling at his head and says "I solved the case." – "Bravo," says his boss, "and how did you do that?" – "I stepped on the rake, too." (Schwarz-Friesel, 2007, p. 10)

As Schwarz-Friesel (2007) also notes, there are many mixed cases where it is difficult to make out a single type of indirect anaphora. In particular, it is not clear where the borderline between thematic semantic and frame-based conceptual anaphora has to be drawn, as we will see in more detail in chapter 7. However, three classes can be clearly distinguished in Schwarz-Friesel's identification of conditions for anchoring referents of indirect anaphora:

> "The referent of an indirect anaphor must be either an identifiable part of the semantic structure of the preceding sentences, or it must be a default value of a specific frame or script, or it must be inferrable on the basis of cognitive plausibility determined by general world knowledge."
>
> (Schwarz-Friesel 2007, p. 11)

At this point, I want to clarify some possible notational confusions. In this thesis, I use the terms "bridging" and "indirect" anaphor synonymously. Schwarz-Friesel prefers to use the term "indirect anaphor" instead of "associative anaphor", "inferrable" or "bridging anaphor" because not all indirect anaphora can be explained by associative relations or involve complex inferencing processes. Schwarz (2000); Schwarz-Friesel (2007) and also Consten (2004) understand the term "inference" as referring to complex, mostly conscious,

inferencing processes involving intentions of the discourse participants. In contrast to this rather narrow understanding of "inference", I use it in a rather general sense referring to all kinds of inferential processes involved in language comprehension. In previous chapters of this thesis, I used the term "inference" for any inference the hearer has to draw beyond purely grammatical knowledge. In particular, defeasible inferences are necessary in all types of conceptual indirect anaphora and also, depending on one's view on what is part of the lexicon, in semantic indirect anaphora. Of course, the inferences drawn in (6.21) are of quite different nature than those of (6.23), as pointed out by Consten (2004, p. 76). Nevertheless, the resolution of mereological types of indirect anaphora often require certain inferences involving non-linguistic knowledge. In sum, I opt for a two-fold distinction of indirect anaphora as follows.

Mereological Indirect Anaphora
This group coincides with Clark's indirect reference by association and with Schwarz-Friesel's mereological semantic indirect anaphora. The anchor is an already established discourse referent and its semantic type is either an entity or a set of entities. The bridging relation can be any mereological relation such as *part-of*, *member-of* etc.

Frame-related Indirect Anaphora
This group, which is located more closely at the semantics/pragmatics borderline, corresponds to Clark's indirect reference by characterization and compromises both Schwarz-Friesel's groups of lexical/thematic semantic and frame-/script-based conceptual indirect anaphora, as it is difficult to draw a clear distinction between them. Here, the bridging anchor is an eventuality or a frame present in the discourse model, and the bridging relation is a thematic role (e.g *agent*, *instrument*, *theme*) played in the eventuality by the entity denoted by the anaphor.

Bridging anaphora involving more complex inferencing processes involving goals and intentions of communication participants (corresponding to Schwarz-Friesel's inference-based conceptual indirect anaphora) are subsumed by the frame-related group. Although in these cases the bridging relation can be a conceptual relation of any type, it is often a causal, temporal, or spatial bridging relation. All of these relations have to do with eventualities, even if they do not play as central a role as thematic relations.

The classification assumed in this thesis is given in Table 6.4.

Bridging class	Anchor type	Typical bridging relations
mereological	(sets of) entities	*part-of*, *member-of*
frame-related	eventualities	thematic: *agent*, *instrument*, *theme* causal: *cause-of*, spatial: *in*, temporal...

Table 6.4: Classification of bridging anaphora

6.2 Bridging Relations

As we have already seen in the preceding sections, the relationship holding between anaphor and anchor can be of various types. In the literature (e.g. Clark 1977; Kleiber 1997; Gardent et al. 2003), a considerable set of different bridging relations can be found. First of

all, there is a whole group of part-whole relationships, e.g. *set/element* ("class/student"), *set/subset* ("a crowd/ten people"), *whole/part* ("room/ceiling"), *whole/piece* ("cake/slice"). Apart from that, there are relations concerning functional aspects, e.g. *object/attribute* ("car/price"), *stuff/object* ("plastic/bag"). Different from these are thematic relations such as *event/argument* ("murder/victim"). Sometimes, relations involving time and location ("today/the news", "a city/the streets") are listed separately.

Kleiber (1997, 1999) examined the nature of the relation between anchor and the referent of the anaphoric expression. His emphasis lies on what Hawkins (1978) called associative anaphora and he excludes inferential bridging anaphora. He identified five kinds of bridging relations (cf. also Nissim 2001), which are given here together with his French and English examples:

- *meronymic* relations

  (6.25) Il s'abrita sous un vieux tilleul. **Le tronc** était tout craquelé.
  'He looked for shelter under an old sycamore. The trunk was full of cracks.'

- *locative* relations

  (6.26) We drove into the village. **The church** was standing on a hill.

- *actantial* relations

  (6.27) Paul cut himself some bread and put **the knife** on the table.

- *functional* relations

  (6.28) La voiture dérapa. **Le conducteur** s'était assoupi.
  'The car was skidding. The driver had dozen off.'

- *member-collection* relations

  (6.29) Un couple m'a rendu visite hier. **Le mari** était insupportable.
  'A couple visited me yesterday. The husband was unbearable.'

Kleiber points out that many bridging relations cannot be explained by meronymy, unless the notion of part-whole relation is considerably extended in order to account for examples like (6.26) to (6.28). We will get to these in the following sections.

However, meronymic relations are very common bridging relations. In knowledge representation languages used in Artificial Intelligence, meronymy is often expressed by the relation *part-of*. Note that in standard terms, an entity x is a meronym of y if x is a part of y or a member of y. This notion compromises Kleiber's first and last class. To clarify the terminology, I will have a closer look on mereological relations.

6.2.1 Mereological Relations

In standard mereology, part-of relations form a partial order with a sum operator. In the cognitive and philosophical literature, various different tentative taxonomies of mereological relations were proposed. Winston et al. (1987) differentiated six relations: *component/integral object* ("handle/cup"), *member/collection* ("tree/forest"), *portion/mass* ("slice/bread"),*stuff/object*, *feature/activity* ("paying/shopping"), and *place/area*.

Iris et al. (1988) argue that four classes of part-whole relations must be distinguished and present a different taxonomy: *functional component* ("engine/car"), *segment* ("slice/bread"), *membership* ("sheep/flock"), and *subset* ("the husband/a couple").

Vieu and Aurnague (2007) propose a characterization of part-whole relationships based on the theory of plurality of Link (1983). There, a basic distinction is made between atomic entities, which can be singular or collective, and plural entities. A collective atomic entity (e.g. "a forest") is constituted of some plural entity (e.g. "the trees"), which is the sum of several atomic entities. Building on this lattice-theoretic account of plurality, Vieu and Aurnague (2007, p. 486f) distinguish the following main classes:

Member/Collection is a non-transitive relation holding between a singular entity and either a plural entity or a collection, e.g. "a tree/a forest".

Subcollection/Collection is a transitive relation holding between two plural entities or collections constituted by them, e.g. "the pines/a forest"

Portion/Whole is a transitive relation between two amounts of the same substance (e.g. "some water / a glass of water"), or between an amount of a substance and a singular entity which is constituted of an amount of the same substance (e.g. "slice/cake").

Substance/Whole is a transitive relation holding between two amounts of different substances (e.g. "the ice / this glass of whiskey"), or an amount of a substance and a singular entity which is constituted of an amount of a different substance ("flour/cake").

Component/Integral Whole is a relation based on functional dependence between two singular entities, e.g. "engine/car".

Note that this classification is restricted to the ontological categories of material object, amount, and substance, leaving out the categories of time and space, which can exhibit part-whole relationships as well. Temporal and spatial relations are crucially involved in the conception of eventualities, which is subject of the next section. A comparison of the classifications considered so far is given in Table 6.5.

Winston et al.	Iris et al.	Vieu & Aurnague	Kleiber
member/ collection	membership	member/ collection	member/ collection
	subset	subcollection/ collection	meronymic
portion/ mass	segment	portion/ whole	meronymic
stuff/ object		substance/ whole	meronymic
component/ integral object	functional component	component/ integral whole	functional
feature/ activity			actantial
place/ area			locative

Table 6.5: Types of bridging relations

6.2.2 Relations Involving Events and Frames

Most taxonomies of bridging anaphora and relations make out a class of thematic relations. These relations are characterized by a role the entity denoted by an anaphor plays in an eventuality.
Thematic relations express the meaning a noun phrase has with respect to the eventuality denoted by the verb of the sentence. For example, in (6.30), Snow-white is the acting person in the eating event, so she is the *agent*, and the apple is the item that is eaten, so it is the *patient* or *theme*.

(6.30) Snow-white ate a poisonous apple.

Closely associated with thematic relations at the syntax-semantics interface are thematic roles, or *theta-roles*, which are assigned to verbal arguments in a sentence and structurally reflecting the thematic relations associated with them.

The terms thematic relation and theta-role can be found in a broad range of work in different areas of linguistics from Fillmore (1968) to Chomsky (1981). Further surveys on this issue and discussions on their status in natural language semantics are, among others, Jackendoff (1987), Rappaport and Levin (1988), and Parsons (1990).

A multitude of different thematic relations has been proposed. The major thematic relations, which are more or less uncontroversial, are

- *Agent*: deliberately performs the action in question,
- *Experiencer*: receives sensory or emotional input,
- *Theme*: undergoes the action but does not change its state,
- *Patient*: undergoes the action and has its state changed, and
- *Instrument*: used to carry out the action.

Often, no clear boundaries between these relations can be drawn, so, for instance, researchers have different opinions on whether the apple in (6.30) is theme, patient, or undergoer.

Apart from these, often other thematic relations are assumed, e.g. *cause*, *purpose*, *goal*, *source*, *time*, *location*, *manner*, and *beneficiary*. As a consequence, virtually any kind of relationship between concepts, be it causal, temporal, spatial, or whatever, can be assigned a thematic relation. A theory that takes this point of view is Frame Semantics (Fillmore, 1976), which will be introduced in section 7.1.

6.3 Resolution of Bridging Inferences

6.3.1 Pragmatic Accounts

Many pragmaticists argue that the consideration of *Relevance* is most important in the interpretation of bridging anaphora. For instance, Erkü and Gundel (1987) discuss examples like the following:

(6.31) We went to a Thai restaurant. The waitress was from Bangkok.

(6.32) We stopped for drinks at the New York Hilton before going to the Thai restaurant. The waitress was from Bangkok.

According to Erkü and Gundel (1987), "the waitress" in (6.31) is understood as the one in the Thai restaurant. In contrast, in (6.32), "the waitress" would be understood as the one in the Hilton, despite the strong association between Thai and Bangkok. Their criterion for the choice of bridging anchors can be characterized as accepting the first candidate referent that yields an overall interpretation that is relevant. A merit of their work is the point that the information structure (i.e. topic-focus-articulation) of the involved sentences plays an important role in the resolution of bridging references, although many of their predictions have turned out to be untenable. Moreover, they do not give an explicit formulation of Relevance.

Matsui (2000) proposed a more sophisticated account of bridging in terms of Relevance Theory as developed by Sperber and Wilson (1986)[3]. The data used by Wilson and Matsui (1998) and Matsui (2000) is drawn from a questionnaire developed by Matsui. Subjects were asked to answer questions about the most appropriate antecedents when presented short texts containing bridging anaphora with competing anchors, such as the following:

(6.33) I prefer London to Edinburgh. I hate the snowy winters.

In this example, 100% of the subjects chose Edinburgh as the preferred antecedent of "the snowy winters". Matsui argues that accounts which base the choice of antecedents on the accessibility of antecedents fail, because, according to Sidner's focus theory (Sidner, 1981), the most accessible antecedent would be the direct object of "prefer", that is London. Thus, addressees may reject the most accessible antecedent candidate in favour of the most factually plausible one. However, another example (6.34) does not confirm this hypothesis.

(6.34) I prefer the restaurant on the corner to the student canteen. The cappuccino is less expensive.

Here, in 100% of the cases the answer was that the cappuccino was served in the restaurant, although world knowledge tells us that the coffee served in student's establishments is usually less expensive than in ordinary restaurants. Thus, not always the most factually plausible antecedent is chosen. To illustrate the working of the relevance-theoretic mechanism, consider the following example.

(6.35) Sara left Australia for England. She hates the sandy beaches.

Australia is the preferred antecedent in Matsui's study (100%), it is also the most accessible antecedent according to Focus theory, and there is world knowledge telling us that there are many sandy beaches is Australia. The reasoning according to Relevance theory goes as follows: The first utterance raises an implicit question ("Why did Sara leave Australia for England"?) which is expected to be answered by an optimally relevant answer. Wilson and Matsui (1998) cite this example in order to argue against a coherence-based account of bridging, but in fact, assuming an implicit question "Why"? amounts to saying that there is reason to suspect a coherence relation EXPLANATION.

3 See section 1.1.2.3 on page 20 for a short introduction to Relevance Theory.

Relevance theory is very flexible and can account for a considerable range of bridging anaphora. However, it tends to give arbitrary explanations, which can be adapted to account for virtually any kind of example. The inferences that have to be drawn in this type of account rely on expectations and intentions of the discourse participants. Although this kind of reasoning gives many plausible answers for phenomena difficult to explain otherwise, it is questionable whether complex reasoning processes as outlined above always play a decisive role in the resolution of indirect anaphora. Moreover, as Nissim (2001) points out, it is not clear why a speaker would choose an expression involving such a complicated way of communicating over others that are possibly easier for the addressee to interpret. I suspect that reference resolution does not always immediately involve a complex modeling of the conversation participants' intentions and can be explained in terms of the information conveyed by the utterances. However, although the idea to assume only one underlying principle of Relevance, from which all types of pragmatic inferences are to be derivable, is intuitively appealing, a serious drawback for relevance-theoretic accounts of bridging is the lack of a formal implementation of the mechanisms proposed. In the next section, I will turn to formal accounts of bridging resolution.

6.3.2 Computational Accounts

6.3.2.1 Minimal Models for Bridging Anaphora

According to dynamic semantic theories (cf. section 3.4.3 above), definite nominal phrases carry the presupposition of a familiar discourse referent (Heim, 1982). Van der Sandt (1992) suggested that these discourse referents can be found in the process of anaphora resolution. This amounts to trying to bind the presupposition to an antecedent, and, if this fails, accommodate it. Anaphoric (i.e. referential) binding is the case of direct anaphora, as exemplified in (6.36). Accommodation takes place if a brand-new discourse referent is introduced (6.37).

(6.36) A monkey is sleeping. The monkey is dreaming.

(6.37) A monkey is sleeping. The lion is watching it firmly.

But there is a third possibility not accounted for in the above mentioned anaphoric theories of presupposition: bridging or indirect anaphora (6.38).

(6.38) A monkey is dreaming. The tail is twitching.

As repeatedly noted in the literature (e.g. van der Sandt, 1992; Bos et al., 1995; Asher and Lascarides, 1998a; Gardent and Konrad, 2000), there is an empirical preference for binding over bridging and for bridging over accommodation. In the following sections, I will review some formal and computational approaches accounting for this preference. First, I want to illustrate briefly the use of model generation (cf. section 5.2) in bridging anaphora resolution.

Gardent and Konrad (2000) suggested that local minimality captures the preferences for resolving definite descriptions mentioned in the last section. Consider first the preference for direct anaphoric binding over bridging and accommodation. Binding means that the entities denoted by the definite description and its antecedent are identical. Models with

this characteristic have fewer entities in the discourse domain than models which do not assume identity. From the following models for example (6.36), it is easy to see that $\mathcal{M}_a$ is domain minimal.

- $\mathcal{M}_a = \{monkey(c_1), sleep(c_1), dream(c_1)\}$
- $\mathcal{M}_b = \{monkey(c_1), sleep(c_1), monkey(c_2), dream(c_2)\}$

Second, concerning the preference for bridging over accommodation, models involving bridging references need fewer assertions than models that accommodate referents. In order to correctly understand example (6.38), it is crucial to have world knowledge available, namely that a defining property of a monkey is that it has a tail. This property can be expressed as shown in (6.39).

(6.39) $\forall x[monkey(x) \rightarrow \exists y[tail(y) \wedge part\text{-}of(y, x)]]$

Again, by model generation, several models can be constructed:

- $\mathcal{M}_a = \{monkey(c_1), dream(c_1), tail(c_2), part\text{-}of(c_2, c_1), twitch(c_2)\}$
- $\mathcal{M}_b = \{monkey(c_1), dream(c_1), tail(c_2), part\text{-}of(c_2, c_1), tail(c_3), twitch(c_3)\}$

Of these models, $\mathcal{M}_a$ is both domain and subset minimal, hence locally minimal. So far, we get the right interpretation for some simple cases. More difficult are cases where the existence of an antecedent is not entailed by the bridging anchor. We will need to model the influence of defeasible knowledge from various information sources on discourse interpretation.

6.3.2.2 Automated Anaphora Resolution (Freitas, 2005)

Freitas (2005) presents a computational methodology to resolve anaphora by means of abductive reasoning over a semantic discourse representation integrating DRT (Kamp and Reyle, 1993) (cf. section 3.4.3) and a variant of Focus Theory (Sidner, 1981) (cf. section 3.4.2.1).

The problem of anaphora resolution can be represented by the relation $R(A, T)$, where A denotes the anaphor, i.e. an entity introduced by a pronoun, an ellipsis or a definite noun phrase; T denotes its (possibly indirect) antecedent; and R is the relation between A and T.

The goal of anaphora resolution then, in general, can be expressed as follows: given A, find T and R. Freitas proposes four different bridging relations R (p. 62):

1. *coreference*: both A and T denote entities and $A = T$,
2. *member_of*: A denotes an entity, T denotes a set of entities and $A \in T$,
3. *part_of*: both A and T denote entities and $A \sqsubseteq T$,
4. *subcategorized_by*: both A and T denote entities, and the entity denoted by A is a "conceptual part of" the entity denoted by T.

The proposed bridging relations are obtained by virtue of the following pragmatic rules (pp. 61ff.):

Determination of coreference Relations

(i) If A has been introduced into the discourse by virtue of a pronoun or an ellipsis, then R is a relation of *coreference*, e.g. (6.40).

(ii) If A has been introduced by a definite NP and A and T have concordance in number and gender, then R can be a *coreference* relation, e.g. (6.41).

(iii) If A has been introduced by a definite NP and A and T have concordance in number and A or T are qualified collections, then R can be a *coreference* relation, e.g. (6.42).

Rules (ii) and (iii) do not rule out a non-coreference relation in cases of definite NPs. They can be seen as default rules. The rules are illustrated by the following Portuguese examples from Freitas (2005).

(6.40) a. Fernando foi a uma festa.
Fernando.m.sg go.3sg.PAST to a party
'Fernando went to a party.'

b. **Ele** nem me escutou.
he.m.sg not me listen.3sg.PAST
'He didn't listen to me.'

(6.41) a. Fernando foi a uma festa.
Fernando.m.sg go.3sg.PAST to a party
'Fernando went to a party.'

b. **O** **idiota** não me ouviu.
the.m.sg idiot.m.sg not me hear.3sg.PAST
'The idiot didn't hear me.'

(6.42) a. Ontem passou uma matilha por aqui.
yesterday pass.3sg.PAST a.f.sg pack.f.sg for here
'Yesterday, a pack of dogs passed by.'

b. **Os** **cães** mataram cinco galhinas.
the.m.pl dog.m.pl kill.3pl.PAST five hens
'The dogs killed five hens.'

Determination of member_of Relations

Let $\mathbb{T}$ be the "type" of an entity $\mathcal{E}$, which is determined as follows:

(i) if $\mathcal{E}$ is in plural, then $\mathbb{T}$ is a unique set formed by the "linguistic head" of $\mathcal{E}$ in singular,

(ii) if $\mathcal{E}$ is a collective entity, then $\mathbb{T}$ is the set of synonyms of $\mathcal{E}$.

If $\mathbb{T}_A$ is the type of A and $\mathbb{T}_T$ is the type of T and if $\mathbb{T}_A \cup \mathbb{T}_T \neq \emptyset$ and $singular(A)$ and $plural(T)$, then it can be assumed that $A \in T$, i.e. $member\_of(A, T)$.

(6.43) a. Ontem passou uma matilha por aqui.
yesterday pass.3sg.PAST a.f.sg pack.f.sg for here
'Yesterday, a pack of dogs passed by.'

b. **Um cão** revirou a minha lata de lixo.
a.m.sg dog.m.sg rummage.3sg.PAST the my bin of waste
'A dog rummaged my waste bin.'

Note that this rule applies to both definites and indefinite noun phrases, as exemplified by (6.43).

Determination of part_of Relations

If T is in singular and A is not a collective entity (including plural), then a *part_of* relation can be assumed.

The relation $part\_of(A, T)$ is only valid if there is nothing in the context contradicting it. Hence, this rule is a default rule. It can be expressed by an *abnormality* predicate (cf. section 1.2.2): $part\_of(A, T) \wedge \neg abnormal(part\_of(A, T))$.

(6.44) a. Wilson trouxe uma cesta de lanche.
Wilson found.3sg.PAST a.f.sg basket of lunch
'Wilson found a lunch basket.'

b. **A cerveja** estava quente.
the.f.sg beer.m.sg be.3sg.PAST warm
'The beer was warm.'

Determination of subcategorized_by Relations

If T is in singular and A is an **animate** definite and T is not a collective entity ($\neg plural(T)$), a *subcategorized_by* relation can be assumed.

(6.45) a. Un ônibus chegou à rodoviária.
a.m.sg bus arrive.3sg.PAST to-the road
'A bus arrived at the road.'

b. **O motorista** era calvo.
the.m.sg driver be.3sg.PAST bold
'The driver was bold.'

Accommodation

If A is a definite NP and it is not possible to establish any of the former relations between A and the antecedent T, then A is *accommodated*.

In Freitas' system, these rules are implemented as integrity constraints on logic programs. He uses a variant of *Prolog* (cf. Blackburn et al., 2006), extended by integrity constraints[4]. The system follows an abductive reasoning scheme (cf. section 1.2.4) in the line of Hobbs et al. (1993) (cf. section 5.1 above) with abnormality predicates known from Circumscription (cf. section 1.2.2).

Freitas made two experiments in order to evaluate his proposal (p. 83ff.). The first one used the automated process following his methodology, and the second one was made by a human annotator.

Results of the first experiment revealed that in almost half of the data, an *accommodation* relation had to be assumed and that the relation *member_of* had a low frequency.

In the second experiment and its comparison with the human annotation, however, considerable differences can be found. While 87 (16%) of annotated examples were manually classified as *part_of*, only 28 (5,4%) were classified as such by Freitas' system. Much more accommodations were made automatically (50,4%) than manually (37%). Finally, in 73 cases (14%) the relation could not be identified by the human annotator as belonging to any of the groups proposed by Freitas. This was the case when the anaphoric expressions was an event nominalization, e.g. "the arrival" in (6.46).

(6.46) The president's airplane landed at 5. The arrival was...

Interestingly, these cases were not subsumed by Freitas' *subcategorized_by* group. These cases, however, are covered by an approach using FrameNet data. I will sketch such an approach in chapter 7.

Freitas' approach is very clearly formulated and certainly attractive from a computational point of view. Drawbacks include, however, that the determination of *member_of* relations presupposes an unusual definition of "types" which seems to be rather ad hoc. Similarly, establishing *subcategorized_by* relations requires animacy of referents, although it surely can be the case that a non-animate entity is, according to Freitas' definition of *subcategorized_by* relations, "a conceptual part of" another entity. Moreover, it is not clear how the two groups *subcategorized_by* and *part_of* are to be distinguished. Another shortcoming emerges as a consequence of Freitas' emphasis on bridging relations between noun phrases: bridging anaphora where the anchor is an eventuality are excluded from the analysis. Freitas proposes his own theory of discourse structure, which is based on relations between nouns only and neglects relations between eventualities and discourse relations.

6.3.3 Bridging in SDRT

In SDRT (cf. section 5.3 on page 151), bridging inferences are seen as "a byproduct of computing how the current sentence connects to the previous ones in the discourse" (Asher and Lascarides, 1998a). The resolution of bridging anaphora relies on four meta-rules:

1. **If possible use identity.**
2. **Bridges must be plausible.**
3. **Discourse structure determines bridging.**
4. **Maximize discourse coherence.**

4 Integrity constraints are also known in Answer Set Programming (e.g. Simons et al., 2002).

The first rule reflects the empirical preference of resolving anaphora to an identical antecedent. This rule is the preferred rule; if resolution to identity is not possible, then the other rules apply in the indicated order.

The second rule means that world knowledge "specifies certain plausible ways of filling the underspecified parameters in the presupposed material".

The third rule states that if a rhetorical relation between the involved discourse segments gives particular clues for resolving the anaphora, then this information is to be used.

The fourth rule is one of the most basic principles assumed in SDRT. In discourse interpretation, there is a preference for resolving bridging anaphora in a way that maximizes discourse coherence[5].

These meta-rules are the basis of the interpretation of bridging anaphora in SDRT. I will illustrate their working in particular cases later in this section. I advise the reader to keep these rules in mind, because I will come back to them in the next chapter and refine them in section 7.3.

6.3.3.1 Representing Bridging Anaphora

To see more formally how bridging inferences are drawn in SDRT, I will first concentrate on the meaning representation of definite descriptions triggering bridging inferences.

In Russellian tradition (Russell, 1905), the denotation of a definite noun phrase can only be given if it fulfills the conditions of existence and uniqueness. The condition of existence is usually expressed by $\exists x[P(x)]$ and the condition of uniqueness by $\exists x \forall y[P(y) \rightarrow x = y]$. These two conditions together can be written in a short form using the iota operator ι, which maps a set containing only one element to this element. A contextual definition of this operator is given in (6.47), where P is the predicate corresponding to the meaning of the noun and Q is the predicate corresponding to the verb.

(6.47) $Q(\iota x[P(x)]) = \exists x[P(x) \wedge \forall y[P(y) \rightarrow y = x] \wedge Q(x)]$

Using this definition, the semantic representation of the definite article "the" can be given as follows:

(6.48) $\lambda P \lambda Q[Q(\iota x[P(x)])]$

Chierchia (1995, p. 221) extends this notion and includes a contextual parameter B for a bridging relation. He claims that "the P" denotes a P that is related by B to an antecedent a to be specified by context. B restricts the domain and must be included in the uniqueness condition. Thus, the definite article gets the refined representation (6.49), and the meaning of a definite noun phrase "the P" is characterized by (6.50a), which is equivalent to (6.50b).

(6.49) $\lambda P \lambda Q[Q(\iota x[B(a, x) \wedge P(x)])]$

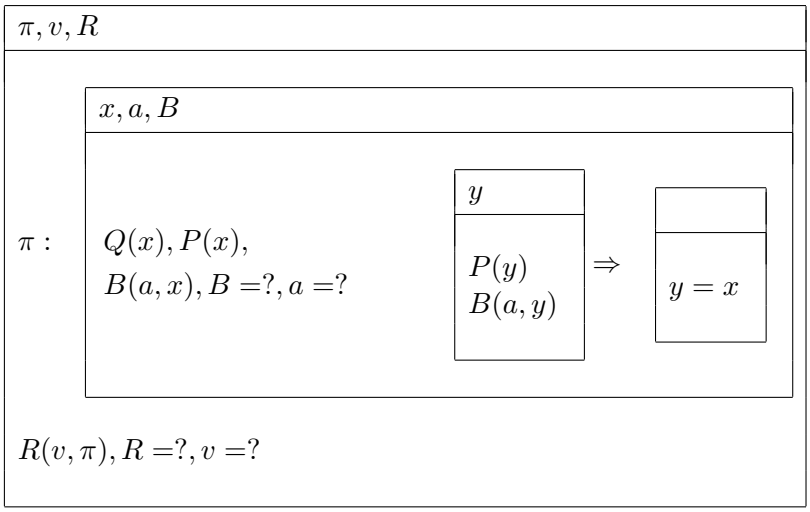
(6.50) a. $\lambda Q[Q(\iota x[B(a, x) \wedge P(x)])]$

b. $\lambda Q[\exists x[B(a, x) \wedge P(x) \wedge \forall y[P(y) \rightarrow y = x] \wedge Q(x)]]$

5 cf. section 5.3.2.5

Building on that, Asher and Lascarides (1998a, p. 87) give a SDRT representation for a definite noun phrase corresponding to (6.50) as shown in (6.51).

(6.51) λQ



The predicates $P(x)$ and $Q(x)$ representing the meanings of the noun and the verb are translated as conditions on a discourse referent x. The bridging relation $B(a, x)$ is represented as an underspecified predicate in the set of SDRS-conditions. The bridging anchor a shows up as an underspecified discourse referent. The condition of uniqueness is now represented by the complex SDRS-condition consisting of the two small SDRSs connected by $\Rightarrow$.

Sometimes, for the sake of clearness of the exposition, I will use a shortcut for representing the uniqueness condition imposed by definites. For a discourse referent x introduced by a definite NP, I will just write $!x$ in the universe, and omit the complex SDRS-condition representing the uniquess condition.

The representation of an indefinite noun phrase triggering a bridging inference would be very similar: we just leave out the uniqueness condition and keep the rest of the conditions. The usual semantic representation for the indefinite article "a" is (6.52). Extending this formula by a bridging relation, we get (6.53).

(6.52) $\lambda P \lambda Q \exists x[P(x) \wedge Q(x)]$

(6.53) $\lambda P \lambda Q \exists x[B(a, x) \wedge P(x) \wedge Q(x)]$

In the SDRS (6.51), two underspecifications are to be specified (or, in other words, presuppositions to be resolved[6]) by pragmatic inference:

Firstly, a coherence relation $R(v, \pi)$ has to be established. According to Asher and Lascarides (1998b), a definite description triggers a coherence relation between the current utterance π and some previous utterance v.

In SDRT, there is no special mechanism for presupposition *accommodation*. Instead, it is modelled by linking a presupposition via the BACKGROUND relation to the context. This

6 The same SDRS can be represented in a different manner in the presuppositional variant of SDRT advocated in Asher and Lascarides (1998b) in the spirit of van der Sandt (1992). Assuming that a sentence (viz. an utterance), in general, yields two labelled SDRSs, one for the asserted information and one for the presupposed information, a definite NP is expressed by the following two SDRSs:

is supposed to be the default way of resolving R in absence of additional information. The semantics of the BACKGROUND relation provides constraints on accommodation.

Secondly, in the bridging relation $B(a, x)$, the parameters B for the bridging relation and a for the bridging anchor have to be specified (Asher and Lascarides, 1998a).

The specification of B can be related to the taxonomy proposed in section 6.1.4. For direct anaphora, B is *identity*. For mereological indirect anaphora (Clark's indirect reference by association), B can be *part-of* or *member-of*. In their article on bridging, Asher and Lascarides (1998a) focus on mereological bridging relations. However, the account of bridging in SDRT can be straightforwardly extended in order to carry over to the other types of bridging relations. For frame-related indirect anaphora (Clark's indirect reference by characterization), B is either a thematic role, e.g. *agent*, *theme*, or *instrument*, or an instance of other conceptual relations such as *cause*, *place*, etc.

6.3.3.2 Resolving Bridging Anaphora

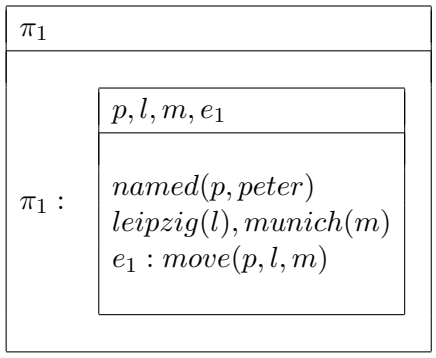
Let us illustrate the working of the SDRT mechanism for bridging resolution with a simple example (6.54).

(6.54) a. Peter moved from Leipzig to Munich.

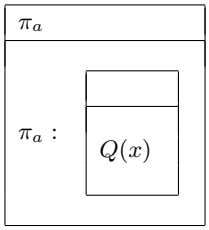
b. The rent was lower.

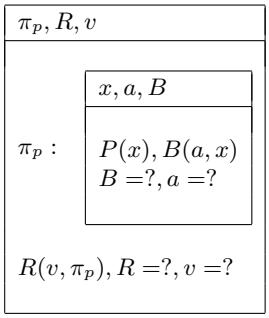
As meaning representations of the sentences in this short text we can assume underspecified SDRSs according to the standard DRS construction rules and the representation scheme (6.51) for definite descriptions. For the sake of a clear exposition, the meaning representation of the comparative "lower" is simplified as a one-place predicate $\lambda x.low(x)$.

(6.55) a.



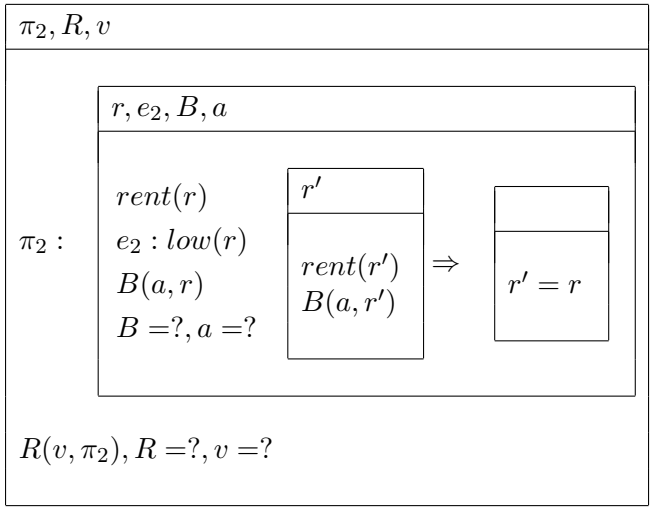
(i) $\lambda Q.$





However, to keep the representations more readable, I stick to the variant proposed in Asher and Lascarides (1998a), spelling out the uniqueness condition and tacitly treating both $R =?$ and $B =?$ as presuppositions to be resolved. In fact, whenever a SDRS-condition contains a question mark, it is a presupposition.

b.



The first meta-rule for bridging – resolve to identity – cannot be applied because there is no available antecedent that could be coreferential to the referent r. For the second rule – plausibility – we have to consider world knowledge. On the one hand, it is well known in Germany that rents are, at least at the moment, much less expensive in Leipzig than in Munich, which is known to be among the most expensive places. This knowledge can roughly be represented by the following two default rules.

(6.56) a. $(rent(r) \wedge leipzig(l) \wedge in(l, r)) > low(r)$

b. $(rent(r) \wedge munich(m) \wedge in(m, r)) > \neg low(r)$

On the other hand, it is common world knowledge that a lower rent is a cause to move. We can express this by another default rule:

(6.57) $(rent(r) \wedge in(r, b) \wedge e_\beta : low(r) \wedge e_\alpha : move(x, a, b)) > cause_D(\beta, \alpha)$

Thus, world knowlegde provides clues for both readings of the text, that the rent was either in Leipzig or in Munich, giving rise to conflicting interpretations. We thus proceed to the third meta-rule – discourse structure deterimines bridging. There are (at least) two discourse relations which are plausible to hold. In terms of the glue logic[7], they can be derived by the following defaults.

Background $(?(\alpha, \beta, \lambda) \wedge event(e_\alpha) \wedge state(e_\beta)) > Background(\alpha, \beta, \lambda)$

Explanation $(?(\alpha, \beta, \lambda) \wedge cause_D(\beta, \alpha)) > Explanation(\alpha, \beta, \lambda)$

Again, we have conflicting default rules. On the one hand, since the second sentence describes a state, we can derive by default a Background relation. On the other hand, the world knowledge rule (6.57) gives rise to the default for inferring Explanation[8].

A way out of this conflict is to impose an ordering $\succ_\tau$ on coherence relations based on their "quality". Background is then supposed to be weak in the sense that it conveys only little thematic continuity:

7 Remember from section 5.3.2.2 on pages 155ff. that the glue logic term $?(\alpha, \beta, \lambda)$ corresponds to $\lambda : R(\alpha, \beta) \wedge R =?$ in the SDRS language.

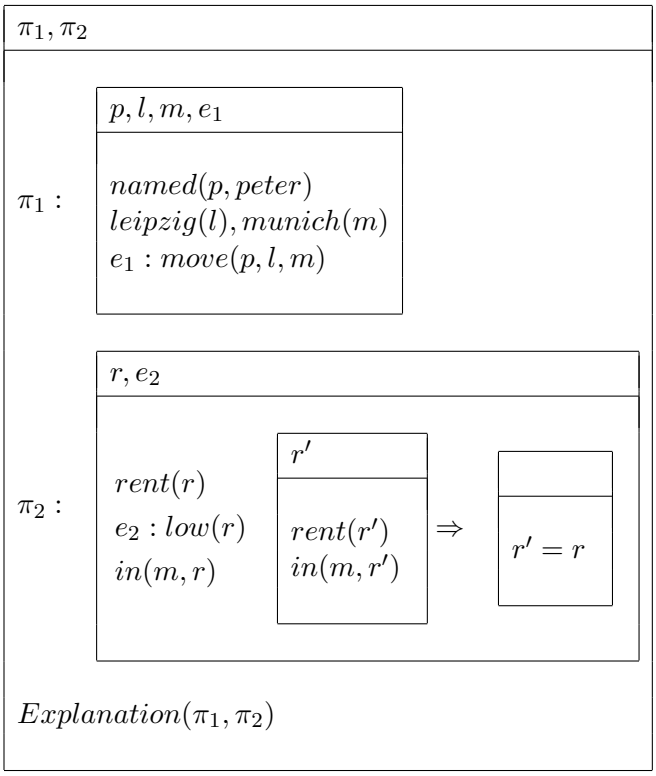
8 $cause_D(\beta, \alpha)$ means that there is evidence in the discourse that β caused α.

(6.58) $Explanation(\alpha, \beta, \lambda) \succ_\tau Background(\alpha, \beta, \lambda)$

The last meta-rule – maximize discourse coherence (MDC) – tells us that "better" relations are to be preferred. So we can specify R as EXPLANATION and resolve B to *in* and a to m. With this, we derive the reading in which it is the rent in Munich that is referred to in the second sentence – contra world knowledge.

For illustration, a pragmatically enriched SDRS for text (6.54) in its preferred reading is the following:

(6.59)



A similar example (6.34), repeated here, was brought into the discussion by Wilson and Matsui (1998) for supporting a relevance-theoretic account of bridging.

(6.34) I prefer the restaurant on the corner to the student canteen. The cappuccino is less expensive.

We can derive the preferred reading, in which the cappucino is taken to be served in the restaurant, by means of a coherence-based account of bridging. MDC guarantees that this interpretation, where (b) is an EXPLANATION for (a), is preferred over an interpretation in which (b) gives BACKGROUND information for (a).

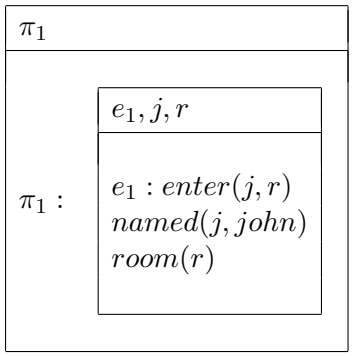
6.3.3.3 Extending SDRT by Equality by Default

Consider now Clark's classical example:

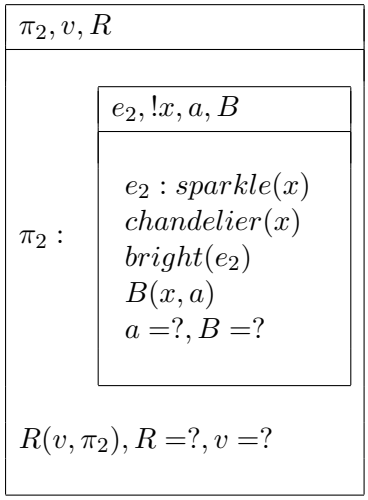
(6.60) John entered a room. The chandelier sparkled brightly.

The content of the sentences in (6.60) can be represented as underspecified SDRSes as follows:

(6.61) a.



b.



Let us try to account for this example in terms of Asher & Lascarides' meta-rules. As usual in indirect anaphora, the first meta-rule cannot by applied since there is no suitable antecedent that could be identified with the referent of "the chandelier".

In order to properly interpret this short discourse, some world knowledge is needed. We will assume that a "room" by default involves the existence of a source of light. For ease of exposition, let us suppose that we have the following Default Logic rule in our knowledge base[9].

(6.62) $$\frac{room(r) : light(l), part-of(l, r)}{light(l), part-of(l, r)}$$

This default can be expressed shorter by means of the nonmonotonic conditional operator used in SDRT's logic Commonsense Entailment:

(6.63) $room(r) > light(l) \wedge part-of(l, r)$

World knowledge also tells us that a "chandelier", unless it is not broken, is a source of light. Thus, we assume the following default, again both in Default Logic (6.64) and in Commonsense Entailment (6.65) notations.

9 A more sophisticated way of integrating essential parts of world knowledge into the semantic representation will be presented in chapter 7.

(6.64) $$\frac{chandelier(x) : \neg broken(x)}{light(x)}$$

(6.65) $chandelier(x) \wedge \neg broken(x) > light(x)$

We can shorten this rule further by assuming that the property of not being broken is part of the normality of chandeliers, which is expressed by the conditional operator:

(6.66) $chandelier(x) > light(x)$

With these pieces of world knowledge, the underspecificied representations (6.61) can be further specified. First, applying (6.63) leads to enriching the SDRS for the first utterance by a referent l for the light source and the predicates $light(l)$ and $part-of(l, r)$. Second, rule (6.66) allows a condition $light(x)$ in the second SDRS.

The meta-rules for bridging, as they are stated, are not sufficient to resolve the bridging relation, because a link between x and l has to be assumed, and neither the discourse structure nor the MDC can provide clues for that.

However, there is a rather straightforward way to solve this problem. In section 5.2.4.2 (page 148), I have introduced Cohen (2006)'s notion of Equality by Default. It consisted in assuming a general, low-priority default (5.33), repeated here as (6.67), or equivalently in terms of SDRT's Commonsense Entailment as in (6.68).

(6.67) $$\frac{: x = y}{x = y}$$

(6.68) $> x = y$

As this default, stated in this way, does not have any precondition, it is likely to overgenerate. A first attempt to adapt it to SDRT is the following:

Equality by Default

$\pi_i : \boxed{x} \wedge \pi_j : \boxed{y} \wedge (x \text{ is available in } \pi_j) > x = y$

SDRT's definition of availability of antecedents for anaphora was given in section 5.3.2.4. It says in essence that an antecedent for an anaphoric expression must be DRS-accessible on the right frontier.

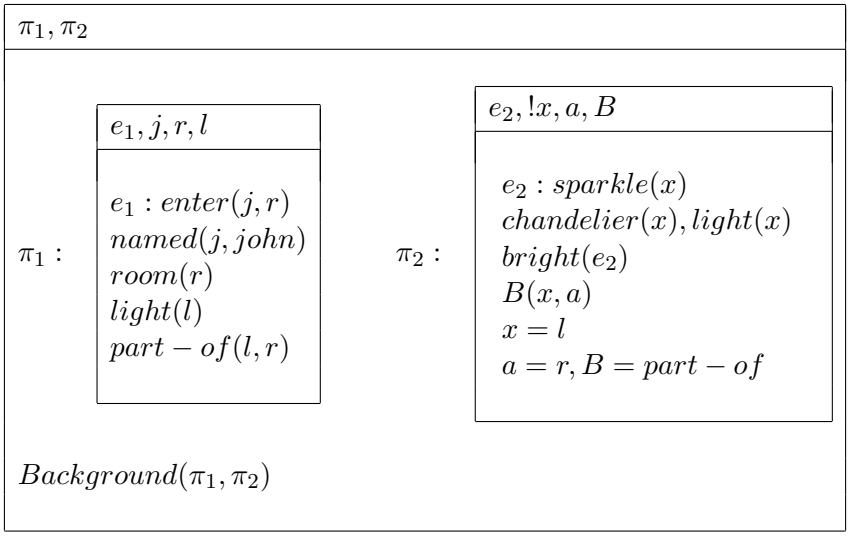
This default provides a way to achieve minimality of discourse representations. Consequently, since a minimal model for the example discourse seeks to minimize the number of discourse referents, this rule permits to assume the equivalences $x = l$ and $a = r$. Furthermore, local minimality includes a minimal number of DRS-conditions, which allows us to establish the equivalence $B = part-of$. A model which does not establish these equivalences is obviously not minimal.

We are left with connecting the second utterance to the preceding discourse by virtue of specifying the coherence relation R. For establishing this relation, let us have a look at the eventualities expressed in the text. Entering a room is an achievement (in the terminology of Vendler, 1957) and thus a proper event. To sparkle is, in a strict sense, an activity. However, as it is a durative instance of emitting light, it can be seen as a special case of a state (in a broad sense). Assuming this, we are able to apply SDRT's default rule for inferring a Background relation (cf. page 183 above). Alternatively, we can simply take

for granted that presuppositions connected with definite descriptions tend to bind with BACKGROUND by default.

A pragmatically enriched SDRS for (6.60), where all underspecifications are resolved, is given in (6.69):

(6.69)



A few remarks are in order at this point. SDRT provides an exact formulation of inferences in discourse interpretation involving knowledge from various information sources. Nevertheless, world knowledge has to be encoded in a suitable fashion in order to match the antecedents of glue logic defaults. It seems rather difficult to straightforwardly extend the approach to arbitrary non-domain-specific naturally-occuring texts. However, there are ways to express (at least parts of) domain and world knowledge in a constrained manner. One possibility makes use of an extended view of lexical knowledge. Another possibility is to use frame-based representations of stereotypical knowledge. In the next chapter, I will present the latter approach (section 7.2) and discuss the former approach (section 7.4.2).

6.4 Conclusion

This chapter was devoted to bridging inferences. Based on existing corpus-based and psycholinguistic studies, I have made a classification of bridging anaphora. In essence, two types of bridging relations can be distinguished: mereological and frame-related relations. Bridging is a challenge for accounts of anaphora resolution. While pragmatic accounts can explain the reasoning behind bridging in terms of intentions of discourse participants, they lack a precise formulation of these inferences. Computational and formal accounts can formally express these inferences with the help of nonmonotonic reasoning: minimal model generation can account for the preferences in anaphora interpretation and automated anaphora resolution can be constrained by suitable default rules. The approach to bridging in SDRT provides the most extensive basis for drawing bridging inferences in an interplay of grammatical and contextual knowledge. It explains how bridging anaphora attach to the existing discourse structure and it provides a basis for integrating world knowledge. Extended by the consideration of minimal models, it can handle a broad range of bridging inferences. In the last two chapters, SDRT will be the underlying model of discourse

interpretation. In chapter 7, I will present a way of integrating knowledge of frames and scenarios into the interpretation. In chapter 8, the proposed account is applied to the interpretation of clitic left dislocated noun phrases in Spanish.

Chapter 7

Bridges Between Events

This chapter sketches a new approach to bridging anaphora, which is particularly aimed at covering not only mereological but also frame-related bridging relations. The proposed account is based on two keystones. First, SDRT (Asher and Lascarides, 2003) will be assumed as the underlying theory of discourse structure and interpretation. This theory, which is formalized in many important details, is the most comprehensive theory of discourse interpretation known so far. Second, I will exploit the idea developed in Frame Semantics (Fillmore, 1976) that world knowledge is organized in frames. This framework provides a suitable cognitive complement to the formal tools provided by SDRT.

In section 7.1, the basic ideas of Frame Semantics are introduced. Essentially, with each eventuality introduced in a discourse, a corresponding frame is evoked in the discourse model. In section 7.2, I will extend the discourse representation of SDRT by including possibly underspecified representations of frame elements, which can give clues for finding suitable antecedents of bridging anaphora. In section 7.3, some general constraints on bridging inferences are determined. Finally, in section 7.4, I will compare the presented framework to related approaches.

7.1 Frame Semantics and FrameNet

7.1.1 Frame Semantics

To get clues for the resolution of bridging relations involving eventualities, I propose to exploit Frame Semantics and subsequent work on FrameNet (Baker et al., 1998; Fillmore et al., 2003; Ruppenhofer et al., 2005).

Frame semantics is based on the central assumption that world knowledge is organized in *frames*. Frames are mental representations of stereotypical situations, whose elements can only be defined relating one to another. The notion of frames goes back to earlier work in Artificial Intelligence (Minsky, 1975), where a frame refers to a data structure representing a stereotyped situation. A similar term, *script*, emerged in cognitive psychology (Schank and Abelson, 1977), referring to knowledge structures for sequences of events, e.g. the restaurant script.

A typical example for a frame is the *Commercial_transaction* frame. The basic idea of frame semantics is that one is not able to understand the word "sell" without knowing anything about the situation of commercial transaction, which involves, among other things, a seller, a buyer, goods, money, and relations between them. Crucially, the meaning of a single word is not defined by single elements of a frame, but by a particular perspectivation of a frame. From the perspective of a buyer, a commercial transaction is referred to by the verb "buy", and from the perspective of a seller by the verb "sell". Thus, the

Commercial_transaction frame enables one not only to understand single words but also to understand similarities and differences between semantically related words.

In short, a text activates or evokes a frame when a linguistic form is conventionally associated with that frame.

Frame semantics has a wide range of applications in various subfields of linguistics and related disciplines. The central and most successful application is in lexicography. In a frame-based lexicon, the frame accounts for related senses of a single word and its semantic relations to other words. Such a lexicon offers more comprehensive information than the traditional lexicon. Consequently, over the last decades, Fillmore and his colleagues in Berkeley developed FrameNet.

7.1.2 FrameNet

FrameNet (Baker et al., 1998; Ruppenhofer et al., 2005) is a lexical resource providing a body of annotated sentences based on frame semantics. At the moment, the database contains around 10,000 lexical units, 800 semantic frames, and over 120,000 example sentences. The original FrameNet is developed for the English language. Furthermore, in the last years, it was ported to other languages such as German, Japanese, and Spanish.

In the following, I will introduce and try to briefly define the essential concepts of FrameNet as far as needed for the purposes of this thesis.

Frame A frame is a structure $\langle N, D, FE, LU \rangle$ consisting of a name N, a definition D, a set of frame elements FE, and a set of lexical units LU.

The name is just an identifier of the frame, and the definition informally describes the concept the frame stands for. Before turning to frame elements, let us define the set of lexical units evoking a frame.

7.1.2.1 Lexical Units

A *lexical unit* is a pairing of a linguistic expression with a frame. Every lexical unit *evokes* a particular frame and can only be understood in relation to that frame. More formally,

Lexical Unit Let Φ be a frame and V a linguistic expression that is potentially frame-evoking, i.e. a verb, a noun, or an adjective[1], so

$V \in LU(\Phi)$ iff V evokes Φ.

For example, the frame *Cooking_creation* is evoked by the lexical units "bake", "concoct", "cook up", "cook", "make", "prepare", "put together", and "whip up".

The set of lexical units of a frame may be empty, as well. This is the case of abstract frames that correspond to larger scripts or scenarios, e.g. the frames *Commercial_transaction* or *Crime_scenario*. These frames are used for structuring the FrameNet resource.

An important question is whether a particular linguistic expression evokes at most one frame, exactly one frame, or more than one frame. Polysemous words are represented by several lexical units: "the separate senses of the word correspond to the different (sets of)

1 Lexical units can also consist of several words, e.g. in the case of phrasal verbs such as "take a bath".

frames that the word can participate in. When a word's sense is based on a particular frame, the word evokes the frame" (Fillmore et al., 2003, p. 236). For example, the verb "break" can evoke, among others, the frame *Experience_bodily_harm* (e.g. in "I broke my leg") or the frame *Render_nonfunctional* (in "I guess I broke the doorknob"). Thus, the interpretation of a text requires assumptions about which frame is relevant in the given context. Take the verb "eat": it could be associated with a set of frames, e.g. a restaurant frame, a family home frame, etc. The question is how the right frame ends up being selected. In the spirit of underspecified semantics as outlined in the first chapter of this thesis, I would suggest to choose the most general frame fitting in the given context. For "eating" this would be the frame *Ingestion*. Due to the hierarchical structure of FrameNet (see below), any frame involving eating would inherit the properties and frame elements of this frame.

7.1.2.2 Frame Elements

A frame consists of various *Frame Elements* (FEs), kinds of entities that can participate in a frame. They are defined in relation to a frame. There is a differentiation of two types of frame elements: core FEs and non-core FEs. Core FEs stand for "conceptually necessary components of a frame, while making the frame unique and different from other frames" (Ruppenhofer et al., 2005, p. 26). They correspond roughly to thematic roles in an eventuality. Non-core FEs include peripheral FEs and extra-thematic FEs. Peripheral FEs, e.g. *Place*, *Time*, *Manner*, *Means*, *Degree* can occur in any frame in which these FEs are semantically appropriate. Extra-thematic FEs introduce additional events that do not conceptually belong to the frames they appear in. They are used, for example, to express causal connections between frames. In case that FEs that are normally considered as non-core FEs are conceptually necessary for defining a frame, they can become core FEs.

For illustration[2], the *Killing* frame is described in Fig. 7.1[3], and one of the lexical units evoking that frame, the verb "murder", is characterized in Fig. 7.2. Here, the numbers in parentheses refer to the number of annotated sentences in the database containing the target, viz. the lexical unit, *murder.v* and showing the respective pattern.

Sometimes, conceptually necessary frame elements, i.e. core FEs, do not show up in a sentence. Theses cases are called *Null Instantiations* and may be of one of three types: *Constructional Null Instantiations* CNI), e.g. omitted agents in passive sentences, *Indefinite Null Instantiations* (INI), i.e. implicit arguments of certain transitive verbs that are used intransitively, e.g. verbs as "eat" or "bake", and *Definite Null Instantiations* (DNI), i.e. missing obligatory elements that can be inferred from the context.

As can be seen in Fig. 7.2, there are three cases among the 23 annotated sentences in the FrameNet database containing the lexical unit *murder.v* in which the *Killer* was not expressed at all, i.e. a CNI, and the *Victim* showed up as external argument of the verb. This configuration is typical for passive sentences like (7.1) and (7.2), where the role *Killer* is not filled by a lexical item.

(7.1) It informed him in letters an inch high that [$_{Victim}$Captain Peter Dawson] had been

2 Frame descriptions are taken from the FrameNet Database, obtainable from the International Computer Science Institute, Berkeley, California (http://framenet.icsi.berkeley.edu/).

3 Inheritance and sorts of FEs are subject of the next section.

Definition: A Killer or Cause causes the death of the Victim.

Core Frame Elements :

FE	description	inherited FE	sort
Killer	The person or sentient entity that causes the death of the Victim	Agent	sentient
Victim	The living entity that dies as a result of the killing	Patient	sentient
Instrument	The device used by the Killer to bring about the death of the Victim	Instrument	physical entity
Cause	An inanimate entity or process that causes the death of the Victim	Cause	
Means	The method or action that the Killer or Cause performs resulting in the death of the Victim	Means	state of affairs

Non-Core Frame Elements: Beneficiary, Manner, Place, Purpose, Time, ...

Lexical Units: annihilate.v, annihilation.n, ..., murder.n, murder.v, murderer.n, ..., terminate.v

Figure 7.1: The *Killing* frame

murdered$_{Target}$ in Cyprus [$_{Killer}$CNI]. (FrameNet, sentence 1761950)

(7.2) [$_{Victim}$John] was murdered$_{Target}$ yesterday [$_{Killer}$CNI].

In one annotated occurrence (7.3), the *Victim* is not expressed. It is an instance of an INI: the victim is existencially bound within the discourse model but left unspecified.

(7.3) He had robbed, [$_{Killer}$he] had murdered$_{Target}$! [$_{Victim}$INI] (FrameNet, sentence 1762446)

In contrast, in a DNI the omitted referent must be specified. In (7.4), the referent of the omitted *Goal* is a location that must be accessible to the discourse participants.

(7.4) [$_{Theme}$A tray of coffee] arrived$_{Target}$ almost immediately, together with an enormous plate of calorie-laden cakes [$_{Goal}$DNI]. (FrameNet, sentence 999633)

Null instantiations are of particular interest for this thesis because they are potential bridging anchors. Consider, for example, text (7.5) from (Clark, 1975), where the role of the murderer is not expressed in (a) but later filled in (b).

(7.5) a. John was murdered yesterday.

b. The murderer got away.

Lexical Entry: murder.v

- Frame elements and their syntactic realizations

Killer	CNI.– (3), NP.Ext (15), PP[by].Dep (5)
Victim	NP.Ext (8), INI.– (1), NP.Obj(14)

- Frame elements and valence patterns

frame element	realized as			
Killer	NP.Ext	NP.Ext	PP[by].Dep	CNI.–
Victim	NP.Obj	INI.–	NP.Ext	NP.Ext
(23)	(14)	(1)	(5)	(3)

Figure 7.2: Lexical entry *murder.v*

7.1.2.3 Relations between Frames

Frame Inheritance

Frames in FrameNet are hierarchically organized: e.g. the frame *Killing* inherits the properties from the more general frame *Transitive_action* which in turn inherits from the abstract frame *Event*. Two types of frame inheritance are assumed by (Baker et al., 2003): full and monotonic inheritance, which I will define as follows:

Full Inheritance Let A and B be frames, so

$$inherits\_full(A, B) \text{ iff } \forall\psi\exists\phi[\psi \in FE(B) \rightarrow \phi \in FE(A)].$$

For $inherits\_full(A, B)$ to hold, the names of corresponding frame elements ϕ and ψ can be different, but there must be some kind of binding between them. The child frame A can have additional frame elements that are not part of B. The coreness status of FEs in child and parent frame can be different, as well. The relation $inherits\_full$ has the following property: if A inherits from B then every instance of A is an instance of B. Note that *multiple inheritance* is explicitly allowed: a child frame can have multiple parents. For illustration, consider the inheritance relations centered around the *Killing* frame as depicted in Fig. 7.3[4].

Monotonic Inheritance Let α and β be frames or frame elements, so

$$inherits\_monotonic(\alpha, \beta) \text{ iff } \alpha \sqsubseteq \beta.$$

Monotonic inheritance makes use of a sort[5] hierarchy based on mereological relations like $\sqsubseteq$ (subtype). It is designed to allow distinguishing different sorts of entities ("semantic

4 This representation is provided by FrameNet's FrameGrapher application, which has been created by Christine Adell Hodges and Carol Hays and is available online under http://framenet.icsi.berkeley.edu/FrameGrapher/.

5 In FrameNet, ontological sorts are called "semantic types". I will avoid this term and stick to the term "sort" in order to avoid confusion with what is usually understood as a semantic type, viz. entities e, truth-values t, relations from entities to truth-values $\langle e, t\rangle$ etc.

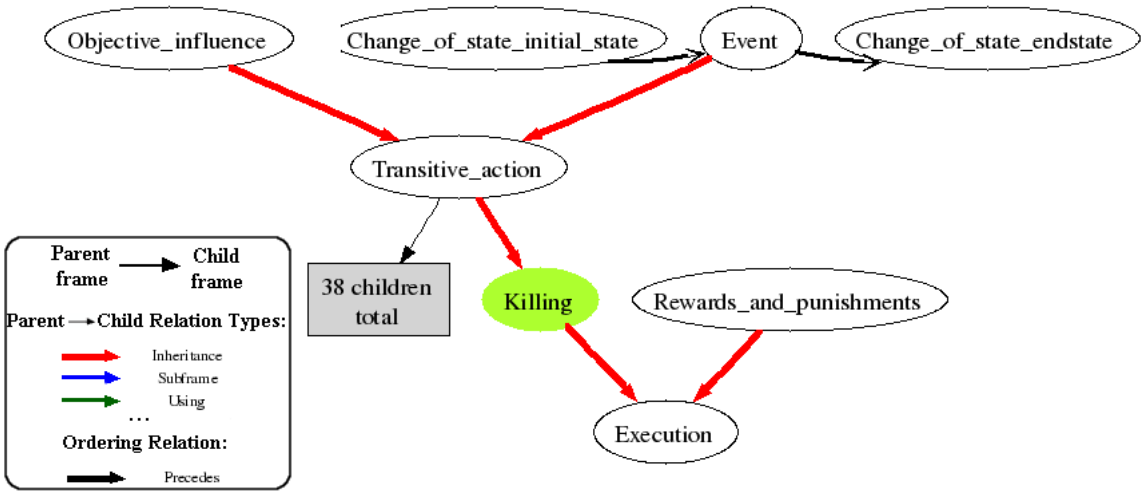


Figure 7.3: Frame inheritance

types") and relations between them, e.g. "human" as a subtype of "animate". The sort hierarchy applies to frames, frame elements, and lexical units: $\alpha \sqsubseteq \beta$ holds iff the sort of α is a subtype of sort β. Note that the sort of a frame element categorizes the sort of its filler, not of the role itself. The development of a fine-grained sort hierarchy for FrameNet is still work in progress.

Uses

The relation *uses* between frames is a kind of weak or partial inheritance. If frame A *uses* frame B then frame A has FEs corresponding to FEs of B but need not have an FE corresponding to each FE of B. Used FE must have compatible sorts. A relation $uses(A, B)$ has the following consequence.

Uses If A, B are frames and $uses(A, B)$ then

$$\exists\phi\exists\psi[\phi \in FE(A) \wedge \psi \in FE(B) \wedge inherits\_monotonic(\phi, \psi)].$$

For example, the frame *Food* uses the frame *Ingestion* implying an inheritance relation between the FEs *Food* in the former and *Ingestibles* in the latter frame. However, in most cases more than one FE is used. In general, $Uses(A, B)$ can be seen as the set of used relations between FEs, viz. a two-place function from a frame A and a frame B to a subset of $FE(A) \times FE(B)$.

Partial Inheritance If A and B are frames and $uses(A, B)$, so

$$Uses(A, B) = U \subseteq FE(A) \times FE(B) = \{\langle\phi, \psi\rangle | \phi \in FE(A) \wedge \psi \in FE(B)\}.$$

For instance, $Uses(Food, Ingestion) = \{\langle Food, Ingestibles\rangle\}$.

Subframes

Many frames express complex events which may be constituted by several subevents. In these cases, the subevents are expressed by frames which are *subframes* of the complex

frame. For instance, the frame *Crime_scenario* has the subframes *Committing_crime*, *Criminal_investigation*, and *Criminal_process*. Subframes can contain FEs that are bound to FEs in the complex frame, e.g. the subframe *Committing_crime* has an FE *Perpetrator*, which is bound to the FE *Perpetrator* in the complex frame *Crime_scenario*. In contrast to inheritance, not all FEs of the superframe must necessarily show up in the subframe.

Often, subframe relations stand in a particular temporal ordering, which is specified by a separate precedence relation, *precedes*. This relation "specifies the sequence of states and events that are definitional for a certain state-of-affairs" (Ruppenhofer et al., 2005, p. 109). For example, in the *Crime_scenario*, the frame *Committing_crime* precedes *Criminal_investigation*, which in turn precedes a *Criminal_process*. In Fig. 7.3, there are precedence relations between *Change_of_state_initial_state*, *Event*, and *Change_of_state_endstate*, indicated by black horizontal arrows.

Note that if a frame inherits from another frame via full inheritance, then the subframe structure of the parent frame is inherited, as well. Since multiple inheritance is allowed, binding of FEs and subframes can get quite complex. Some details concerning this issue can be found in the recent literature, e.g. in the FrameNet book (Ruppenhofer et al., 2005); the elaboration of exact definitions is still under consideration and beyond the scope of this thesis. However, two things can be said about subframes: if A is a subframe of B then an instance of A implies (i) the existence of an instance of B, and (ii) that all subframes of B which precede A are instantiated. I will write "$e : \Phi$" for "e instantiates Φ", where e denotes an eventuality[6].

Subframe If A and B are frames and $subframe(A, B)$ and $e_1 : A$, then

$$\exists e_2[e_2 : B], \text{ and}$$
$$\forall A' \exists e'[subframe(A', B) \wedge precedes(A', A) \rightarrow e' : A'].$$

Other Relations between Frames

Baker et al. (2003) mention another relation between frames, *see_also*, which, however, is only informally defined as "a pointer from one or more frames to another frame, with a definition that includes a detailed discussion of the differences among the frames in the group" (Baker et al., 2003, p. 287). This relation is mainly used for simplifying the manual annotation of sentences.

Recent refinements of FrameNet include additional information concerning perspectives on frames and causal relations between frames. These relations, $perspective\_on$, $causative\_of$, and $inchoative\_of$, are still in development and I will not treat them in this thesis.

7.1.2.4 Relations between Frame Elements

Apart from relations between frames as a whole, there are also relations between frame elements. These can be relations between FEs in a single frame or relations between FEs in different frames.

Within a frame, FEs can be grouped together in *coreness sets*. For example, *Source*, *Path*, and *Goal* are core FEs in the *Motion* frame. It is not necessary, and in fact not usual, that all three FEs are expressed in one sentence. Moreover, sometimes the occurrence of a

6 See the definition of Frame Evocation in section 7.2.2.

core FE *requires* another FE to be instantiated, and sometimes a particular FE *excludes* that another FE of the same group co-occurs in the same sentence.

More interesting for this thesis are relations between FEs in different frames, especially inheritance relations between corresponding FEs in inherited frames. Full inheritance of frames implies that corresponding FEs are inherited, too. The partial inheritance relation *uses* also involves inheritance of FEs. Taking these two cases into account, the relation *inherits*, which is the FE-to-FE relation corresponding to inheritance between frames, can be defined as given below. The definition of its transitive closure $inherits^*$ is straightforward.

Frame Element Inheritance Let A, B, and C be frames and $\phi \in FE(A)$ and $\psi \in FE(B)$, so

1. $inherits(\phi, \psi)$ iff
   a. $inherits\_full(A, B)$, or
   b. $\langle \phi, \psi \rangle \in Uses(A, B)$.
2. $inherits^*(\phi, \psi)$ iff
   a. $inherits(\phi, \psi)$, or
   b. $\exists \chi[\chi \in FE(C) \wedge inherits^*(\phi, \chi) \wedge inherits(\chi, \psi)]$.

In short, a FE ϕ in a frame A inherits from a FE ψ in another frame B if there is either full inheritance between A and B, or if A uses B and ϕ is used as ψ. Inheritance relations between FEs in the frames *Transitive_Action* and *Killing* are illustrated in Fig. 7.4, which is an expansion of Fig. 7.3.

7.2 Building Bridges using FrameNet and SDRT

7.2.1 Integrating FrameNet and SDRT

Each eventuality introduced in a discourse evokes a corresponding frame in the discourse model. Its frame elements correspond to all relevant (necessary or optional) thematic roles of the eventuality. I propose to include for all core frame elements a representation in the discourse model, i.e. in the SDRS of the current utterance.

In case that some participant of a frame is not expressed linguistically, its representation remains underspecified. These elements can be further specified by subsequent information, provided that the discourse referent for the eventuality remains accessible for anaphoric reference. I will spell out in more detail how this works in section 7.3. Before that, I will discuss how frame elements can be represented in SDRT and how they help to determine discourse relations. The focus will lie on frames evoked by verbs, which introduce discourse referents for eventualities. Of course, in (S)DRT, also noun phrases introduce discourse referents. As they are also lexical units, they evoke frames with roles to be reflected in the discourse model. To keep things simple, I assume here that the frame evoked by the main verb in a sentence is the dominant frame, while frames evoked by other lexical units provide additional information. However, the questions of integrating different frames within a single utterance in particular and of compositionality in frame semantics in general deserve more attention, but that is beyond the scope of this thesis. Some issues concerning these questions were discussed by Nissim et al. (1999).

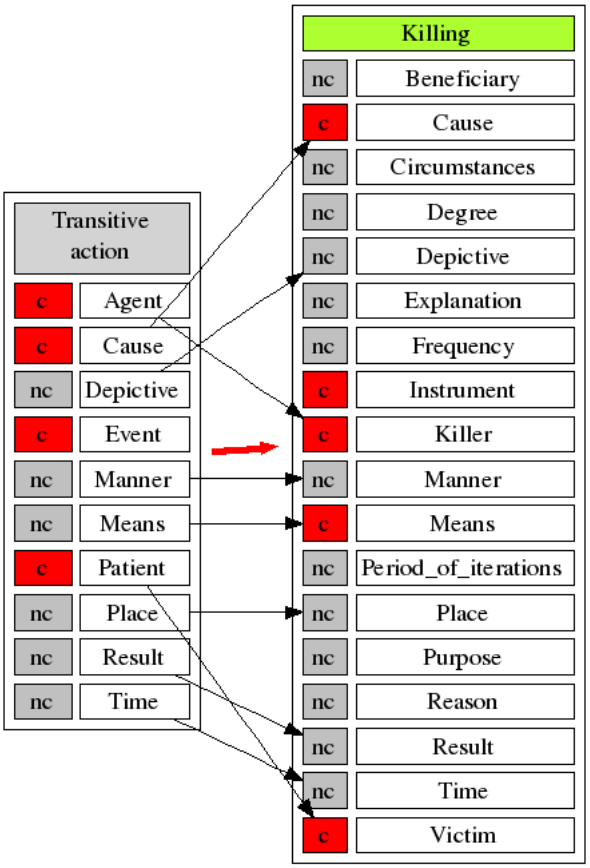


Figure 7.4: Frame element inheritance

7.2.2 Representing Frame Elements in SDRT

In order to integrate FrameNet data in SDRT, a neo-Davidsonian style of event semantics (Parsons, 1990) is adopted, assuming that verbs include an implicit eventive argument in their semantic representation. Thematic roles in an eventuality are represented as conditions in form of predicates with a first argument which is this eventive argument. For instance, the sentence "John eats an apple" gets a semantic representation $\exists e, j, a[eat(e) \wedge agent(e, j) \wedge theme(e, a) \wedge named(j, john) \wedge apple(a)]$.

For this to work, the DRS construction rule for verbs has to be modified accordingly[7]. Instead of re-inventing the wheel, I will build on previous work, e.g. the DRT implementation developed by Johan Bos (Curran et al., 2007). In that framework, the sentence (7.2) is represented by the (slightly adapted) DRS (7.6).

(7.2) John was murdered yesterday.

7 In their standard work on DRT, Kamp and Reyle (1993) make use of a Davidsonian style of representation (Davidson, 1967).

(7.6)

e, j
$named(j, john)$ $murder(e)$ $patient(e, j)$

Let us see now how frames enter the semantic representation. If a verb $[_V \text{V}]$, e.g. "murder", introduces a condition $V(e)$ in a DRS, together with a discourse referent e for the corresponding eventuality, then the frame to which the verb is mapped by the corresponding lexical unit is evoked.

In Fillmore's frame semantics, (core) frame elements are defined as entities corresponding to thematic roles in an eventuality. In the representation language of (S)DRT, these entities will show up as discourse referents x_i. The thematic role played by such an entity, however, is represented as a relation ϕ between the discourse referent x_i and the discourse referent e corresponding to the eventuality, viz. $\phi(e, x_i)$. These thematic relations will be represented in a DRS K as ordinary conditions in C_K on corresponding discourse referents in U_K. In this way, an unnecessary augmentation of structural complexity of discourse representations is avoided[8].

In order to differentiate between discourse referents that are linguistically introduced and discourse referents that are merely evoked, I propose to assume two different kinds of discourse entities: *regular* discourse referents introduced by linguistic expressions, and *weak* discourse referents which are not (yet) expressed linguistically. *Weak Discourse Referents*[9] are abstract entities which are evoked or activated in course of the interpretation process. A linguistic expression does not introduce them directly, but rather indirectly by virtue of the frame evoked by a lexical unit. They often remain underspecified, but can be specified by subsequent anaphoric reference. The universe of a DRS K is divided into two sets of referents:

- U_K^r: regular discourse referents, and
- U_K^w: weak discourse referents,

with $U_K = U_K^r \cup U_K^w$. In the box-style representation of DRSs, this distinction is reflected by a delimiter "|" between the two types of referents:

(7.7)

U_K^r | U_K^w
C_K

8 In fact, the term *frame element* as used in FrameNet is ambiguous: it can refer either to the entity that plays the role corresponding to the frame element, i.e. the role-filler, or to the relation this entity has with respect to the frame.

9 The distinction between two types of discourse referents is not entirely new, e.g. Kamp and Rossdeutscher (1994) assume "schematic discourse referents", which correspond to underspecified discourse referents. Furthermore, this assumption could be generalized in the sense that all discourse referents are assigned finer-grained weights on a scale according to their salience, instead of differentiating just two kinds of referents. I will leave this point to further investigation.

With this distinction, Frame Evocation can be defined.

Frame Evocation

If $V \in LU(\Phi)$ and $e \in U_K$ and $V(e) \in C_K$ then

$\forall\psi[\psi \in FE(\Phi) \rightarrow \exists x_i[x_i \in U_K^w \wedge \psi(e, x_i) \in C_K]]$ and

$e : \Phi \in C_K$.

Let us illustrate with an example how the proposed mechanism works. According to FrameNet data (Baker et al., 1998), in the course of interpreting the utterance, the *Killing* frame is evoked by the verb "murder"[10]. We thus get from a DRS (7.8a) via frame evocation to a DRS (7.8b).

(7.8) a.

e
$murder(e)$

b.

e | x_1, x_2, x_3
$e : Killing$ $killer(e, x_1)$ $victim(e, x_2)$ $instrument(e, x_3)$

Due to the inheritance between frames, the *Killing* frame inherits the frame elements from the parent frame *Transitive_Action*. The frame element *Victim* of the *Killing* frame is a specification of the frame element *Patient* in the parent frame. In order to avoid duplifications of DRS-conditions, a transition from a DRS K to a DRS K' matching inherited frame elements is defined.

Matching Inherited Frame Elements

If $x \in U_K^w \wedge \phi(e, x) \in C_K$ then

a. $y \in U_{K'}^r \wedge \phi(e, y) \in C_{K'}$ if $y \in U_K^r \wedge \psi(e, y) \in C_K$ and $inherits^*(\phi, \psi)$,

b. $x \in U_{K'}^w \wedge \phi(e, x) \in C_{K'} \wedge x =? \in C_{K'}$ otherwise.

For example (7.2), the DRS-condition $victim(e, x_2)$ in (7.8b) matches the condition $patient(e, j)$ in (7.6). Hence, the weak discourse referent x_2 is eliminated and replaced by the regular discourse referent j. The condition evoked by the frame, i.e. $victim(e, j)$, is adapted and the more general condition $patient(e, j)$ is dropped. Since nothing is known about the remaining weak referents x_1 and x_3, they remain underspecified, as indicated by the conditions $x_1 =?$ and $x_3 =?$. All other discourse referents and conditions remain untouched. The resulting SDRS is depicted in (7.9).

10 Note that the *Killing* frame includes the core frame elements *Killer* and *Cause*, which have in fact a complementary distribution. Thus, example (7.2) can be seen as evoking the more specific frame *Murdering* with a representation of the core frame element *Killer*, whereas a representation of (i) would include the frame element *Cause* instead of *Killer*.

(i) John was killed yesterday. There was a terrible thunderstorm.

In fact, the FrameNet book (Ruppenhofer et al., 2005, p. 33) indicates this possibility.

(7.9)

$e, j \quad \mid \quad x_1, x_3$
$e : Killing$ $killer(e, x_1), x_1 =?$ $victim(e, j), named(j, john)$ $instrument(e, x_3), x_3 =?$

Prepared in this way, let us turn now to bridging anaphora. Consider Clark's classical example:

(7.10) a. John was murdered yesterday.

b. The knife lay nearby.

The underspecified semantic content of (7.10) can be expressed as shown in (7.11). The core frame elements of the *Killing* frame show up in the SDRS as $killer(e_1, x_1)$, $victim(e_1, j)$ and $instrument(e_1, x_3)$. Similarly, the verb "lie" (in its sense "lie nearby") evokes the frame *Being_located*, with only one core frame element $theme(e_2, k)$.

Recall from section 6.3.3 (cf. (6.51) on page 181) that definite noun phrases introduce a discourse referent, e.g. k for "the knife" (which is unique, represented as $!k$ in the universe) together with a bridging condition $B(a, k)$ involving a referent for the bridging anchor a. Both B and a are underspecified. Furthermore, the second utterance is to be connected via a coherence relation R to a previous utterance v. Also these two have to be specified in the course of interpretation. For the sake of clearness of the exposition, additional information, e.g. about tenses, is ignored.

(7.11)

$\pi_1, \pi_2 \quad \mid \quad R, v$		
$\pi_1 :$	$e_1, j \quad \mid \quad x_1, x_3$ $e_1 : Killing$ $killer(e_1, x_1), x_1 =?,$ $victim(e_1, j), named(j, john)$ $instrument(e_1, x_3), x_3 =?$	$\pi_2 :$
		$e_2, !k \quad \mid \quad B, a$ $e_2 : Being\_located$ $theme(e_2, k)$ $knife(k)$ $B(a, k), B =?, a =?$
$R(v, \pi_2), R =?, v =?$		

7.2.3 Establishing Discourse Relations by FrameNet Data

As already mentioned, thanks to the hierarchical structure of the FrameNet database, the *Killing* frame inherits the properties of the more general abstract frame *Transitive_action*, which in turn inherits from *Event*. The frame *Being_located* inherits the frame elements of the abstract frame *State*. As assumed in Asher and Lascarides (2003), the occurrence of an event followed by a state is a strong indicator for the presence of a BACKGROUND relation between the discourse segments containing the eventualities, expressed by the default rule from page 183 (cf. also section 5.3.2.2), repeated here for convenience.

Background $(?(\alpha, \beta, \lambda) \wedge event(e_\alpha) \wedge state(e_\beta)) > Background(\alpha, \beta, \lambda)$

With this rule, the underspecified relation R in (7.11) can be specified by default as BACKGROUND. Since the text (7.10) consists of only two sentences, v must be specified as π_1. However, the specification of the discourse segment to which a given utterance is to be attached is not always that trivial. In more complex discourse structures, additional constraints play a role in determining how an utterance triggering a bridging inference connects to the preceding discourse. I will turn to these constraints shortly, but first I want to illustrate how other discourse relations can be inferred from information provided by FrameNet.

Take, for example, NARRATION. Recall from section 5.3.2.2 that NARRATION can defeasibly be inferred if an event provides an occasion for another. In the glue logic, this default is expressed as follows.

Narration $(?(\alpha, \beta, \lambda) \wedge occasion(e_\alpha, e_\beta)) > Narration(\alpha, \beta, \lambda)$

Occasion, in turn, can normally be inferred "if two event types of a certain kind are to be related." (Asher and Lascarides, 2003, p. 201). More specifically, it is suggested that the relation between the events can be derived from script or frame knowledge. How can we establish such a relation by virtue of information encoded in FrameNet? Let us consider an example.

(7.12) a. Last April he claimed that more than $750,000 in cash was stolen from his Jerusalem residence.

b. When police investigated the theft, Capucci insisted that the money had been returned, and asked them to drop the matter.

(TIME Magazine, Sep. 02, 1974[11])

In the first sentence, the verb phrase "was stolen" corresponds to the lexical unit *steal.v*, which evokes the frame *Theft*. This frame inherits from the frame *Committing_crime*. In the second sentence, the verb "investigate" evokes the frame *Criminal_investigation*. Both *Committing_crime* and *Criminal_investigation* are subframes of the abstract frame *Crime_scenario*, where *Committing_crime* precedes *Criminal _investigation*. Hence, there is reason to suspect that the event corresponding to *Committing_crime* provides an occasion for the eventuality corresponding to *Criminal_investigation*. More generally, if Φ_1 and Φ_2 are both subframes of a third frame Ψ and precede each other then normally an occasion relation holds between the two eventualities corresponding to the frames. I propose to express this knowledge by the following fairly general default rule:

Inferring Occasion from Subframes

If Φ_1, Φ_2, Ψ are frames with $e_1 : \Phi_1$ and $e_2 : \Phi_2$ then

$(subframe(\Phi_1, \Psi) \wedge subframe(\Phi_2, \Psi) \wedge precedes(\Phi_1, \Phi_2)) > occasion(e_1, e_2)$.

By virtue of this rule, an occasion relation can be assumed in example (7.12), and the axiom for NARRATION can be applied, yielding a NARRATION relation between (7.12a) and (7.12b). Similar rules can be postulated for other discourse relations. In section 8.3.2 will be shown how clues for establishing an ELABORATION relation can be inferred from FrameNet data.

11 http://www.time.com/time/magazine/article/0,9171,943729,00.html

7.3 Constraints on Bridging Inferences

Resolving bridging anaphora requires two problems to be solved:

(i) the correct anchor to which the anaphor is to be connected has to be found,

(ii) the nature of the bridging relation itself must be identified.

For solving (i), possible anchors must be identified, and impossible ones must be ruled out. Accessibility of antecedents (cf. section 3.3.2) must be further restricted by a number of constraints in order to rule out dispreferred interpretations of bridging anaphora. For solving (ii), I suggest to restrict possible relations to conditions on discourse referents already present in the discourse model.

7.3.1 Bridging Constraints

In the following, I will discuss which factors constrain bridging inferences, taking (7.10) as example, and relate them to SDRT's meta-rules on bridging. Recall from section 6.3.3 that the resolution of bridging anaphora in SDRT relies on four meta-rules, as stated in Asher and Lascarides (1998a):

1. If possible use identity.
2. Bridges must be plausible.
3. Discourse structure determines bridging.
4. Maximize discourse coherence.

I will take these rules as a starting point in the following discussion of constraints on bridging inferences.

7.3.1.1 The Preference for Coreference

The first rule "if possible use identity" expresses the empirical preference that the referent of an anaphoric expression is identical to another previously introduced referent. This preference for identical referents of anaphor and antecedent seems to be subsumed by a very general constraint in discourse interpretation, sometimes called DOAP ("Don't overlook anaphoric possibilities", Williams, 1997), essentially stating that if there is an anaphoric trigger, we must try to find an antecedent. As we have already seen at various points in this thesis (cf. sections 5.2.4.2 and 6.3.3.3), this preference for coreference can be expressed by Equality by Default (Cohen, 2006, 2007): unless there is evidence for the contrary, two discourse referents can be assumed to be equal. This default accounts well for the resolution of direct anaphora. However, in order to extend it to cover bridging anaphora, it must be adapted because bridging involves a relation between anchor and anaphor that is different from coreference. In the account of bridging in SDRT, an anaphoric expression introduces a bridging condition $B(a, x)$ involving a weak discourse referent a. This referent can be seen as being coreferential to the bridging anchor α introduced earlier in the discourse.

In general, if a is an underspecified weak discourse referent in π_j, then it wants to be resolved to an available discourse referent α introduced earlier. For direct anaphora, α must be a regular discourse referent, and for indirect or bridging anaphora, α is a weak discourse referent. Both cases are captured by the following adapted default rule:

Equality by Default for Discourse Anaphora

$$a \in U^{w}_{K_{\pi_j}} \wedge \alpha \in U_{K_{\pi_i}} > a = \alpha$$

This default, which is to be understood as a last-resort principle, can (and often must) be blocked by higher-ranked constraints in order to avoid overgeneration. Let us look at the other conditions on bridging in SDRT.

7.3.1.2 Plausibility and Consistency

Asher and Lascarides (1998a)'s second meta-rule, "bridges must be plausible", says that the specification of bridging relations is influenced by world knowledge, i.e. general encyclopedic knowledge about entities in the world, properties and relations between them, the occurrence of events and actions, causal, temporal, and spatial relations, etc. Language users make assumptions and have expectations about specific situations drawing on knowledge of previous experiences. In this way, in a given situation, some eventualities are more plausible than in another situation. Since FrameNet encodes this kind of information in its descriptions of specific schemas of stereotypical situations and scenes, it can provide a suitable basis for determining if a bridging relation in question is plausible or not.

It was suggested in section 7.2.3 that FrameNet data can give important clues for establishing discourse relations. However, the knowledge of discourse structure is not always sufficient to resolve bridging anaphora. In example (7.10), the presence of a BACKGROUND relation alone is not enough to motivate the bridge. Let us examine whether there is further information provided by FrameNet that can be used for establishing plausible bridging relations. The frame element *Instrument* in the *Killing* frame must be of the sort *physical_entity*. It can be a weapon, but in principle any other physical entity could be used for killing, e.g. hands (7.13) or a lamp (7.14) .

(7.13) John killed Mary. He strangled her.

(7.14) John killed Mary. He stunned her with a lamp.

On the one hand, the lexical unit "knife" evokes the frame *Weapon* of a sort *artifact*, indicating the possibility that it could serve as an instrument in a killing event. But, on the other hand, as noted in the informal FrameNet description, knives are not necessarily designed as weapons. So this knowledge does not really help us to resolve the bridging relation, at least in the present state of FrameNet. The only knowledge we can use is that there is no clash of sorts: both knives and killing instruments are physical entities. As far as that we can capture the intuition behind the plausibility constraint. It amounts to demanding that interpretations must be *consistent*. In fact, as Zeevat (2006) suggests, selecting the most plausible interpretation of an utterance in a given context entails a preference for consistent over inconsistent interpretations.

But what does it mean exactly for a discourse to be consistent? To answer this question, let us look at two examples. The discourses in (7.15) and (7.16) are unacceptable because they are inconsistent.

(7.15) # John is a musician. John is not a musician.

(7.16) # Mary is married. Mary does not have a husband.

Note that while the inconsistency of (7.15) is a matter of pure logic, in (7.16) it is not. In the latter example, it depends on additional background knowledge, namely that Mary is a woman and that married women have husbands. While logical consistency just consists in rejecting $p \wedge \neg p$, there seems also to be a kind of extra-logical, "pragmatic" consistency. In fact, there are cases that are logically consistent and pragmatically inconsistent. Consider, e.g., Moore's paradox[12]: it is absurd to say something like "It's raining outside but I don't believe that it is." ($p \wedge believe(\neg p)$), although, in a formally strict sense, this statement is logically consistent.

On the one hand, the definition of logical inconsistency is straightforward. In SDRT, it is part of the principle MDC (Maximize Discourse Coherence, cf. section 5.3.2.5 above). I will turn to this principle shortly.

On the other hand, a pragmatic notion is much more difficult to state explicitly. Recall from section 3.4.1 that a notion of "pragmatic consistency constraints" can be found in neo-Gricean theories of anaphora resolution, for example, in the works of Huang (1994) and Blackwell (2003), who start from the assumption that anaphora interpretation is constrained by a set of consistency conditions, which are (i) background knowledge, (ii) semantic constraints, and (iii) antecedent salience.

In fact, all these conditions play their role in bridging anaphora resolution. As for background knowledge, we have already seen that FrameNet data provides a rich information source. Semantic constraints comprise both logical consistency and pragmatic consistency regarding general ontological properties of entities. This kind of knowledge is encoded in FrameNet in form of information about sorts ("semantic types") of frame elements. As for antecedent salience, it is promising to encode information about the attentional state in the discourse model. We have seen in section 3.3.2 that there are many different ways of dealing with salience of discourse referents. It is in no wise trivial to combine a sophisticated account of salience such as Centering Theory (Grosz et al., 1995) with a theory of discourse structure such as SDRT. A first step to account for the relative salience of bridging antecedents is taken by the distinction of two kinds of discourse referents. I will return to this issue in section 7.3.2 below.

7.3.1.3 The Right Frontier Constraint

Accessibility for anaphoric reference is constrained by general discourse principles such as the Right Frontier Constraint (RFC, introduced in section 4.1.3 on page 100, for RFC in SDRT, see sections 4.2.4 and 5.3.2.4).

Basically, the right frontier of a discourse consists of the last discourse segment and all segments dominating it. This constraint draws a distinction between coordinating and subordinating discourse relations: a coordinating relation pushes the right frontier to the right, closing off its attachment point, and a subordinating relation extends the right frontier downwards, leaving open its attachment point. In SDRT, an available antecedent for an anaphoric expression must be DRS-accessible on the right frontier (Asher and Lascarides, 2003; cf. also section 5.3.2.4 on page 158). I will take this constraint to cover Asher & Lascarides' meta-rule "discourse structure determines bridging".

In example (7.10), the anchor for "the knife" must be in a DRS-accessible segment on the right frontier. Recall that a BACKGROUND relation between the two utterances can

12 This paradox is named after G.E. Moore (1873-1958), who is supposed to have it discussed in one of his lectures.

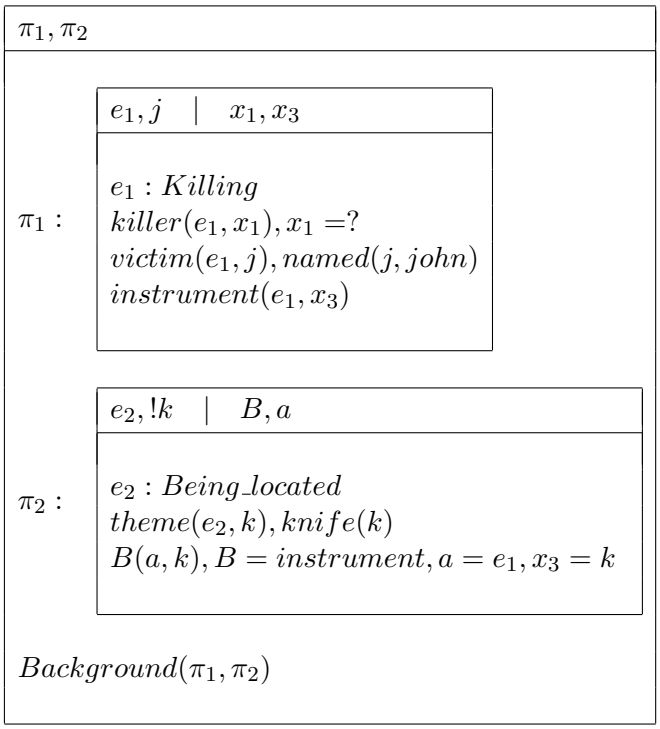
be assumed by default. Recent work on SDRT (Vieu and Prévot, 2004) has revealed that BACKGROUND should be considered as subordinating by default. Accordingly, in (7.10), π_1 lies on the right frontier of the discourse, and e_1 is available for anaphoric reference in π_2. Hence, the discourse structure tells us that, in principle, a bridging relation can be established. We are left with the question of how to build the bridge between the knife and the killing event.

With the presence of a discourse relation between π_1 and π_2, the discourse referents in π_1 are available for anaphoric reference in π_2. So, with Equality by Default, it can be assumed that a is equal to e_1. Thus, the bridging relation $B(a, k)$ can be specified as $instrument(e_1, k)$. As a byproduct, the underspecified variable x_3 in the condition $instrument(e_1, x_3)$ in π_1 can be resolved to k, yielding that instrument and knife refer to the same entity. Although k is not accessible in π_1, it is available in the superordinated SDRS comprising both utterances, and therefore, after processing the second utterance, the underspecification can be resolved.

7.3.1.4 Maximize Discourse Coherence

For illustration, the SDRS (7.11) for discourse (7.10) is pragmatically enriched as shown in (7.17). Note that, since the murderer is not mentioned at all, his referent could not be resolved and its representation remains underspecified.

(7.17)



I have already mentioned in the previous section that weak discourse referents often remain underspecified, with the possibility to be specified by subsequent anaphoric reference. This is what happens with the killing instrument. Its identification with the knife helps to render the discourse more coherent. If the knife in the second sentence had nothing to do with the first sentence, the discourse would be rather incoherent, at least after uttering the second sentence.

So far, bridging resolutions are preferably identity relations, they must be plausible, i.e. semantically and pragmatically consistent, and they must obey the right frontier constraint. But sometimes, neither plausibility nor availability are sufficient to establish a bridging relation. Consider discourse (7.18).

(7.18) a. John was murdered yesterday. b. # The book lay nearby.

This discourse is – in a neutral context – less coherent than (7.10), and I will explain why. In example (7.10), the knowledge that a knife is a kind of weapon that can serve as an instrument in a killing event licenses the bridging inference. In example (7.18), such a connection cannot be found. Also here, a BACKGROUND relation can be inferred, but the role which "the book" could play in the killing event is less clear than that of a knife. Although there is no clear semantic connection between "the book" and any evoked core frame element, there is no clash of sorts, and a bridging relation to the instrument could be plausible. Nevertheless, as no sense of "book" evokes a frame similar to *Weapon*, it remains unclear what nature has the bridging relation, and as a consequence, the discourse seems less coherent. Note that if the context provides additional evidence that the book is a probable killing instrument, e.g. by being contaminated with poison (for instance in Umberto Eco's novel "The name of the rose"), the bridging inference indeed can be drawn.

The inferences drawn so far are defeasible and can be overridden by subsequent information. Nevertheless, there is a preference for bridging relations to be resolved, because the existence of bridging anaphora makes a discourse more coherent. This is the intuition behind one aspect of Asher & Lascarides' fourth meta-rule "maximize discourse coherence" (MDC). Recall from section 5.3.2.5 that the MDC consists of four parts:

- minimize labels,
- be consistent,
- maximize rhetorical connections,
- resolve underspecifications.

The first constraint just counts the nodes in a discourse update. This is very easy to compute and does not add much complexity to the MDC.

The second constraint is logical consistency, which is subsumed by the notion of plausibility discussed above.

The third constraint is the heart of the MDC: it states a preference for interpretations which maximize the number and the quality of rhetorical relations.

The last constraint seeks to resolve underspecifications if possible, including anaphoric conditions. At first sight, this constraint looks very similar to "Don't overlook anaphoric possibilities" (DOAP) disussed above. However, it seems[13] that one should draw a distinction between the resolution of ambiguity (including anaphora) and a constraint that says that anaphora must be bound (DOAP, that is, essentially, Equality by Default). While it is true that we may want to leave some ambiguities, we normally do not want to leave a pronoun unbound. For example:

13 Thanks to Arik Cohen (p.c.) for pointing out this issue.

(7.19) John studied, but he went drinking.

If "he" is interpreted as referring to someone other than John, the sentence gets a very plausible and strong rhetorical interpretation: John studied, but, in contrast, somebody else did not. If "he" is interpreted as coreferential with John, the rhetorical connection is much less obvious. The MDC predicts both interpretations. But it seems that in this case, unless there is an indication for a deictic interpretation (pointing, intonation, etc.), there is a preference to bind the pronoun to the antecedent.

7.3.2 Weak Discourse Referents as Bridging Anchors

To summarize the principles we need for bridging resolution, we remain with the following general constraints on anaphoric reference:

- DOAP (via Equality by Default)
- PLAUSIBLE (subsumes CONSISTENT)
- RFC
- MDC (minimize labels,

  maximize rhetorical connections,

  resolve underspecifications)

Note that they are not meant to be special meta-rules designed for bridging resolution, they rather seem to be more general constraints to be obeyed in discourse interpretation[14]. More formally, the resolution of a bridging anaphor must obey the following constraints:

Constraints on Bridging Anaphora Resolution

If $x \in U^r_{K_{\pi_j}}$ and $B, a \in U^w_{K_{\pi_j}}$ and $B(a,x) \in C_{K_{\pi_j}}$ and $\alpha \in U_{K_{\pi_i}}$ then
by default, $a = \alpha$ holds for the following cases: (DOAP)

1. if $\alpha \in U^r_{K_{\pi_i}}$ then $B = \mathit{identity}$ with $x = a$,
2. if $\alpha \in U^w_{K_{\pi_i}}$ then either
   a. $B = \sqsubseteq^+$ with $x \sqsubseteq^+ \alpha$, or
   b. $B = \phi$ with $\phi(\alpha, x) \in C_{K_{\pi_i}}$,

provided that

- the sort of α is compatible to the sort of a, (CONSISTENT)
- α is available for anaphoric conditions in π_j, and (RFC)
- MDC is not violated. (MDC)

14 These principles could be seen as constraints in optimality theoretic pragmatics, but I will not adopt a particular framework here, as I leave open the question whether the ranking of the constraints should be left as stated above. For a related and much more sophisticated discussion on discourse constraint ranking, see Zeevat (2006).

Thus, Equality by Default is considerably restricted by a number of additional constraints. Direct anaphora are accounted for by the first case, where the relation B is identity. As elaborated in section 6.2, a (proper) bridging relation can be either mereological or frame-related. In the former type, B is $\sqsubseteq^+$, which stands for any type of mereological relation (e.g. member/collection or portion/whole etc.). In the latter type, B is a condition ϕ that is already present in π_i.

In a nutshell, my proposal is to restrict the search space for suitable antecedents for bridging anaphora to take into account only accessible regular and weak discourse referents. Conditions on discourse referents already present in the discourse model are considered as preferred bridging relations. Both indirect anaphora (involving roles in eventualities as well as mereological relations) and direct anaphora are covered by this assumption. In this way, the resolution of bridging anaphora can be considerably constrained. In the proposed account, new entities are (weakly) introduced with every eventuality that is talked about, with the potential to be strengthened, to remain in the background, or even to be dropped.

7.4 Related Approaches

7.4.1 Implicit Arguments as A-definites (Koenig & Mauner, 1999)

Important work on the discourse status of non-expressed event participants was presented by Koenig and Mauner (1999). These authors report results of psycholinguistic experiments concerning implicit verbal arguments. In an experiment carried out by Mauner et al. (1995), reading times of sentences like (7.21) following one of the sentences in (7.20) were compared.

(7.20) a. A ship was sunk

b. A ship sank

c. A ship was sunk by someone

(7.21) ... to collect settlement money from the insurance company.

Subjects take longer to process rationale clauses like (7.21) when they follow intransitive sentences like (7.20b) than when they follow short passives (7.20a) or agentive passives (7.20c). Thus it seems reasonable to assume that verbs like "sink" in (7.20a) include an implicit actor argument as part of the representation of the lexical item, making it easier for readers to anchor the implicit anaphoric (PRO) subject of "collect" in (7.21) in their discourse model.

Koenig and Mauner (1999) further argue that nonquantificational NPs and pronouns have three main distinct functions in discourse: they can introduce new discourse referents to which subsequent NPs can refer back; they satisfy one of a main predicate's arguments; and they impose a restriction on the discourse referent they introduce. It is observed that the French subject clitic "on", as well as the German "man" and indefinite uses of English "they" do not fulfill the first function. For instance, discourse (7.22) with a coreferential reading of "on" and "il" is not felicitous, while (7.23) is fine (Koenig and Mauner, 1999, p. 213).

(7.22) a. On$_i$ a assassiné la présidente.
CL.subj.indef have.3sg.PAST kill.PTCP the.f president.f

b. #Il$_i$ était du Berry, paraît-il.
he.m be.3sg.PAST from.the Berry, seem.3sg-it
'Someone$_i$ murdered the (woman) president. He$_i$ was from the Berry, it seems.'

(7.23) a. Quelqu'un$_i$ a assassiné la présidente.
someone.m.indef have.3sg.PAST kill.PTCP the.f president.f

b. Il$_i$ était du Berry, paraît-il.
he.m be.3sg.PAST from.the Berry, seem.3sg-it
'Someone$_i$ murdered the (woman) president. He$_i$ was from the Berry, it seems.'

It seems that in general, implicit arguments in short passive sentences, as well as words like French "on" (*a-definites* in their terminology) cannot serve as antecedents of anaphora. Koenig and Mauner (1999) claim that implicit arguments do not introduce any discourse referent at all. Their DRT representation for sentence (20a) is (7.24):

(7.24)

y
$ship(y), sink(x, y)$

In this representation, it remains unclear how the apparently free variable x, representing the actor, is model-theoretically interpreted. Moreover, as noted in their paper, indirect references to implicit arguments *are* indeed possible, e.g. consider example (7.25).

(7.25) a. They killed the president.

b. The terrorists were merciless.

Koenig and Mauner (1999) do not give any details on how such an inference can be drawn according to their theory. The interpretational apparatus of DRT (Kamp and Reyle, 1993) would have to be changed in order to allow uninstantiated variables in final DRSs. Such an attempt is made by Farkas and Swart (2003) who, however, have to make a major modification of truth conditions in DRT.

7.4.2 Bridging as Coercive Accommodation (Bos et al., 1995)

Bos et al. (1995) presented a lexical approach to bridging. Basically, they combine an extension of van der Sandt (1992)'s theory of presupposition in DRT (cf. section 3.4.3) with the *Generative Lexicon* (Pustejovsky, 1995).

To the two ways of resolving anaphora proposed by van der Sandt (1992), linking and accommodation, they add the third option of bridging (cf. section 6.3.2.1). DRS-construction is modelled as a two-stage process (cf. page 90). A single sentence is represented as an underspecified DRS with respect to anaphoric information. In a second stage, this DRS is added to the context DRS resulting in a proper DRS without unresolved anaphoric information, which can be model-theoretically interpreted as in standard DRT.

The central notion in the Generative Lexicon, *qualia structure*, can be seen as a set of *lexical entailments* of a lexical entry. For example, the word "book" lexically entails the events of *reading* and *writing* it and consists of several separate parts, like the cover, pages, etc. Four qualia roles are distinguished to represent such knowledge: *formal*, *constitutive*, *telic* and *agentive*. The agentive qualia involves a writing event, and the telic role can be attributed to a reading event.

(7.26) Yesterday, John began a new book.

In example (7.26), the verb "begin" expects an *event*, but the only entity expressed in the sentence is a "book", i.e. an *artifact*. In order to resolve this mismatch, it is assumed that *coercion* takes place. Coercion means that whenever a word or a phrase is not of a desired sort[15], it can be coerced into one of its entailments that is of the appropriate sort.

In their account of bridging with qualia structure, Bos et al. (1995) propose the following extension of DRT. Two new condition types are defined, α-DRSs, which represent unresolved anaphoric information, and Q-DRSs, which represent qualia structures. The syntax of DRSs (cf. rule 2 of the definition on page 87) is extended straightforwardly by the following rule:

Extended Syntax of DRSs

If K is a DRS, then $\alpha : K$ and $Q : K$ are DRS-conditions.

Anaphoric information is resolved in compliance with van der Sandt (1992)'s algorithm. Qualia information, as the authors state, "is normally not accessible and does not affect the truth-conditions of a DRS" and "is introduced in the lexicon and brought into discourse via the DRS bottom-up construction algorithm" (Bos et al., 1995, p. 3). Whenever necessary, for example to play the role of an antecedent, the qualia structure is put forward to the surface by a process called *coercive accommodation*, a function from DRSs to sets of DRSs[16].

Coercive Accommodation

Let $K = \langle U_K, C_K \rangle$ be a DRS. Then $CA(K) = \{K \oplus K_Q \quad | \quad Q : K_Q \in C_K\}$.

In DRT, for a discourse referent to be the antecedent for an anaphor, it must be *accessible* from the DRS in which the anaphor is represented. Accessibility of discourse referents is defined as in standard DRT (cf. section 3.4.3 on page 91). An additional constraint on the choice of antecedent suggested by the authors is *suitability*: "A DRS is *suitable* to another DRS if there is a way finding a match between discourse referents and conditions between both" (Bos et al., 1995, p. 7). To exemplify this account, the following example (7.27) is presented.

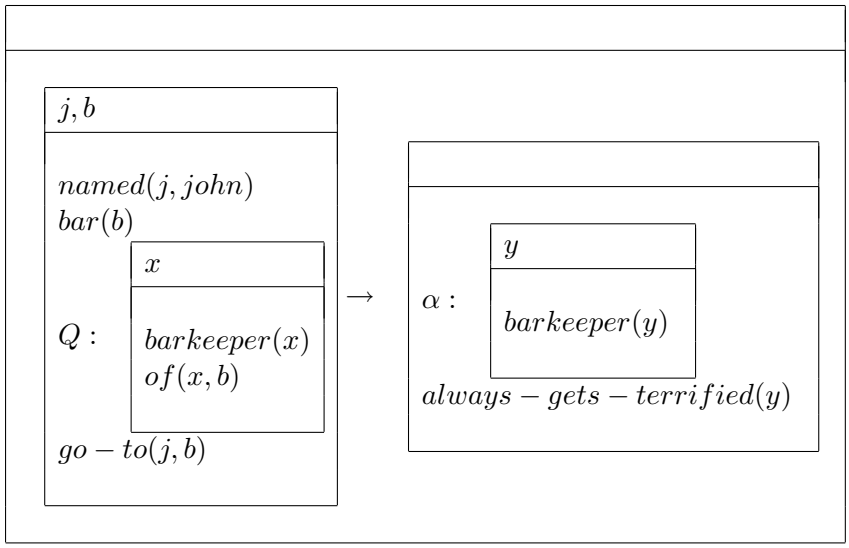
(7.27) When John goes to a bar, the barkeeper always gets terrified.

15 Pustejovsky (1995) uses the term "type" for sorts such as *artifact*, *event*, etc. In fact, his notion is very similar to the notion of "semantic type" as used in Frame Semantics. In this thesis, however, I will stick to the term "sort".

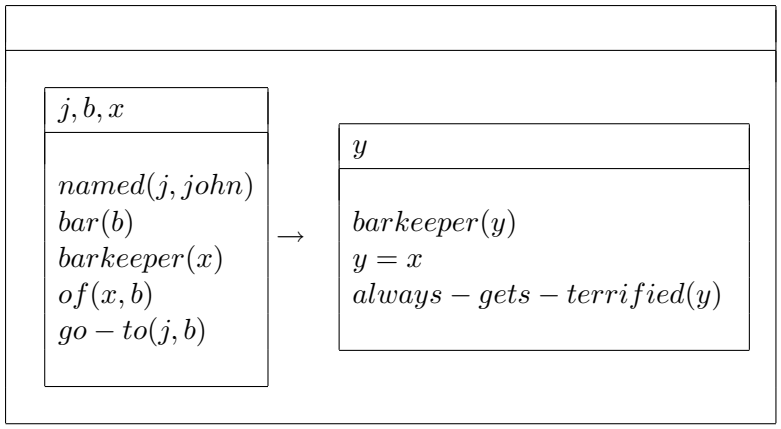
16 The merge-operator $\oplus$ was defined on page 87 in terms of the set unions of the universes and sets of DRS-conditions in question, i.e. $K_1 \oplus K_2 = \langle U_{K_1} \cup U_{K_2}, C_{K_1} \cup C_{K_2} \rangle$.

The qualia structure for "bar" includes a representation of a barkeeper. It is a constitutive qualia, indicated by the condition of. An unresolved preliminary DRS for this sentence is given in (7.28a). In the second step of DRS construction, it is resolved to (7.28b). Note how the qualia information is brought into the DRS representing the first clause via coercive accommodation.

(7.28) a.



b.



This approach is based on a convincing and straightforward formal definition. However, there are some major problems to be observed. First, the theory is presented as an extension of DRT and thus suffers from the same drawbacks which we already discussed in connection with DRT. So the salience of potential candidates for anaphora resolution is not modelled. Moreover, the approach fails to account for complex discourse structures beyond simple subordination.

Second, the treatment of bridging as a lexical phenomenon is not free of problems. It is limited to lexically induced bridging inferences. Even if we take, as in example (7.27), knowledge about bars and barkeepers and their relationship as lexical knowledge, there are cases involving knowledge far beyond the lexicon. Bos et al. (1995) cite the following example as a limitation case of their approach:

(7.29) Probably, if Jane takes a bath, Bill will be annoyed that there is no more hot water.

Interpreting this short discourse involves the inference that taking a bath involves using a hot water reservoir. This inference is difficult to explain in Bos et al. (1995)'s framework because a hot water reservoir is not part of the qualia structure of "take a bath" in the lexicon.

Let us examine if the above sketched approach using FrameNet can account for this kind of bridging inferences. In the present state of English FrameNet it is unclear whether phrasal verbs are lexical units and how they evoke frames, e.g. whether "take a bath" counts as a lexical unit, or just "take". However, in other versions of FrameNet, such knowledge is encoded; an equivalent sentence in Spanish using the verb "bañarse" ("to take a bath") is analyzable in FrameNet terms[17]. There, it evokes the frame *Cause_to_be_wet* with a core frame element *Liquid*, which can be instantiated by "hot water". Still better is a suggestion made by the developers of Polish FrameNet[18], according to which both "wziąć kąpiel" (like in English) and "wykapać się" (like in Spanish) evoke the frame *Grooming*, where an *Agent* engages in personal body care. An *Instrument* can be used in this process as well as a *Medium*. Thus, if "take a bath" is treated as a lexical unit, we can draw the inference that the water in the second clause is used for the bath in the first clause.

Another problem with the lexical approach to bridging was noted by Piwek and Krahmer (2000). Not all bridging antecedents are lexical entailments. Sometimes, background knowledge clearly beyond the lexicon is necessary. These authors give the following example:

(7.30) Yesterday, Chomsky analyzed a sentence on the blackboard, but I couldn't see the tree.

To correctly understand this utterance, the hearer has to rely on specific background knowledge, in particular on the knowledge that a generative syntactic analysis typically involves a tree-like representation of the sentence. It is questionable whether highly context-sensitive information of this kind is part of the lexicon, as I had already discussed in section 6.1.3 on psycholinguistic research on bridging involving implicit event participants. Although neither a lexical approach as taken by Bos *et al.* nor my approach as presented in this chapter can account for such complex inferencing processes, I suggest a way of dealing with contextual knowledge that can be fairly easily extended (via inclusion of non-core FEs) in order to include this kind of encyclopedic knowledge.

7.4.3 FrameNet and DRT (Bos & Nissim, 2008)

Very recently, Bos and Nissim (2008) presented a sketch of a combination of DRT and FrameNet, which shares many commonalities with my proposal as presented here. It builds on the basic idea of representing frame elements as DRS-conditions. They observe also that in order to represent frame elements in DRT, it is best to replace the event semantics of standard DRT (Kamp and Reyle, 1993) with a neo-Davidsonian representation.

Furthermore, they point out another representational problem in combining DRT and FrameNet. As FrameNet assumes that frames are evoked by any type of lexical units, that is to say, essentially verbs, noun phrases, and adverbs, it is straightforward to assume that frames evoked by noun phrases and adverbs enter the discourse representation as well. This

17 Subirats Rüggeberg (2005); see http://gemini.uab.es/
18 Zawisławska et al. (2008); see http://www.ramki.uw.edu.pl/

is unproblematic for noun phrases, but since adverbs do not introduce discourse referents, the representation of frame elements in the discourse model seems rather tricky in these cases. Another difficulty to be mentioned in this context is that when combining frames evoked within a single utterance, possible inconsistencies that can arise between different evoked frames must be dealt with. However, due to its limitation to DRT, which abstracts over discourse segments, the account of Bos and Nissim (2008) cannot account for the interplay of discourse structure and bridging phenomena as our extension of SDRT, which preserves discourse segment ordering, does.

An important point elaborated by Bos and Nissim (2008) is that there is no straightforward mapping of the grammatical function of lexical expressions to the thematic role they play in a frame. As we have seen before (cf. Fig. 7.2), FrameNet data includes valence descriptions for lexical entries, where possible matchings of grammatical function and expressed frame element, as well as their frequencies in the annotated examples, are listed. Very interestingly, Bos & Nissim used a supervised machine learning system to automatically classify semantic roles on the basis of annotated FrameNet examples. Their system reached an accuracy of over 90% in assigning semantic role labels (frame elements) to lexical expressions in a test data set of around 10,000 instances, given that the corresponding frame was already known. These results are very promising and show that an integration of formal semantics and cognitive modeling bears not only theoretical but also practical advances.

7.5 Conclusion

In this chapter, I presented an extension of SDRT's account of bridging to cover reference to eventualities. It was spelled out how world knowledge, represented in frames, contributes to the interpretation process, both for establishing discourse relations and for resolving indirect anaphora. Although some shortcomings in integrating the two lines of research still have to be resolved, in particular the choice of suitable frames for more complex verbal expressions and the interaction of different frames evoked within a single utterance, the integration of FrameNet and SDRT works quite straightforwardly, assuming a neo-Davidson representation of eventualities and distinguishing two kinds of discourse referents, the regular and weak discourse referents. In addition, I was able to indicate that the meta-principles for bridging of Asher and Lascarides (1998a) can be put down to more general constraints to be obeyed in discourse interpretation.

Chapter 8

Bridging by Clitic Left Dislocation

In this chapter, the suggested approach to bridging is applied on the realization of bridging inferences triggered by a specific construction: clitic Left Dislocation in Romance languages. Section 8.1 investigates dislocation construction in various languages. Various subtypes can be made out, involving different grammatical phenomena. In section 8.2, discourse functions of clitic Left Dislocations are examined, including familiarity, discourse topic, contrast, and constraints on discourse structure. Topic of section 8.3 is the discourse integration of clitic Left Dislocation constructions involving bridging inferences.

8.1 Dislocation Constructions Across Languages

Left Dislocation is a construction, or more generally a family of similar constructions, that can be found in many languages, although there are many subtle differences both between particular constructions and across languages. The term was first introduced by Haj Ross (1967). Generally, Left Dislocation is characterized as follows. A phrase is moved out of its original position to the left periphery of a sentence. In the main clause, a resumptive element, which is coreferential with the dislocated phrase, is left behind.

(8.1) Jerry$_i$, I don't see him$_i$ very often.

Left Dislocation must not be confused with some superficially similar constructions such as Topicalization and Focus-Fronting. Furthermore, various types of Left Dislocation can be distinguished. We will turn to these issues in the following sections.

8.1.1 Left Dislocation vs. Topicalization

Ross's programme was to establish a series of constraints on syntactic rules, which are now known as "island"-constraints. In essence, they state that extraction (A'-movement) of constituents which are embedded in other constituents such as complex DPs or coordinate structures ("islands") is not possible (cf. Ross 1967; Chomsky 1973). Ross used these constraints to distinguish Left Dislocation (8.2a) from another superficially similar construction called *Topicalization* (8.2b).

(8.2) a. My father, the man **he** works with in Boston is going to tell the police that that traffic expert has set that traffic light on the corner of Murk Street far too slow.

b. *My father, the man ∅ works with in Boston is going to tell the police that that traffic expert has set that traffic light on the corner of Murk Street far too slow.

(Ross 1967, p. 233)

Topicalization differs from Left Dislocation in two major respects: First, in a Topicalization construction, a constituent is extracted from a clause and moved to the left periphery, and what remains is an incomplete clause. The detached constituent is not resumed by a placeholder in the main clause. Second, Topicalization is sensitive to island constraints, thus leading to ill-formedness of the main clause in (8.2b).

Topicalization is a phenomenon common to many languages. For instance, in the French example (8.3), the object DPs "aux filles" and "aux garçons" are topicalized. Similarly, the German sentence (8.4a) in canonical form can occur in a topicalized version (8.4b). For comparison, the corresponding sentence in a Left Dislocation construction is given in (8.4c).

(8.3) Marie a réuni les élèves. Aux filles, elle a donné des exercices d'algèbre et aux garçons, elle a dicté un problème de géométrie.

'Marie reunited the pupils. To the girls, she gave exercises in algebra, and to the boys, she set a problem of geometry.' [French]

(Doetjes et al. 2002)

(8.4) a. Ich sah diesen Film, als ich ein Kind war.
I saw this.acc film when I a child was
'I saw this movie when I was a child.'

b. Diesen Film sah ich, als ich ein Kind war.
this.acc film saw I when I a child was
'This movie I saw when I was a child.'

c. Diesen Film$_i$, den$_i$ sah ich, als ich ein Kind war.
this.acc film it.acc saw I when I a child was
'This movie, I saw it when I was a child.' [German]

8.1.2 Left Dislocation vs. Focus Fronting

Focus fronting consists in placing a focussed constituent to the beginning of a sentence. In the following English (8.5a), German (8.5b), and Spanish (8.5c) examples, the focus constituent is marked by $[_F]$, and t_i marks the original position of that constituent in canonical word order from where it is moved to the left periphery.

(8.5) a. $[_F$ Three hundred dollars$]_i$ we raised t_i last week.

b. $[_F$ Fleisch$]_i$ esse ich nicht t_i.
meat eat.1sg I not
'I don't eat meat.' [German]

c. $[_F$ Dos cañas$]_i$ me he tomado t_i.
two beer REFL.1sg have.1sg taken
'I had two beer.' [Spanish]

Focus fronting is similar to Topicalization in that the fronted constituent remains a complement of the verb. It differs prosodically from Topicalization in the sense that it has a characteristic intonational pattern, at least in English and German. It differs pragmatically from Topicalization in that it conveys a contrastive or a new information focus. The fronted constituent pertains to the focal part of a sentence, whereas a topicalized phrase is not part of the focus but constitutes the topic of a sentence.

8.1.3 Hanging Topic Left Dislocation vs. Clitic Left Dislocation

So far we have seen that Left Dislocation involves, apart from a detachment to the left sentence periphery, a placeholder in the main sentence which is coreferential with the dislocated element. This placeholder, or resumptive element, can be a regular personal pronoun as in English, a demonstrative pronoun as in Dutch and German, or a clitic pronoun in Romance languages like French, Spanish, Catalan, Italian, as well as in Greek. Various kinds of Left Dislocation have been distinguished, depending on the category of the dislocated element, the nature of the resumptive element, and the syntactic relation between them. In the generative literature (for an overview, see e.g. Anagnostopoulou et al. 1997), two types of Left Dislocation constructions have been distinguished on the basis of syntactic properties: Clitic Left Dislocation (CLLD) and Hanging Topic Left Dislocation (HTLD). Based on Italian data, Cinque (1990) proposed a series of criteria for telling the two constructions apart. These criteria are summarized in Table 8.1. Virtually any

	HTLD	CLLD
category of dislocated phrase	DP	DP, PP, AP, CP
type of resumptive element	epithet, or clitic or tonic pronoun	clitic pronoun
case matching	no	yes
island sensitivity	no	yes

Table 8.1: Syntactic properties of CLLD and HTLD

maximal XP can be dislocated in CLLD constructions (Cinque 1990, p. 57f.) (8.6,8.7), provided that the corresponding clitic exists in the given language. For instance, Spanish is more restricted with respect to the category of dislocated elements and lacks locative and partitive clitics, which can be found in French ("y","en") and in Italian ("ci","ne"). Catalan has even more clitics, e.g. "ho" refers to a copula complement (8.8).

(8.6) $[_{PP}$ Al mare$]_i$, ci$_i$ siamo già statti.
At sea CL.loc be.1pl already be.PTCP.pl
'To the seaside / we have already been there.' [Italian]
(Cinque 1990)

(8.7) $[_{PP}$ Sobre la taula$]_i$, no l'hi$_i$ he posat.
over the table not CL.obj'CL.loc have.1sg put.PTCP
'I have not put it on the table.' [Catalan]
(López 2009)

(8.8) $[_{AP}$ Intel.ligent$]_i$, no ho és.
intelligent not CL.adj be.3sg
'Intelligent / he is not.' [Catalan]
(López 2009)

Hanging Topic constructions only allow specific DPs to be dislocated, and neither case nor preposition is copied to the left. Example (8.9a) is a case of CLLD, whereas in (8.9b) we are dealing with HTLD.

(8.9) a. $[_{PP}$ A la campagne$]_i$, Paul n'y$_i$ reste jamais longtemps.
In the countryside Paul not'CL.loc stay.3sg never long
'In the countryside / Paul never stays long there.'

b. $[_{DP}$ La campagne$]_i$, Paul n'y$_i$ reste jamais longtemps.
the countryside Paul not'CL.loc stay.3sg never long
'In the countryside / Paul never stays long there.' [French]
(Doetjes et al. 2002)

Resumptive pronouns in CLLD are exclusively clitics while HTLD allows for an epithet (8.10a), as well as a tonic (8.10b) or a clitic pronoun (8.10c).

(8.10) a. Mario$_i$, non darò più soldi a quell'imbecille$_i$.
Mario not give.1sg.FUT more money to this-idiot
'Mario / I won't give more money to this idiot.'

b. Mario$_i$, non darò più soldi a lui$_i$.
Mario not give.1sg.FUT more money to him
'Mario / I won't give more money to him.'

c. Mario$_i$, non ne$_i$ parla più nessuno.
Mario not CL.part speak.3sg more nobody
'Mario / nobody speaks anymore of him.' [Italian]
(Cinque 1990)

Clitic left dislocated phrases are sensitive to island constraints[1] (8.11a) whereas hanging topics (8.11b) are not.

(8.11) a. *A Marie$_i$, je connais le flic qui lui$_i$ a retiré son permis.
To Marie I know.1sg the cop who her have.3sg take-away.PTCP her license
'Marie / I know the cop who took away her driver's license.'

b. Marie$_i$, je connais le flic qui lui$_i$ a retiré son permis.
Marie I know.1sg the cop who her have.3sg take-away.PTCP her license
'Marie / I know the cop who took away her driver's license.' [French]
(Doetjes et al. 2002)

1 Complex DPs with a CP complement are "islands", out of which extraction (A'-movement) is not possible (cf. Ross 1967; Chomsky 1973 and section 8.1.1 above).

Cinque's criteria carry over to other Romance languages, though note that there are subtle differences between left detachment constructions across languages. For instance, Cinque restricts the number of HTLD phrases to one, while in French there are no restrictions in this sense.

In other languages than Romance, the distinction between CLLD and HTLD has similar correlates. In English, there are mainly two constructions, Topicalization and Left Dislocation. German and Dutch have Contrastive Left Dislocation and Left Dislocation. However, it seems that there are no one-to-one correspondences of constructions across languages. Frey (2005) notes that German Left Dislocation corresponds more to English Topicalization while German HTLD is more similar to English Left Dislocation. Roughly, the English type of Left Dislocation can be regarded as HTLD, and the canonical Romance type as CLLD (cf. van Riemsdijk 1997).

8.1.4 Clitic Left Dislocation in Spanish

In this chapter, I will concentrate on (clitic or non-clitic) Left Dislocation in Spanish, although many of the observations can be applied to other Romance languages and to equivalent constructions in other languages, as well. Note, however, that there are distributional, formal, and functional divergences of Left Dislocations across languages. Furthermore, I will restrict my examinations to the case of left dislocated DPs, which is by far the most frequent type of CLLD.

A typical Spanish example is presented in (8.12). In this example and from now on, the dislocated phrase is underlined and the resumptive pronoun is set in **bold face**. Additionally, in case that there is a bridging anchor, it is underlined by dots.

(8.12) a. Juan preparó la comida.
Juan prepare.3sg.PAST the meal
'Juan prepared the meal.'

b. La carne, **la** quemó.
the meat CL burn.3sg.PAST
'The meat, he burned it.'

c. Las verduras, **las** olvidó.
the vegetables CL forget.3sg.PAST
'The vegetables, he forgot them.'

In utterance (8.12b), the noun phrase "la carne" ("the meat") is moved to the left periphery, and the clitic pronoun "la" is left behind. An utterance without CLLD in canonical form "Quemó la carne." ('He burned the meat.') would also be acceptable, though it fits differently into the surrounding discourse. It seems that the dislocated phrase must be connected somehow to a preceding utterance. It is this difference that I attempt to explain in the remainder oft this chapter.

While in many other languages the use of Left Dislocation is restricted to informal conversations involving colloquial registers, CLLD in Spanish occurs both in written and in spoken language. In written language it is not very frequent, though it can be found already in 19the century texts (8.13).

(8.13) [...] cultiva, ensancha, agita y desenvuelve en el ánimo de sus lectores el odio a los europeos, el desprecio de los cuerpos que quieren conquistarnos. A los franceses, **los** llama titiriteros, tiñosos; a Luis Felipe, guarda chanchos, unitario, y a la política europea, bárbara, asquerosa, brutal [...]

'[He] grows, expands, agitates and triggers in his readers' soul the hate for Europeans, the contempt for the bodies that pretend to conquer us. The French / he calls them puppet players, scabby; Luis Felipe he calls pig shepherd, unitarian, and European politics he calls barbarian, disgusting, evil [...]'

(Sarmiento, Domingo Faustino (1850): Facundo [from Davies, 2002])

On the other hand, in spoken Spanish it is a widely used utterance organizing device. It is very frequent in spontaneous spoken discourse, but also in rather formal genres such as radio and television, academic and political speeches (8.14), etc.

(8.14) De todas maneras la reflexión sobre el fondo de la política parlamentaria de nuevo **la** pongo... o sobre la legitimidad democrática, **la** pongo en relación con lo que he dicho, nosotros hemos defendido...

'Anyway, I would like to relate again the reflection on the backgrounds of parlamentary politics, or on democratic legitimicity, to what I have already said, we have defended...'

(Felipe González, 21-12-91, Marín 1992, pol/ppol008f.asc (cinta 008))

The data I will take as a basis for the further investigations in this chapter is extracted from corpora of both spoken and written naturally occurring Spanish discourses. Examples are taken from the "Corpus de referencia de la lengua española contemporanea (CRLEC)" (Marín, 1992)[2], from the "Corpus del Español" (Davies, 2002)[3], from the "Corpus de referencia del español actual (CREA)" (Real Academia Española, 2008)[4], and from the spoken language corpus "El habla de la ciudad de Madrid" compiled by Esgueva and Cantarero (1981).

8.2 Discourse Functions of Left Dislocation

In the literature on Left Dislocation, a great amount of work deals with its discourse pragmatic functions (cf. *inter alia* Keenan and Schieffelin 1976a; Ziv 1994; Prince 1997; Hidalgo Downing 2001; López 2009).

In earlier works, Left Dislocation is characterized as an indicator of topicality. A problem with this view is that the notion "topic" can be conceived in diverse ways (cf. section 4.3). As Prince (1997) points out, a conception of sentence topic as the initial part of a sentence leads to a circularity in the treatment of Left Dislocation. A view that takes topic as "what the sentence is about" (Reinhart, 1982) is intuitively appealing, but is faced with certain difficulties. López (2009) discusses at length the tests for topicality proposed by Reinhart

2 http://www.lllf.uam.es/corpus/corpus_oral.html
3 http://www.corpusdelespanol.org
4 http://corpus.rae.es/creanet.html

and concludes that dislocations in Romance are not well characterized as aboutness-topics. Just to mention some issues, there can be more than one dislocation in a sentence (8.15). What the sentence is about in this case?

(8.15) Raúl, a mi prima, no **se** **lo** presentes.
Raúl to my cousin not CL.obl CL.acc present.3sg
'Raúl / to my cousin / do never introduce him to her.' [Spanish]
(Elena Valdés Luxán, p.c.)

Moreover, it is difficult to state how the following sentence can be about "intelligent".

(8.16) El Joan, intel.ligent, no **ho** és gaire.
the Joan intelligent not CL.adj be.3sg much.
'John is not very intelligent. [Catalan]
(López 2006)

Another point made by Prince is that it is difficult to define exactly what is about what. However, as we will see below, a treatment of these sentences in terms of Vallduví (1992)'s *link* is possible. Before, I will discuss the three distinct discourse functions that Prince (1997) attributes to Left Dislocation.

First, Left Dislocations serve to simplify discourse processing by removing a discourse-new entity from a position in the clause which favours discourse-old entities, and replacing it with a Discourse-old entity, i.e. a pronoun. A separate processing unit is created for the dislocated phrase. This can be seen as a meta-discourse function in the sense that the incrementation of the discourse model under construction is facilitated without affecting the contents of the model.

Second, Left Dislocations mark an entity as already evoked in the discourse or as standing in a certain relation to an already evoked discourse entity. In this case, Left Dislocation triggers an inference that the entity represented by the dislocated phrase stands in a salient partially-ordered set relation[5] to some entity already present in the discourse model. Prince calls this a true discourse function in that substantive aspects of the discourse model being constructed are signalled, in particular mereological relations among entities in the discourse model. I will come back to this important function below.

A third property of Left Dislocations is mentioned by Prince. They can serve to amnesty an island violation in enabling the completion of a sentence that would otherwise be grammatically ill-formed. I should note here that this function can only be fulfilled by Hanging Topic Left Dislocations (HTLD) and not by CLLD, since CLLD is sensitive to island constraints (see section 8.1 above).

As Ziv (1994) points out, the first two functions indicate the discourse organizational nature of Left Dislocations. Left Dislocations indicate that the discourse is shifted in a particular direction. That means on the one hand that the felicity of the construction depends on the preceding discourse and on the other hand that there are consequences for the following discourse. I will examine both points in more detail in the next two sections.

5 Partially-ordered sets ("posets") (Hirschberg, 1991; Prince, 1997) are defined by a partial ordering R on a set of entities S, such that for all elements of S, R is either reflexive, transitive, and antisymmetric or, alternatively, irreflexive, transitive, and asymmetric. Poset relations include, besides the identity relation and the usual set relations, relations like *part-of* and *subtype-of*. It should be noted here that neither all types of *part-of* relations nor the *member-of* relation are necessarily transitive. See the discussion in chapter 6, section 6.2.1.

8.2.1 CLLD and Familiarity

The first of Prince's functions only applies to left dislocated discourse-new entities. However, as various corpus-based investigations reveal (Hidalgo Downing, 2001; Padilla García, 2001), dislocated constituents in spoken Spanish can have any familiarity status (cf. Prince 1981). It can (i) resume or readopt given information, (ii) draw the attention to inferrables, or (iii) introduce new information.

8.2.1.1 Given Entities

Given information can be expressed by the phorical demonstrative pronouns "esto", "eso", and "aquello" ("this / that"), as exemplified in (8.17). The phorical element connects the dislocation with the preceding discourse (unless it is not deictically used) and in addition, it summarizes the already conveyed content. But also referential expressions in form of full noun phrases (8.18) can refer to given information.

(8.17) No quiero que lo dejemos. Eso sí que **lo** tengo claro.
not like.1sg that CL let.1pl that yes that CL.obj have.1sg clear
'I don't want that we stop it. That's what I am sure about.' (Padilla García 2001, p. 258)

(8.18) [Context: An abbreviation ("SRM") was mentioned before, and the speaker tries to remember it.]

...tres siglas, a ver si **las** acierto.
three letters to see if CL.obj.pl spot.1sg

'Three letters / let's see if I spot them.' (Padilla García 2001, p. 259)

8.2.1.2 Inferrable Entities

A large part of the occurrences of Left Dislocations involve inferrable entities, i.e. they are cases of indirect or bridging anaphora. The relation between dislocated element and some entity mentioned before (the bridging anchor) can be of any of the kinds of bridging relations identified in chapter 6, section 6.2. In example (8.19), the referent of the dislocated element stands in a *member-of* relation to the models mentioned in the first sentence. Similarly, a *subset* or *subcollection* relation is established in (8.20). In (8.21), there is a *part-of* relationship between the meniscus and the leg. More precisely, it is a *component-integral object* relation. Note that here the direction of the relation is inverse, i.e. the before-mentioned meniscus is part of the dislocated referent.

(8.19) a. La gente estaba asombrada con muchos de los modelos que tenías colgados por allí.

'The people were amazed by many of the models you had hanging around.'

b. Uno de los modelos hasta me **lo** llevé yo a mi casa.
one of the models until CL.poss.1sg CL.obj take.1sg I to my house
'One of the models / I took it home.'

(Marín 1992, pub/apub004c.asc (cinta 004))

(8.20) a. A: ¿A conocer el caviar? Ah, pues sí. Eso...

'To get to know the caviar? Ah, well. That ...'

b. B: Uno de esos vicios...

'One of those vices.'

c. A: Un vicio más. Un vicio más. Luis García Berlanga...

'One more vice. One more vice. Luis García Berlanga...'

d. B: Otros vicios no me **los** ha enseñado. Los guarda todavía
other vices not me CL.obj.pl have.3sg show.PTCP them keep.3sg still
en secreto.
in secret

'More vices / he didn't show them to me. He keeps them still in secret.'

(Marín 1992, deb/adeb002a.asc (cinta 002))

(8.21) Me he roto el menisco. [...] La pierna, **la** tengo muy bien.
me have.1sg break.PTCP the meniscus the leg CL.obj have.1sg very good

'I broke my meniscus. The leg / it is very well.'

(Hidalgo Downing 2001, p. 270)

Apart from these mereological relations, also other types of bridging relations can hold between the dislocated referent and an anchor mentioned before. In example (8.22), the dislocated noun phrase "las clases" ("the classes") stands in a close relationship to "aprender a esquiar" ("to learn skiing") mentioned in the first sentence. In terms of FrameNet (cf. section 7.1), the occurrence of "aprender" ("to learn") evokes the frame *Education_Teaching* including a frame element *Course*, which is instantiated by "las clases". Similarly, the serial mentioned in (8.23a) evokes a frame containing the concept of its idea, to which the dislocated expression in (8.23b) refers to.

(8.22) a. A: ¿Es verdad eso que se dice de que los niños canarios van a aprender a esquiar?

'Is that true that they say that the Canarian children are going to learn skiing?'

b. B: Pues sí, y tanto... y tanto que es verdad. Esta mañana se ha presentado este asunto, se llama "la Semana Blanca Canaria en el Pirineo". Toma ya. [...]

'Indeed, and so much... so much that it is true. This morning, this subject was presented, it is called "The White Canarian Week in the Pyrenees". Take that.'

c. A: Oye, una cosa, pero las clases no **las** van a dar en el Teide, me decías en el Pirineo.

'Listen, one thing, but the classes, they won't give them on the Teide, you said in the Pyrenees.'

(Marín 1992, conv/ccon012b.asc (cinta 012))

(8.23) a. Es el tercer espacio de la serie que estamos dedicando a la alimentación y su relación con la salud.

'It is the third space of this serial that we are dedicating to alimentation and its relationship to health.'

b. La idea de esta serie se **la** debemos a Socorro Calvo, profesora de bioquímica en la UNED.

'The idea of this serial, we owe to Socorro Calvo, professor of biochemistry in the UNED.'

(Marín 1992, doc/adoc012a.asc (cinta 012))

The bridging relation between the referent of the dislocated phrase and its anchor can be locative, as well. In (8.24a), "some prehistorical ruins" are introduced, and the dislocated DP in (8.24b) stands in a locative relation *object/place* to them.

(8.24) a. Una desviación de dos kilómetros conduce a unas ruinas prehistóricas excavadas por la Escuela Británica de Atenas.

'A detour of two kilometres leads to some prehistorical ruins that were excavated by the British School of Athens.'

b. El lugar excavado **lo** vinculan los arqueólogos con la población que fundó la primera ciudad de Troya.

'the excavated site / the archeologists connect it to the settlement that founded the first city of Troya.'

(El País, 30-10-2004)[6]

As we can see from these examples, the bridging relations occurring in connection with left dislocated referents can be of any type, and not only mereological or "poset" relations, as claimed by Prince (1997) or López (2009).

8.2.1.3 New Entities

Contrary to claims made in some earlier theoretical works, even new information can occur in left dislocated positions. In (8.25), a referent is freshly introduced without any mention in the preceding discourse. Similarly, (8.26) is an instance of the use of a Left Dislocation to start a story, obviously involving the first mention of Juana.

(8.25) Los cantos de sirena de Telefónica no me **los** quiero oír.
the singing of siren of Telefónica not me CL.obj.pl want.1sg hear
'The siren's calls of Telefónica, I don't want to hear them.'

(Padilla García 2001, p. 259)

(8.26) A Juana, **le** dieron un premio.
A Juana CL.obl give.3pl a prize
'Juana / they gave her a price.' (Ricardo Apilánez Piniella, p.c., 19-05-2007)

6 http://elviajero.elpais.com
/articulo/viajes/isla/despierta/pasiones/elpviavia/20041030elpviavje_3/Tes

While Padilla García (2001) only distinguishes between new and given information and does not regard inferrable entities at all, Hidalgo Downing (2001) found in her study of two corpora of spoken Spanish (Esgueva and Cantarero, 1981; Marín, 1992) that 47% of all occurrences of dislocated noun phrases referred to already evoked entities, 33% involved inferrable entities, and 14% of the cases introduced new entities. In sum, referents of left dislocated expressions can have any familiarity status, though most are evoked or inferrable.

8.2.2 CLLD and Discourse Topic

In this section, I will examine the relationship between left dislocated referents and the notion of discourse topic (cf. the discussion in chapter 4, section 4.3). According to various scholars of discourse analysis (cf. e.g. Keenan and Schieffelin 1976b), there are two basic strategies of discourse sequentiality, discourse continuity and discontinuity, which correspond to the intentions of discourse participant of either maintaining or changing the subject matter of a conversation. Discontinuity is expressed by the change of the discourse topic. Hidalgo Downing (2001, p. 283) distinguishes four types of topic change: (i) topic introduction, (ii) progressive topic change, (iii) topic reintroduction, and (iv) topic closure. Discourse continuity consists of topic incorporation or repetition. The distribution of the discourse functions regarding the discourse topic in the corpus study of Hidalgo Downing (2001) is given in Table 8.2.

Discourse strategy		Occurrences	
topic change		68%	100%
	topic introduction		36%
	progressive change		52%
	topic reintroduction		6%
	topic closure		7%
topic continuity		32%	100%
	incorporation		49%
	collaborate topic		24%
	repetition		27%

Table 8.2: Distribution of CLLD w.r.t. discourse topic (Hidalgo Downing, 2001)

8.2.2.1 Topic Change

Topic Introduction

As revealed by this data, topic change is the primary function of Left Dislocations. In the first place, a topic is changed when a new topic is introduced. The most obvious occurrence of Left Dislocation used to introduce a new discourse topic is in the beginning of a story. Consider for example (8.27), where the Left Dislocation contains the first mention of Juan.

(8.27) A Lorca, **lo** fusilaron en el año 1936.
A Lorca CL.obj shoot.3pl in the year 1936
'Lorca / they shoot him in 1936.' (Ricardo Apilánez Piniella, p.c., 19-05-2007)

The introduction of a new discourse topic is often accompanied by the use of meta-discourse markers like "now", "by the way", or "listen", as illustrated by (8.28).

(8.28) [Context: A and B are hanging a curtain on the wall.]

a. A: Sujétamela así.
hold.3sg-CL.obl.1sg-CL.obj so
'Hold it like that.'

b. B: Oye y la revista esta, ¿cada cuánto tiempo **la** sacáis?
listen and the magazine this every how-much time CL.obj publish.2pl
'Listen, and this magazine / how frequently do you publish it?'
(Marín 1992, conv/ccon013d.asc (cinta 013))

Progressive Topic Shift

The most frequent topic changing strategy is what Hidalgo calls "progressive topic shift" ("cambio progresivo de tópico"). It covers more than a third of all cases of Left Dislocations in the corpus of Marín (1992). It is a very natural device of continuing a conversation, letting it flow smoothly from one subject to another without making sudden changes.

As Hidalgo (*op.cit.*, p. 304) notes, this is the function of Left Dislocations that corresponds most closely to its informational status. Thus, a progressive topic change is signalled by the use of an inferrable entity, which we have seen to share properties of both given and new information (see the discussion in chapter 6). In the following conversation, CLLD is used in (b) to shift the topic from the clothes introduced in (a) to one particular piece of clothing, which is then resumed in (d) by means of another occurrence of CLLD. The bridging relation involved in (b) is a *member-of* relation between the coat and the cloths in the store, while in (d) an *identity* relation holds between dislocated phrase and the before-mentioned coat.

(8.29) [Context: A and B talk about A's cloth store.]

a. B: La gente estaba asombrada con muchos de los modelos que tenías colgados por allí.
'The people were amazed by many of the models you had hanging around.'

b. B: Uno de los modelos hasta me **lo** llevé yo a mi casa.
one of the models until CL.obl.1sg CL.obj take.1sg I to my house
'One of the models / I took it home.'

c. A: Sí, picaste.
'Yes, you took the bait.'

d. B: Este mismo abrigo, [A: affirmative gesture] **lo** ves colgado en
this same coat CL.obj see.2sg hang.PTCP in
una casa con un nombre más caro.
a house with a name more expensive
'This very coat / you see it hanging around in stores with more expensive names.'
(Marín 1992, pub/apub004c.asc (cinta 004))

Hidalgo attributes an important argument for the view that the use of inferrable entities in Left Dislocations corresponds to a strategy of topic shift (and not topic continuity) to Hobbs (1990). A progressive topic shift can be indicated by discourse markers that are similar to those which mark a topic introduction. Moreover, as we have discussed in chapter 4 (section 4.3), often it is not possible to make out a precisely defined discourse topic until the conversation is over. The use of inferrable entities does not continue a topic; rather a discourse topic is created by constructing it from various subtopics.

Topic Reintroduction

A function that is often considered typical for Left Dislocations is the reintroduction of a discourse referent which has been mentioned before but has lost attention in the course of the conversation. A referent from the background is brought back into the foreground. Topic reintroduction seems to fulfil the optimal conditions of referent accessibility in marked topical positions. However, in Hidalgo Downing (2001)'s study, only a small part of all Left Dislocation were used to reintroduce a topic. In the attested cases, a considerable distance between first and second mention of a dislocated referent is observed. Topic reintroduction shares with the introduction of new topics the fact that a shift or a change regarding the immediately preceding discourse segments is presupposed. The course of conversation is modified and the Gricean Maxim of Relevance is violated at a local level. In Spanish, a Clitic Left Dislocation with a dislocated phorical element "eso" ("that") often reintroduces and summarizes a discourse topic, as in example (8.17), repeated here as (8.30).

(8.30) No quiero que lo dejemos. Eso sí que **lo** tengo claro.
not like.1sg that CL.obj let.1pl that yes that CL.obj have.1sg clear

'I don't want that we stop it. That's what I am sure about.'

(Padilla García 2001, p. 258)

Topic Closure

Another peripheral function of Left Dislocations is to close a given discourse topic. It is a marked form of concluding a subject. In normal conversations, a subject is talked about and continously changed by progressive shifts. When a topic is exhausted, the conversation terminates without the need to close a topic explicitly. However, some cases can be found, as (8.31) illustrates.

(8.31) a. Mi posición prodría ser o trabajar fuera de casa o trabajar dentro, yo he escogido dentro de casa.

'My position could be either working outside or at home; I have chosen at home.'

b. Mi vida **la** dedico a hacer hogar.
my life CL.obj dedicate.1sg to make.INF household

'My life / I dedicate it to house-keeping.'

(Marín 1992, doc/cdoc008a.asc (cinta 008))

8.2.2.2 Topic Continuity

Although Left Dislocations generally seem to mark topic changes or shifts, a considerable amount of the cases in Hidalgo's study (32%) can be attributed to a strategy of discourse continuity. As discourse continuity can be seen as the default case in a conversation (for a detailed account of this view, see Jasinskaja 2006), it can be identified whenever there are no signs of topic change. In difference to topic change, which can be identified thanks to linguistic expressions and constructions signalling modifications of the course of a conversation, there are no specific forms which mark topic continuity. Even though, topic continuity is characterized by one or more of the following aspects: referential continuity, temporal continuity, and the presence of discourse markers which can signal continuity, e.g. "and", "also", "while", "moreover". Following Keenan and Schieffelin (1976b), Hidalgo Downing (2001) makes out three discourse functions of continuity: incorporation, collaborative topics, and repetition.

Incorporation means that a part of the immediately preceding discourse segment is included or incorporated in a new segment that is created as an expansion or continuation of the former segment. In example (8.32), the speaker first introduces "la comprensión de la realidad en el entorno", and then constructs the second utterance with a Left Dislocation, indicating that this referent is the discourse topic. The continuity is enhanced by the conjunction "y" ("and") between the segments.

(8.32) a. Porque en esta etapa de la vida de diez años, está aumentando la comprensión de la realidad en el entorno,

'Because in this phase of life of ten years, the comprehension of the surrounding reality is increasing,'

b. y esta comprensión puede ... cualquiera de nosotros, cuando analizamos nuestra vida, podemos ver**la** a la vez eh... con un gran optimismo y con gran euforia...

'and this comprehension can ... everyone of us, when we analyze our life, we can see it at the same time ... with a great optimism and with great euphoria.'

(Marín 1992, deb/adeb033a.asc (cinta 033))

Note that this case of topic continuity can actually be seen as an instance of a progressive topic shift. Recently introduced discourse referents are installed as potential future discourse topics.

In the other forms of discourse continuity, collaborate topics and repetitions, a speaker continues the immediately preceding contribution. While a collaborate topic involves an interactive process of two speakers (8.33), a repetition does not necessarily involve more than one speaker. Here, the speaker does not pretend to promote an already mentioned referent and elevate it to a topic, but the current discourse topic is maintained. Often, a special emphasis on the dislocated referent is expressed, as we can see in example (8.34).

(8.33) a. A: Y nada, no me ha puesto nada de trabajo, sólo una hoja de apuntes. Y eso.

'Anyway, he didn't give me any work, only a sheet of notes. That's all.'

b. B: Y los apuntes **los** tomáis en clase, o sea ...
and the notes CL.obj take.2pl in class or be.3sg
'And the notes, you take them in class, I mean ...'

c. A: Sí.

'Yes.'

d. B: ... que tenéis clase teórica.

'... you have a theoretical class.' (Marín 1992, con/ccon018b.asc (cinta 018))

(8.34) a. A: Total. Un diseño precioso.

'Superb. A beautiful design.'

b. B: Ése es muy bonito. Y ese abrigo **lo** teníamos ...
that be.3sg very nice and that coat CL.obj have.1pl.IPFV ...
'This is very nice. And that coat, we had it ...'

c. A: ¿A cuánto está?

'How much is it?'

d. B: Ese abrigo, mira **lo** teníamos en doscientas sesenta y nueve mil.

'This coat, look, we had it at two hundred sixty-nine thousand.'
(Marín 1992, pub/apub004c.asc (cinta 004))

8.2.2.3 Conclusion

As we have seen in this section, speakers fundamentally use Left Dislocations in order to change the discourse topic and to mark discourse discontinuity. In particular, the most preferred functions are topic introduction and progressive topic shift. Apart from topic introduction, which is the case when a new referent appears in a left dislocated position, the largest part of the occurrences of Left Dislocations involve direct and indirect anaphora to already mentioned discourse referents. However, Left Dislocations are not exclusively used to mark a change, but they can occur also in contexts of discourse continuity. Hidalgo suggests that the diversity of contexts in which Left Dislocation occurs can be explained in terms of discourse coherence. On the one hand, in a strategy of topic shift, a global aspect of coherence is emphasized, in the sense that referents which stand in some relation to former subjects of a conversation are (re-)introduced. On the other hand, strategies of topic continuity create links between adjacent discourse segments and are thus a means to express local discourse coherence.

8.2.3 CLLD, Contrast, and Constraints on Discourse Structure

8.2.3.1 CLLD and Contrast

The above mentioned corpus studies (Padilla García, 2001; Hidalgo Downing, 2001) agree on the conclusion that a contrastive effect often observed with CLLD is independent of its familiarity status. For instance, Padilla García (2001) writes that

> "El valor contrastivo [...] no depende exclusivamente de la posición, ni de la aplicación del rasgo +/-nuevo al segmento dislocado, sino de la relación que el

> elemento mantenga con otros elementos, estén éstos presentes o no en el discurso."[7] (Padilla García 2001, p. 260)

There is an ongoing discussion in the literature on the subject whether CLLD itself has a contrastive semantics or not[8]. Some authors, like Chafe (1976), Arregi (2003), and López (2009) claim that CLLD is contrastive by itself. Others, like Prince (1997) and Brunetti (2007) argue against this view. In the following, the two views are presented.

López (2009) attributes two pragmatic functions to CLLD. First, dislocated elements are *anaphoric*, that is there is an anaphoric connection to some previously introduced discourse entity. This property is assumed by most studies on Left Dislocation, so it can be found in Frey (2005). It applies to all instances of Left Dislocation, except for the cases when discourse-new entities are introduced by this construction. López assigns a feature [+a] to the dislocated phrase. We will see below in section 8.3.1 how this kind of feature contributes to the syntactic derivation of sentences involving CLLD.

Second, López assigns a feature [+c], *contrastive*, to dislocated phrases. He follows the analysis made by Arregi (2003), who sees CLLD as contrastive topicalization. This property implies that the construction evokes a set of alternatives to the entity denoted by the dislocated phrase. The entity denoted by the dislocated constituent is that element of the alternative set which makes the predication of the sentence true.

Prince (1997) rejects the claim that contrast is an inherent property of Left Dislocation. She says that

> "[...] contrast is not a primitive notion but rather arises when alternate members of some salient set are evoked and when there is felt to be a salient opposition in what is predicated by them."
>
> (Prince 1997, p. 125)

Brunetti (2007) presents a more detailed account of this view. Following Vallduví (1992), she claims that (clitic) Left Dislocation in Romance has the discourse property of a *link*. In Vallduví's approach, the information structure of a sentence is divided in a *focus* and a *ground*, and the ground consists of a *link* and a *tail*. Focus is the informative part of the sentence, and the link is an expression that directs the hearer to a given address in his mental discourse representation. The information carried by the utterance is entered under this address. In Left Dislocations, the detached phrases are links: "they have the 'aboutness' feeling typical of linkful structures and satisfy the poset relation condition on preposed phrases" (Vallduví, 1992, p. 109). Vallduví analyzes Catalan data like the following example (8.36). In clause (b), "Joan" is the link, and it stands in a poset relation

7 "The contrastive value [...] does not exclusively depend on the position, nor on the application of the feature +/-new to the dislocated segment, rather on the relationship which the element may have with other elements, be them present or not in the discourse."

8 Note that similar constructions in other languages, e.g. Left Dislocation in German (8.35), neither have a unequivocally contrastive reading (Frey, 2005).

(8.35) a. Die Kinder hatten heute ihren ersten Ferientag.
The children had today their first holiday
'Today, the children had their first holiday.'

b. Den Otto, den hatte Maria abgeholt.
The.acc Otto PRON.acc had Maria come-for
'As for Otto, Maria had come for him.'

member-of to the plural entity "Joan and Isidora" mentioned in clause (a). Thus, this analysis is compatible with Prince's second discourse function of Left Dislocations that they mark an entity as evoked or inferrable.

(8.36) a. Quant al Joan i la Isidora no t'ho sé dir,
as-for the Joan and the Isidora no CL.obl'CL.obj know.1sg say

b. doncs el Joan [$_F$ **el** veiem ben poc].
since the Joan CL.acc see.1pl quite little

'As for Joan and Isidora, I can't say, since Joan / we see him very little.'
[Catalan]

As Brunetti claims, a link always implies the existence of an alternative set: the hearer has to select the address among a set of possible ones in the relevant context. Brunetti further argues that a link can have a contrastive interpretation or not, depending on the context. On the one hand, in a non-contrastive interpretation, the contextual alternatives to the link are simply not taken into account by the hearer. This is the case when the dislocated phrase is really discourse-new, e.g. when CLLD is used to start a story. On the other hand, in a contrastive interpretation, the members of the alternative set are contrasted with each other.

To conclude the discussion of contrastiveness of CLLD construction, it can be seen that no matter whether contrast is seen as an inherent property or as arising from independent grounds, a set of alternatives seems to be involved.

I claim that it is the discourse structure that determines the felicity of CLLD constructions in the sense that a contrastive interpretation presupposes the existence of subordinating discourse relation in the discourse context. To gather arguments for this view, let us look at the constraints imposed by Left Dislocation on the surrounding discourse structure.

8.2.3.2 Constraints on Discourse Structure

López (2009) discusses the question in what kind of contexts CLLD is felicitous. Consider the following Spanish texts which correspond to López' Catalan data:

(8.37) a. Juan ha traido los muebles en un camión.
Juan have.3sg bring.PTCP the furniture in a truck.
'Juan brought the furniture in a truck.'

b. Abre el camión y
open.3sg the truck and
'He opens the truck and'

c. la mesa, **la** lleva a la cocina.
the table CL.obj bring.3sg to the kitchen
'the table / he brings it to the kitchen.'

This discourse conveys a Narration relation between (b) and (c), as indicated by the discourse marker "y" ("and") and the temporal sequence of events. In contrast, (8.38) exhibits a Narration relation between (b) and (c), and (d) is an Elaboration of (c).

(8.38) a. Juan ha traido los muebles en un camión.
Juan have.3sg bring.PTCP the furniture in a truck.
'Juan brought the furniture in a truck.'

b. Abre el camión y
open.3sg the truck and
'He opens the truck and'

c. empieza a subirlos a casa.
start.3sg to put-CL.obj to house
'starts to put them into the house.'

d. La mesa, **la** lleva a la cocina.
the table CL.obj bring.3sg to the kitchen
'The table / he brings it to the kitchen.'

López argues that discourse (8.37) is infelicitous while discourse (8.38) is perfectly acceptable. His explanation for this kind of data is as follows. Because in (8.37) only a coordinating NARRATION discourse relation can be established, there is no justification for the use of CLLD in this context. But in discourse (8.38), the subordinating ELABORATION relation allows to identify the dislocated phrase "la mesa" ("the table") as part of "los muebles" ("the furniture") mentioned in (a) and (c). As a consequence, it seems that CLLD needs a subordinating discourse structure where the utterance containing it can be embedded.

However, many Spanish speakers do not judge (8.37) as bad, rather as not even marked. Indeed, a look at the SDRT graphs for these discourses reveals that both texts actually involve a subordination. Fig. 8.1a shows the graph for (8.37), which by CDP[9] can be rewritten as shown in Fig. 8.1b, involving a subordinating BACKGROUND relation between (a) and (c). Hence, (a) is on the Right Frontier of the disourse and, as a consequence of RFC, the referents in (a) become accessible for anaphoric reference in (c). For comparison, the graph for (8.38) is given in Fig. 8.2.

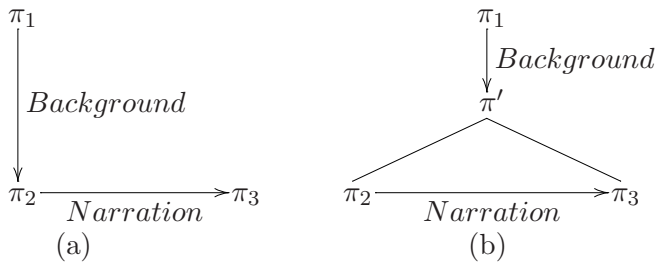


Figure 8.1: SDRT graph for (8.37)

9 CDP (Continuing Discourse Patterns) was defined in section 4.2.4 on page 121. In short, this principle says that coordinated constituents must behave homogeneously with respect to a third superordinated constituent.

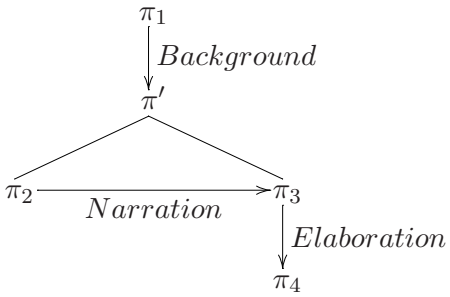


Figure 8.2: SDRT graph for (8.38)

Thus, López' data does not help much to exclude coordinating structures involving CLLD. However, in the data presented so far, we find instances of Elaboration (8.12), (8.13), Commentary (8.17), Background (8.19), (8.21), (8.23), (8.24), Explanation (8.36), which all are subordinating relations. This data tentatively supports López' claim that CLLD cannot occur in a coordinating discourse environment, but surely a more exhaustive corpus analysis is needed to confirm this hypothesis. In addition, what also has to be taken into account is the fact that CLLD can be used to introduce a new discourse referent viz. a new discourse topic.

What we can do for now is softening and inverting López' hypothesis to the effect that if CLLD occurs in a subordinating structure then it has a contrastive reading. Stated in this way, this constraint is perfectly compatible with Brunetti (2007)'s account.

Given the contrastive nature of subordinated CLLD utterances, it seems that the referent of the dislocated constituent is an element of an alternative set evoked by the superordinated constituent. It is this referent which makes the predication of the sentence true. In example (8.12), repeated here, the alternative set consists of parts of the meal introduced in (a), and in (b) it is true just for the meat that Juan burned it, not for the vegetables or any other part of the meal. Likewise, in (c) it is the vegetables that Juan forgot.

(8.12) a. Juan preparó la comida.
Juan prepare.3sg.PAST the meal
'Juan prepared the meal.'

b. La carne, **la** quemó.
the meat CL burn.3sg.PAST
'The meat, he burned it.'

c. Las verduras, **las** olvidó.
the vegetables CL forget.3sg.PAST
'The vegetables, he forgot them.'

8.2.3.3 Discussion

To conclude this section, I will repeat and discuss the discourse functions and constraints discussed. First, CLLD can involve given, inferrable, and new discourse referents. Second, it can be used to (re-)introduce a discourse topic, to change it progressively, to incorporate or repeat it, or to close it.

Until now, it was tacitly assumed that the two distincions are orthogonal. However, in order to see if the two distinctions are really independent from each other or pattern together in some way, a cross-classification can be made. In fact, in Fig. 8.3, only three different patterns of the distribution of Left Dislocation can be detected: first, topic introduction can occur with both new and inferrable entities; second, as already noted above, progressive change and (collaborate) topic incorporation pattern together allowing both given entities and cases of bridging. And so does topic closure, which, although it needs a known topic, this one can remain implicit before its explicit closure, cf. example (8.31). Third, topic reintroduction and repetition allow only given entities.

	new	inferrable	given
topic introduction	+	+	-
progressive change	-	+	+
topic reintroduction	-	-	+
topic closure	-	+	+
incorporation	-	+	+
collaborative topic	-	+	+
repetition	-	-	+

Table 8.3: Discourse functions of CLLD

As a consequence, taking the third group as subsumed by the second group (as the special case of bridging involving coreferentiality), two main discourse functions can be made out:

1. CLLD can introduce a new discourse topic involving a new discourse referent, or
2. CLLD can progressively change or continue a discourse topic, involving given or inferrable entities giving rise to a bridging anaphor.

In the second case, CLLD occurs in a subordinating environment and the anchor of the bridging anaphor triggered by CLLD must be evoked in the superordinating constituent. The dislocation has a contrastive reading if the anchor provides a suitable alternative set.

Let us relate these findings on contrastive CLLDs with the different notions of discourse topic discussed in section 4.3. The propositional topic of a contrastive CLLD is the superordinated segment. If the topic is conceived as a question, then the contrastive dislocation marks a subquestion. And if the topic is seen as an entity, then CLLD marks a new or modified discourse topic.

8.3 Semantics and Discourse Integration of Left Dislocations

The last part of this chapter is devoted to the question of how utterances containing clitic Left Dislocations are integrated in the surrounding discourse context. First, in section 8.3.1, a syntactic derivation of a sentence with CLLD is given a compositional semantic representation. Then, in section 8.3.2, this representation is adapted in terms of SDRT in order to fit with the preceding discourse. Finally, sections 8.3.3 and 8.3.4 spell out how mereological and frame-related bridging anaphora triggered by CLLD are established.

8.3.1 From Syntax to Semantics

In the literature on the syntax of CLLD, three main approaches can be found. Clitics can be seen as the spellout of a trace of the dislocated phrase (López, 2006, 2009). Another possibility is proposed by Uriagereka (1995), who states that the clitic and the doubled argument together form a 'Big DP'[10]. A third way is taken by approaches that treat dislocated phrases as base generated in the dislocated position (Cinque, 1990).

It is not a matter of this thesis to decide which theory is most suitable to account for CLLD. However, for the sake of explicitness, we will give an analysis compatible at least with the first two groups of theories[11]. A syntactic derivation of (8.12b) treating clitics as a spellout of a trace is illustrated in Fig. 8.3.

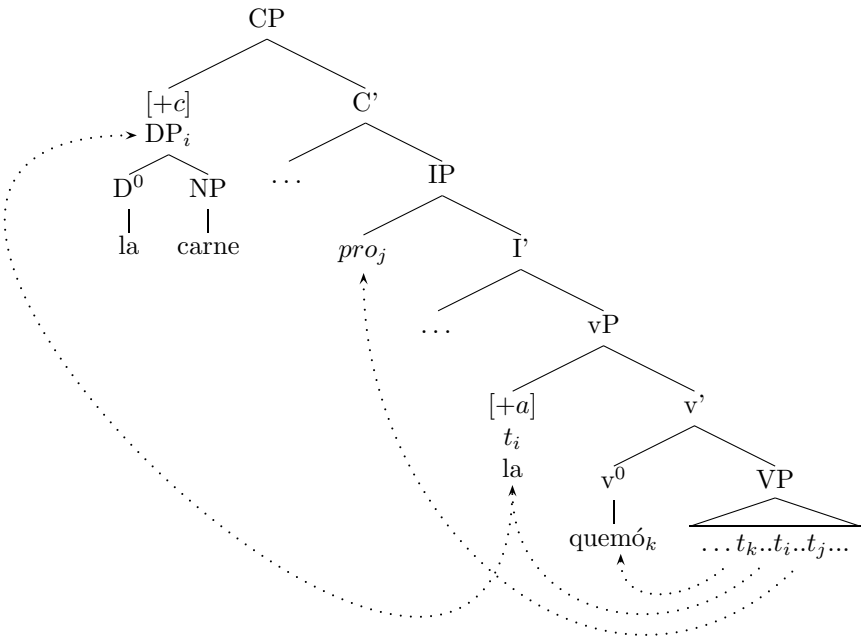
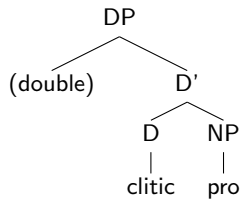


Figure 8.3: Syntactic derivation of example (8.12b).

Following López account, in a first step, CLLD involves movement of the DP to be dislocated to the specifier position of vP, where it gets the [+a] feature. At this position a trace t_i is left that is actually spelled out as a clitic pronoun. A second step to the specifier of CP is responsible for the assignment of [+c]. This derivation is made following mainly López (2009), but if one rejects the claim that [+c] is an inherent feature of CLLD, this derivation remains valid[12], just assuming that [+c] is assigned from independent grounds.

10 A 'Big DP' has the following structure:

(8.39)



11 See López (2009) or Villalba (2000) for further discussions.

12 In a cartographic approach *à la* Rizzi (1997), where CP is divided into (multiple occurrences of) phrases

Let us sketch how the meaning of a sentence with CLLD is computed compositionally according to this syntactic derivation. In chapter 6, it was argued that a definite DP involves a bridging presupposition, resulting in the following meaning characterization of a definite determiner as the head of the phrase (cf. (6.49)):

(8.40) $\lambda P \lambda Q[Q(\iota x[B(x,a) \wedge P(x)])]$

Starting from this representation, the meaning of sentence (8.12b) can be compositionally derived à la Heim and Kratzer (1998) as shown in Fig. 8.4.

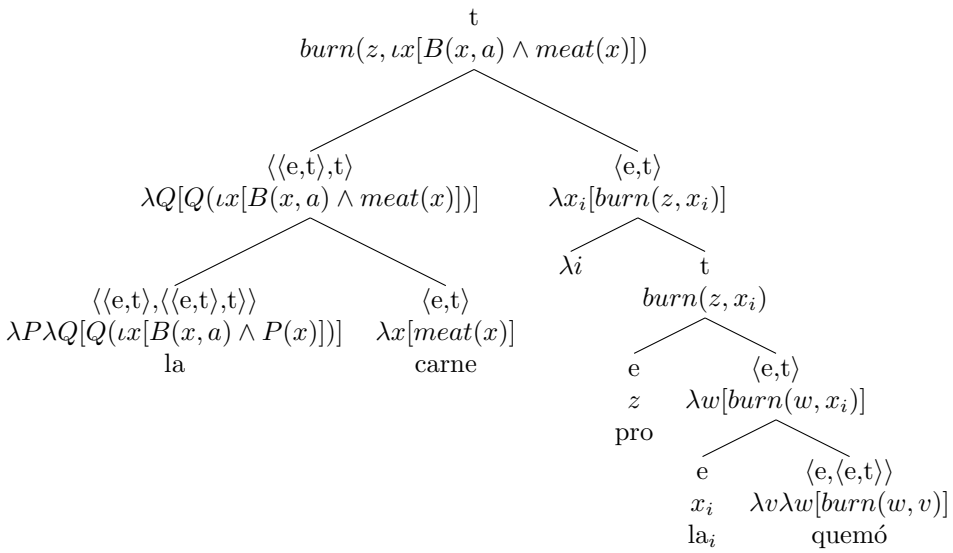


Figure 8.4: Meaning composition of example (8.12b).

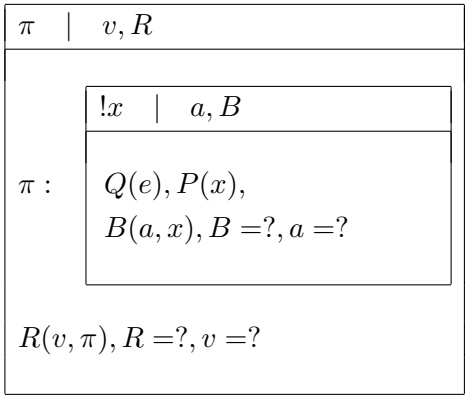
So far, I have spelled out how a left dislocated DP contributes to the meaning of a sentence. In essence, the dislocated constituent gets the normal semantic representation for a definite DP. However, as noted in the previous sections, the most important functions of Left Dislocations are not separable from discourse semantics.

8.3.2 Discourse Semantic Representation of CLLD

Recall from chapter 6 how the semantics of a definite DP is cast in terms of SDRT. There (cf. (6.51)), I have assumed, in line with Asher and Lascarides (1998a, b), that definite DPs carry two presuppositions: (i) Russell's uniqueness condition is extended by a bridging relation B to an anchor a (as in (8.40) above), and (ii) there is a discourse relation R between the current utterance and a previous utterance v containing the bridging anchor. The discourse representation of a definite DP, according to the extension of SDRT proposed in chapter 7, is repeated here:

ForceP and FinP, the CLLDed constituent attaches to the specifier of FinP.

(8.41) $\lambda e \lambda Q$



As assumed in the last section, a left dislocated DP gets the standard semantic representation for definites. However, the constraints on discourse structure imposed by the dislocation must be taken into account. A conclusion of section 8.2 was that CLLD, in some way or another, is related to a (possibly implicit) discourse topic. More precisely, if there is no obvious discourse topic, e.g. if CLLD is used to start a discourse, then the dislocation introduces a new topic. If CLLD is not used at the beginning of a text, then it shifts a given or inferrable topic. In this case, it gives rise to a bridging anaphor and the surrounding discourse must provide a superordinated constituent containing the bridging anchor. A contrastive effect arises if the anchor provides an alternative set in relation to which the dislocated referent is contrasted.

In SDRT, the discourse topic is conceived as a discourse constituent, which stands in a Topic relation (or "$\Downarrow$" in terms of Asher and Lascarides, 2003) to the segments it is topic of. If we take such a topic π' as the segment that is superordinated to the CLLD-containing segment and that contains the bridging anchor, then the following conditions on a CLLD-containing segment π can be stated: There is a relation $\Downarrow^* (\pi', \pi)$, and the underspecified discourse referent a introduced by the bridging condition can be identified with a bridging anchor α, which must have been introduced in segment π', i.e. $\pi' : \boxed{\alpha}$ (viz. $\alpha \in U_{K_{\pi'}}$). We will see below why we need the transitive closure of the topic relation, $\Downarrow^*$, which is recursively defined as follows:

(8.42) $\Downarrow^* (\pi_\alpha, \pi_\beta)$ iff $\Downarrow (\pi_\alpha, \pi_\beta)$ or $\exists \pi_\gamma [\Downarrow^* (\pi_\alpha, \pi_\gamma) \wedge \Downarrow (\pi_\gamma, \pi_\beta)]$

As for the bridging relation B, recall from section 6.2 that B can be either mereological, i.e. some kind of $part-of$ relation, or frame-related, i.e. a thematic or conceptual relation. For the first type of bridging relations, let us take again (as in section 7.3.2) $\sqsubseteq^+$ as a placeholder for any type of mereological relations, i.e. member/collection, subcollection/collection, portion/whole, substance/whole, or component/integral whole. For the second type, it was assumed in section 7.3.2 that the bridging relation is a (possibly underspecified) condition already (explicitly or implicitly) present in the discourse model. More formally, the following constraint on CLLD can be formulated:

Topic Constraint on CLLD

If $?(\pi, v, \lambda)$ and $CLLD(x)$ and $x \in U_{K_\pi}$ and $B(a, x) \in C_{K_\pi}$ then $\exists \pi', \alpha$ such that $\Downarrow^* (\pi', \pi)$ and $\alpha \in U_{K_{\pi'}}$ with either

a. $B = \sqsubseteq^+$ with $a = \alpha$ and $x \sqsubseteq^+ \alpha$, or

b. $B = \phi$ with $a = \alpha$ and $\phi(\alpha, x) \in C_{K_{\pi'}}$.

In short, a CLLD-containing segment can only be successfully attached to the preceding discourse if there is a topic π' that contains the bridging anchor α, which can be identified with the underspecified discourse referent a. The bridging relation B is to be resolved either as a mereological relation or as a thematic/conceptual relation already present in the discourse model. Note that resolving B to identity is allowed in both types of bridging relations. This is the correct resolution in case that the dislocated referent constitutes given information. The case of introducing a new discourse topic via CLLD emerges if π' is taken to be new and x percolates to π', resulting in a topic constituent where subsequent segments can attach to.

8.3.3 Resolving Mereological Bridging Anaphora

Let us examine how bridging anaphora imposed by dislocations are resolved in the process of pragmatic enrichment according to the proposed account. For the sake of explicitness, a detailed interpretation of (8.12), repeated here, will be given.

(8.12) a. Juan preparó la comida.
Juan prepare.3sg.PAST the meal
'Juan prepared the meal.'

b. La carne, **la** quemó.
the meat CL burn.3sg.PAST
'The meat, he burned it.'

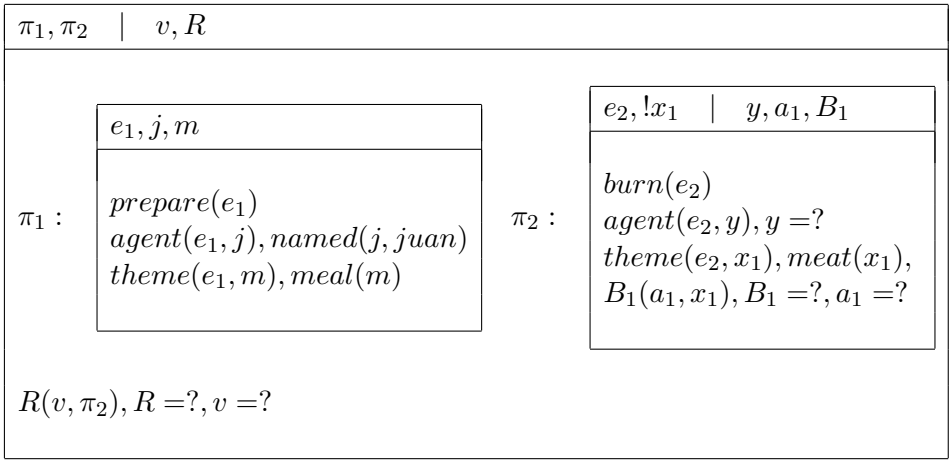
c. Las verduras, **las** olvidó.
the vegetables CL forget.3sg.PAST
'The vegetables, he forgot them.'

First, I want to concentrate on the question of what parts of world knowledge are needed to fill in the underspecified material, before examining how information encoded in FrameNet can help to establish bridging relations.

8.3.3.1 Building Bridges via CLLD

For the linguistic content of these utterances, according to a neo-Davidsonian style of event semantics, underspecified semantic representations as in (8.43) can be assumed. To resolve the references of the underspecified variables y, a_1, B_1, R, v, in addition to the semantics explicitly expressed by the linguistic input, apart from information about the discourse structure, some lexical knowledge and general knowledge about the world is necessary.

(8.43)



Let us first examine how the discourse relation R can be specified. Following the approach that Asher and Lascarides (2003, pp. 249–291) take (cf. also chapter 5 on page 157), lexical knowledge involves that "meat" and "meal" both are of a sort (or lexical type) *food*, "meat" being more specific, such that a subtype relation $\sqsubseteq$ holds between them. This knowledge can be expressed by a default rule (8.44). Moreover, it is part of basic world knowledge that a meal typically consists of a collection of dishes, e.g. soup, rice, meat, vegetables, dessert etc. So it can be assumed that if something is meat, then it can be part of a meal, possibly but not necessarily. More specifically, according to the taxonomy of mereological relations indicated in section 6.2.1, the relationship is a *member/collection* relation (8.45).

(8.44) $meat(x) \wedge meal(y) > x \sqsubseteq y$

(8.45) $meat(x) \wedge meal(y) > member\_collection(x, y)$

Similarly, "burning" a dish is a subtype of cooking in the sense of preparing a meal. Thus, if e_α is a cooking event and e_β is a burning event, then $e_\beta \sqsubseteq e_\alpha$ can be assumed (8.46).

(8.46) $prepare(e_\alpha) \wedge burn(e_\beta) > e_\beta \sqsubseteq e_\alpha$

Equipped with this knowledge, the preconditions for SDRT's *subtype* rule (cf. section 5.3.2.2, repeated here as (8.47)) are met.

(8.47) **Subtype**: $(\theta_i(x, e_\alpha) \wedge \theta_i(y, e_\beta) \wedge y \sqsubseteq x \wedge e_\beta \sqsubseteq e_\alpha) \rightarrow subtype_D(\beta, \alpha)$

Informally, this axiom states that if x and y play the thematic roles θ_i and θ_j in the eventualities e_α and e_β, respectively, and x is a subtype of y and e_α is a subtype of e_β, then a subtype relation holds between the corresponding discourse segments. Information about subtypes, in turn, provides clues for assuming an Elaboration relation (8.48).

(8.48) **Elaboration**: $(?(\alpha, \beta, \lambda) \wedge subtype_D(\beta, \alpha)) > Elaboration(\alpha, \beta, \lambda)$

From (8.44) we know that the sort (viz. lexical type) of "meat" is a subtype of that of "meal" ($x_1 \sqsubseteq m$), and from (8.46) we know that $e_2 \sqsubseteq e_1$. The eventuality e_1 has two participants and therefore two theta-roles: an *agent* and a *theme*. Obviously, the meal is the

theme of e_1, and the meat is the theme of e_2. So both have the same theta-role *theme*, and $\theta_1(e_1, m)$ and $\theta_1(e_2, x_1)$ hold with $\theta_i = theme$. Now we fulfill all preconditions of rule (8.47), and with (8.48) we can assume $Elaboration(\pi_1, \pi_2)$, resolving the underspecified parameter v to π_1. The presence of an ELABORATION relation, in turn, has certain consequences (cf. Asher and Lascarides, 2003, p. 460):

(8.49) $Elaboration(\alpha, \beta) \vdash \Downarrow (\alpha, \beta)$

The entailment (8.49) establishes $\Downarrow (\pi_1, \pi_2)$. We know from the topic constraint on CLLD that $\Downarrow^* (\pi', \pi_2)$. What about the relation between π_1 and π'? Both segments are superordinated to π_2. The assumption that the number of labels is to be minimized (part of the MDC, cf. section 5.3.2.5) gives reason to assume $\pi' = \pi_1$.

Having established $Elaboration(\pi_1, \pi_2)$, we can now turn to the resolution of the bridging relation. It might be assumed that ELABORATION relations imply mereological bridging relations by default. This assumption is supported by the data exhibited in this chapter, but unless it is confirmed by a large-scale corpus analysis, a frame-related bridging relation is not a priori excluded.

The topic constraint demands that the bridging anchor α must have been introduced in π_1, i.e. $a_1 = \alpha$ and $\alpha \in U_{K_{\pi_1}} = \{e_1, j, m\}$. For the resolution of a_1, there are thus the possibilities listed in Table 8.4. The first possibility ($a_1 = j$) leads to a sort conflict: the

a_1	B_1	$B_1(a_1, x_1)$	CONSISTENT
$a_1 = j$	*mereological*	$x_1 \sqsubseteq^+ j$	$*$
$a_1 = m$	*mereological*	$x_1 \sqsubseteq^+ m$	$\surd$
$a_1 = e_1$	*mereological*	$x_1 \sqsubseteq^+ e_1$	$*$
	frame-related	$agent(e_1, x_1)$ $x_1 = j$	$*$
	frame-related	$theme(e_1, x_1)$ $x_1 = m$	$(\surd)$

Table 8.4: Potential bridging relations in (8.43)

referent j standing for Juan as the agent of the cooking activity must be *human*, thus a subtype of *sentient*, which is not compatible with *physical_entity* as the sort of the referent x_1 standing for the meat. The same conflict arises in the fourth option ($a_1 = e_1$ with $x_1 = j$).

Similarly, the third possibility ($a_1 = e_1$ with $x_1 \sqsubseteq^+ e_1$) is ruled out: since x_1 is of sort *food*, which is a subtype of *physical_entity*, it cannot be a subtype of e_1, which is an *eventuality*.

We are left with two possibilities, which do not give rise to sort inconsistencies: the second one ($a_1 = m$ with $x_1 \sqsubseteq^+ m$) and the last one ($a_1 = e_1$ with $x_1 = m$). The latter option is more specific then the former. However, the former possibility is more reasonable because we already know from (8.44) that $x_1 \sqsubseteq m$. Furthermore, from (8.45) we can infer that the mereological relation can be specified as $member\_collection(x_1, m)$.

There is still one underspecified parameter, the referent y of the anaphoric null pronoun in utterance (8.12b) referring to the agent of the burning activity. Equality by Default for discourse anaphora (cf. section 7.3 on page 203, repeated below; see also section 6.3.3.3) helps us to establish the missing connection: There is nothing that contradicts that y (in

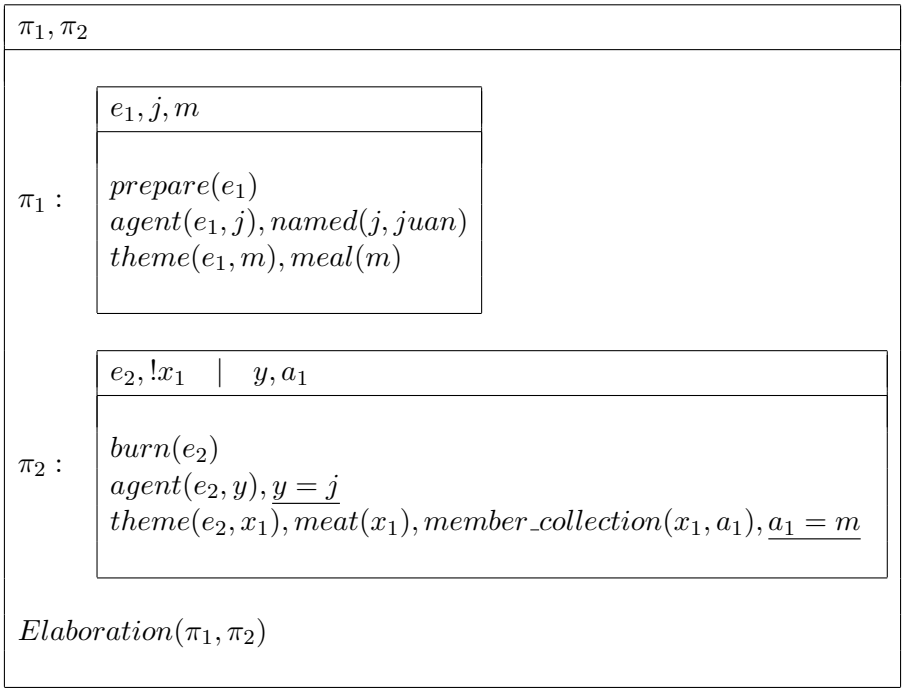
$U^{w}_{K_{\pi_1}}$) can be assumed to be equal to the referent j (in $U_{K_{\pi_2}}$) for Juan, hence $y = j$ can assumed.

Equality by Default for Discourse Anaphora

$$a \in U^{w}_{K_{\pi_j}} \wedge \alpha \in U_{K_{\pi_i}} > a = \alpha$$

The assumptions made so far, $y = j$ and $a_1 = m$, can be justified in the following way. As argued in earlier chapters, interpretation of a discourse can be seen as the process of finding a minimal model for the discourse. If we take as minimality the assumption of minimizing the number of different discourse referents (including labels because they are speech act discourse referents), unifying them whenever possible, then minimality is achieved by Equality by Default. A minimal model for the underspecified representation (8.43) of (8.12ab) is specified by (8.50). There is no other model for the discourse that is more minimal than the proposed one.

(8.50)



So far, utterance (8.12b) is connected to (8.12a). Utterance (8.12c) is interpreted in a similar way. The dislocated definite DP "las verduras" ("the vegetables") introduces a referent x_2 in π_3, connected via a bridging relation B_2 to some antecedent a_2. A successful discourse integration of the dislocation needs a topic constituent π'' with $\Downarrow^{*}$ (π'', π_3) and $a_2 \in C_{K_{\pi''}}$.

According to the SDRT definition of available attachment sites (cf. section 5.3.2.4), the preferred attachment site is the last utterance, e.g. for π_3, it would be π_2. Since both sentences (8.12b) and (8.12c) exhibit a clitic Left Dislocation, there is a partial isomorphism between the structures of the two sentences. Hence, R can be specified as Parallel or Contrast. Since these are coordinating relations and $Elaboration(\pi_1, \pi_2)$ holds, there must be, according to the principle CDP (Continuing Discourse Patterns, cf. page 121), an Elaboration relation between π_1 and π_3, as well. CDP further states that

this configuration results in $Elaboration(\pi_1, \pi')$ with a common topic node $\pi'(= \pi'')$ for π_2 and π_3 with $\Downarrow(\pi', \pi_2)$ and $\Downarrow(\pi', \pi_3)$. The resulting structure is depicted in Fig. 8.5.

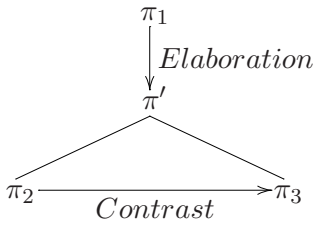


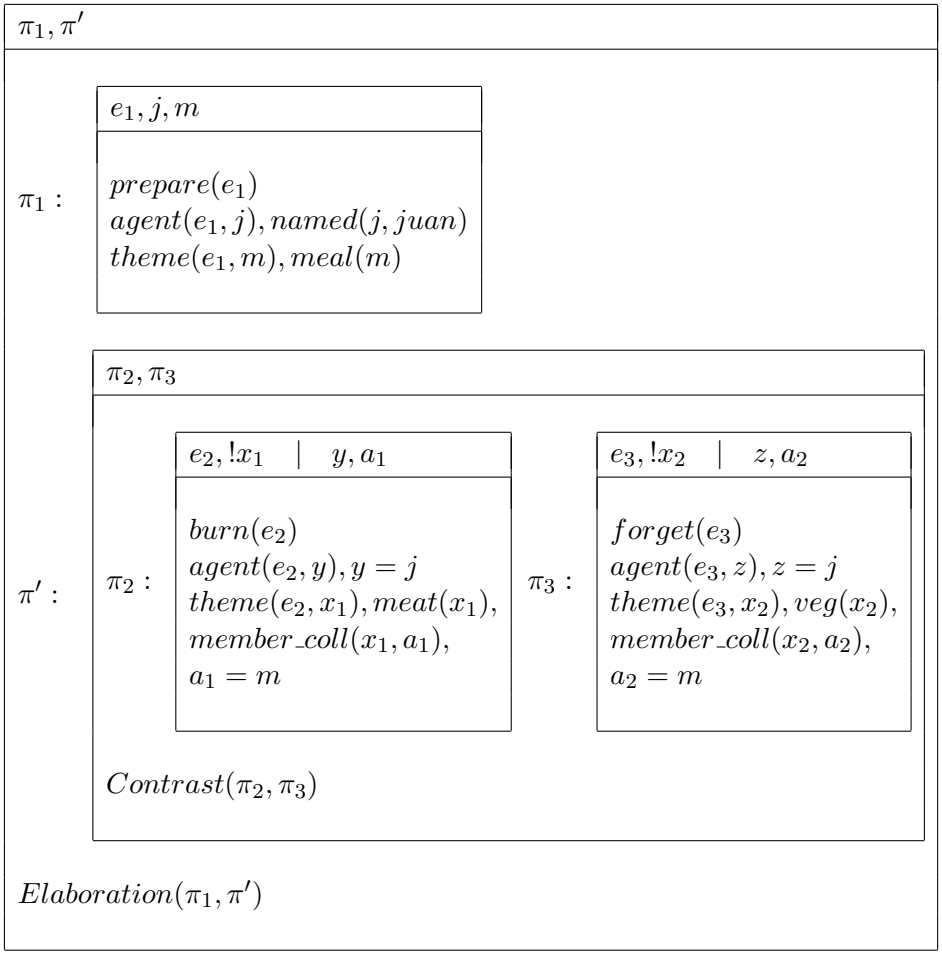
Figure 8.5: SDRT graph for (8.12).

At this point, it becomes clear why we need the transitive closure of $\Downarrow$ in the definition of the topic constraint on CLLD: from the fact that $Elaboration(\pi_1, \pi')$ entails $\Downarrow(\pi_1, \pi')$, together with $\Downarrow(\pi', \pi_3)$, can be inferred that $\Downarrow^*(\pi_1, \pi_3)$ holds. Hence, the referents in π_1 are available as anchors for the bridging anaphor in π_3. Consequently, the bridging anchor a_2 for x_2 is identified as m in π_1, just like a_1 for x_1 above. The bridging relation B_2 can be resolved to $member\_collection(x_2, m)$.[13]

As above suggested, a contrastive effect arises if the anchor provides an alternative set in relation to which the dislocated referents are contrasted with each other. More precisely, if x is a dislocated referent, $P(x)$ the condition introduced by the dislocated phrase, a the bridging anchor and $Alt(a)$ an alternative set, then the following holds: $x \in Alt(a) \wedge P(x) \wedge \forall x'[x' \in Alt(a) \wedge P(x') \rightarrow x' = x]$. Since a meal can be understood as a collection of food items, $Alt(m)$ would be something like $\{x_1, x_2, x_3, ...\} = \{x | member\_collection(x, m)\}$ with $meat(x_1), vegetables(x_2), rice(x_3)$ etc. Thus it is true only for the meat that Juan burned it, only for the vegetables that he forgot it, etc. With this knowledge, a CONTRAST relation can be established. For illustration, the final SDRS for (8.12) is given in (8.51).

13 Note that now the resolution of the bridging anchors a_1 and a_2 as e_1 with $B_1(a_1, x_1) = theme(e_1, x_1)$ and $B_2(a_2, x_2) = theme(e_1, x_2)$ is no more plausible because something that is meat cannot be a vegetable at the same time. Hence, after processing the third utterance, the only consistent resolution of both bridging anaphora involves member/collection relations.

(8.51)



8.3.3.2 Using Frame Information for Building Bridges

For the interpretation of example (8.12), some world knowledge was necessary: (i) a meal is a collection of food items that meat can be a member of, viz. (8.44) and (8.45), and (ii) an activity of burning a dish is a subtype of preparing a meal, viz. (8.46). Let us examine if this information can be obtained from FrameNet.

According to the approach presented in chapter 7, the meaning representation of utterances (8.12ab) can be enriched using FrameNet data in the following way. Let us assume that the verb "preparar" ("to prepare") evokes the frame *Cooking_creation*[14] with two core frame elements *Cook* and *Produced_food*. In the second utterance, the frame *Apply_heat* is evoked by "quemar" ("to burn"), with five core FEs *Food*, *Container*, *Cook*, *Heating_instrument*, and *Temperature_setting*. By Frame Evocation (cf. section 7.2, repeated here), we get (8.52a) and (8.52b) as underspecified representations for (8.12a) and (8.12b), respectively.

Frame Evocation

If $V \in LU(\Phi)$ and $e \in U_K$ and $V(e) \in C_K$ then

$\forall\psi[\psi \in FE(\Phi) \rightarrow \exists x_i[x_i \in U_K^w \wedge \psi(e, x_i) \in C_K]]$ and

14 Spanish FrameNet (Subirats Rüggeberg, 2005) in its present state does not include lexical items "preparar" and "quemar", nor the frame *Cooking_creation*, which, however, exists in English FrameNet.

$e : \Phi \in C_K$.

(8.52) a. π_1:

e_1, j, m
$e_1 : Cooking\_creation$ $cook(e_1, j), named(j, juan)$ $produced\_food(e_1, m), meal(m)$

b. π_2:

$e_2, !x_1 \quad \mid \quad y_1, y_2, y_3, y_4, a_1, B_1$
$e_2 : Apply\_heat$ $food(e_2, x_1), meat(x_1), B_1(a_1, x_1), B_1 =?, a_1 =?$ $container(e_2, y_1), cook(e_2, y_2),$ $heat\_instr(e_2, y_3), temp(e_2, y_4)$ $y_1 =?, y_2 =?, y_3 =?, y_4 =?$

In FrameNet, there is a *uses* relation between *Cooking_Creation* and *Apply_heat*, involving the following set of inheritance relations between FEs:

(8.53) $Uses(Cooking\_Creation, Apply\_heat) =$
$$\left\{ \begin{array}{l} \langle cook, cook\rangle, \langle produced\_food, food\rangle, \langle container, container\rangle, \\ \langle degree, degree\rangle, \langle heating\_instrument, heating\_instrument\rangle, \\ \langle manner, manner\rangle, \langle means, means\rangle, \langle purpose, purpose\rangle, \langle time, time\rangle \end{array} \right\}$$

This knowledge, according to the definition of frame element inheritance (cf. section 7.1.2.3), permits to establish $inherits(produced\_food, food)$.

It was stated above that basic world knowledge tells us that a meal typically consists of various dishes, e.g. soup, rice, meat, vegetables, dessert etc. In FrameNet, the noun "meat" evokes the frame *Food* with only one core FE *Food*, while, unfortunately, the noun "meal" is not yet a lexical unit in FrameNet. It might be assumed, for now, that "meal" evokes a frame, say, **Meal* with various *Food* subframes. Then, a *member/collection* relation between the corresponding referents x_1 and m could be assumed by default. In general, subframe relations might be used to establish mereological relations between bound frame elements. This could be captured by a default rule like the following.

Inferring Mereological Relations from Subframes If e_1: Φ_1 and e_2: Φ_2 and $\phi \in FE(\Phi_1)$ and $\psi \in FE(\Phi_2)$ then

$$\phi(e_1, x) \wedge \psi(e_2, y) \wedge subframe(\Phi_1, \Phi_2) \wedge inherits\_monotonic(\phi, \psi)$$
$$> member\_collection(x, y).$$

In words, if e_1 and e_2 instantiate the frames Φ_1 and Φ_2, respectively, and x and y are referents of the FEs ϕ and ψ in these frames, then from the knowledge that Φ_1 is a subframe of Φ_2, together with sort inheritance between ϕ and ψ, can defeasibly be inferred that x is a member of the collection y.

Now the question is how these bits of world knowledge contribute to the interpretation of our example. We have seen in the last section that information about subtypes can indicate an ELABORATION relation (8.48). A $subtype_D$ relation, in turn, can be inferred by virtue of axiom (8.47), repeated here:

Subtype $(\theta_i(x, e_\alpha) \wedge \theta_i(y, e_\beta) \wedge y \sqsubseteq x \wedge e_\beta \sqsubseteq e_\alpha) \rightarrow subtype_D(\beta, \alpha)$

In order to capture FrameNet information, this axiom has to be adapted. I suggest to state it in the following way:

Inferring Subtypes from Frames

If $\pi_1 : \boxed{e_1}$ and $\pi_2 : \boxed{e_2}$ and e_1: Φ_1 and e_2: Φ_2 and $\phi \in FE(\Phi_1)$ and $\psi \in FE(\Phi_2)$ then

$$\phi(e_1, x) \wedge \psi(e_2, y) \wedge inherits^*(\phi, \psi) \wedge e_1 \sqsubseteq e_2 \rightarrow subtype_D(\pi_1, \pi_2).$$

The first two conjuncts in this axiom state that ϕ and ψ are FEs of frames evoked in π_1 and π_2, respectively. The conjunct $inherits^*(\phi, \psi)$ is needed because in FrameNet, names for thematic roles, viz. frame elements, are always specific to a frame. The third conjunct of the original subtype definition is subsumed by $inherits^*(\phi, \psi)$ because both full inheritance and monotonic inheritance between FEs imply that the corresponding role fillers x and y stand in a subtype relation. The last conjunct, as before, requires the two eventualities to stand in a subtype relation.

Let us examine for example (8.12) if the FrameNet information mentioned so far is sufficient for fulfilling the preconditions of this axiom. From the *uses* relation, we get $produced\_food(e_1, m)$ and $food(e_2, x_1)$ and thus $inherits(produced\_food, food)$. However, this knowledge is not very useful for inferring $subtype_D(\pi_2, \pi_1)$ because what is needed is an inheritance relation in the opposite direction. Moreover, FrameNet does not tell us anything about the fact that the activity of burning a dish is actually a subtype of a cooking activity. A possible way out might be a refinement of the relevant frame information. For instance, one could assume a more abstract frame **Cooking_scenario*, on which both *Cooking_creation* and *Apply_heat* provide a particular perspective. In fact, the informal frame description of *Apply_heat* suggests that "this frame differs from *Cooking_creation* in focussing on the process of handling the ingredients, rather than the edible entity that results from the process" (cf. frame report on *Apply_heat*).

In sum, more research on FrameNet is needed in order to get conclusive results. For now, a particular default rule like (8.46) has to be stipulated, by means of which $subtype_D(\pi_2, \pi_1)$ can finally be derived and, as a consequence, an ELABORATION relation between π_1 and π_2 can be assumed. The remaining steps in the interpretation process are the same as those elaborated in the last section, except that the thematic roles *agent* and *theme* are now replaced by *cook* and *(produced) food*, respectively.

8.3.4 Resolving Frame-related Bridging Anaphora

Finally, let us turn to cases of clitic Left Dislocation involving eventualities as bridging anchors. In section 8.2.1, we have seen that the involved bridging relation can be thematic or conceptual. Two more examples are the following.

(8.54) a. Hemos alcanzado una de las muchas ciudades perdidas de Siria.

'We have reached one of the many lost cities in Syria.'

b. Apamea, a diferencia de Palmira, no está en medio de un desierto, sino dominando el valle del río Oronte y la llanura del Ghab.

'Apamea, unlike Palmira it is not in the midst of the desert, but dominating the valley of the river Oronte and the Ghab plain.'

c. El coche **lo** dejamos a la entrada de un pequeño bar de carretera y vamos caminando hasta la Puerta Sur, a los pies del Cardo.

The car / we left it at the entrance of a small road pub and are walking towards the South Door, in the base of the Cardo.'

(El País, 14-07-2007[15])

In this example, the dislocated DP "el coche" ("the car") in utterance (8.54c) triggers a bridging relation. A suitable bridging anchor is provided by the arriving event in (8.54a) denoted by the occurrence of the verb "alcanzar" ("to reach"), such that the bridging relation can be specified as a thematic relation: the car is the vehicle by means of which the arrival takes place.

(8.55) a. También había volado sobre el desierto almeriense en busca del lugar adecuado para el aterrizaje [...].

'He had also flown over the Almeria deserts looking for the proper place for landing [...]'

b. El avión **lo** consiguió Lataquia en África: un viejo Aviocar C-212 destinado al transporte de pasajeros entre Malabo y Bata, procedente de la ayuda española a Guinea Ecuatorial, construido en 1978 y que todavía volaba.

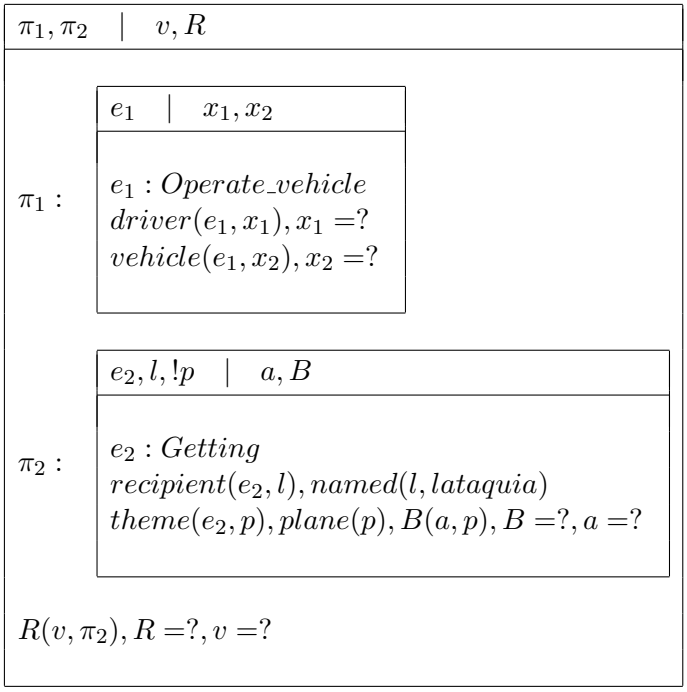
'The plane / Lataquia got it in Africa: an old Aviocar C-212 intended for the transport of passengers between Malabo and Bata, originally part of the Spanish support to Equatorial Guinea, built in 1978 and still flying.'

(Arturo Pérez-Reverte: La Reina del Sur. Alfaguara, Madrid, 2002, pp. 395–396.)

Let us examine how the bridging relation in (8.55) is established according to the account sketched in chapter 7. The verb "volar" ("to fly") in the first utterance (8.55a) evokes the frame *Operate_vehicle*, with two core frame elements: *Driver* and *Vehicle*. The FE *Driver* is filled by the implicit subject of the sentence (in Spanish FrameNet, Subirats Rüggeberg, 2005: existential implicit FE). The FE *Vehicle* is not instantiated at all (Definite Null Instantiation, DNI). Similarly, in utterance (8.55b), the frame *Getting* is evoked by the verb "conseguir" ("to get"), also with two core FEs: *Recipient*, filled by "Lataquia", and *Theme*, filled by the dislocated DP "el avión" ("the plane"). By Frame Evocation (cf. section 7.2), we get (8.56) as an underspecified representation for (8.55). For the sake of a clear exposition, only the information that is relevant for resolving the bridging anaphor is spelled out.

15 http://elviajero.elpais.com /articulo/viajes/Mil/columnas/linea/recta/elpcanviaasi/20070714elpviavje_5/Tes

(8.56)



Here, we have the underspecified variables x_1, x_2, B, a, R, and v. For resolving them, the underspecified semantics of the linguistic input shown in (8.56) must be enriched by information about the discourse structure and world knowledge.

According to the topic constraint on CLLD, the bridging anchor must be in a topic constituent π' superordinated to π_2. Since the preferred discourse constituent for attaching π_2 is the last utterance, the discourse referents therein are available for anaphora (cf. the definition of available antecedents in section 5.3.2.4 on page 158f). Thus, in this short example, $v = \pi' = \pi_1$ can be assumed by default, and as a consequence the bridging anchor is to be found in $C_{K_{\pi_1}}$.

As for the discourse relation R, the assumption of $\Downarrow^* (\pi_1, \pi_2)$ together with the fact that π_2 contains a definite description gives reason to suspect that a BACKGROUND relation holds between π_1 and π_2.

In π_1, three (regular and weak) discourse referents are introduced: e_1, x_1, and x_2. Thus, we have the possibilities for resolving $B(a, p)$ shown in Table 8.5.

a	B	$B(a,p)$	CONSISTENT	MINIMAL
$a = x_1$	*mereological*	$p \sqsubseteq^+ x_1$	*	
$a = x_2$	*mereological*	$p \sqsubseteq^+ x_2$	√	*
$a = e_1$	*mereological*	$p \sqsubseteq^+ e_1$	*	
	frame-related	$driver(e_1, p) \quad x_1 = p$	*	
	frame-related	$vehicle(e_1, p) \quad x_2 = p$	√	√

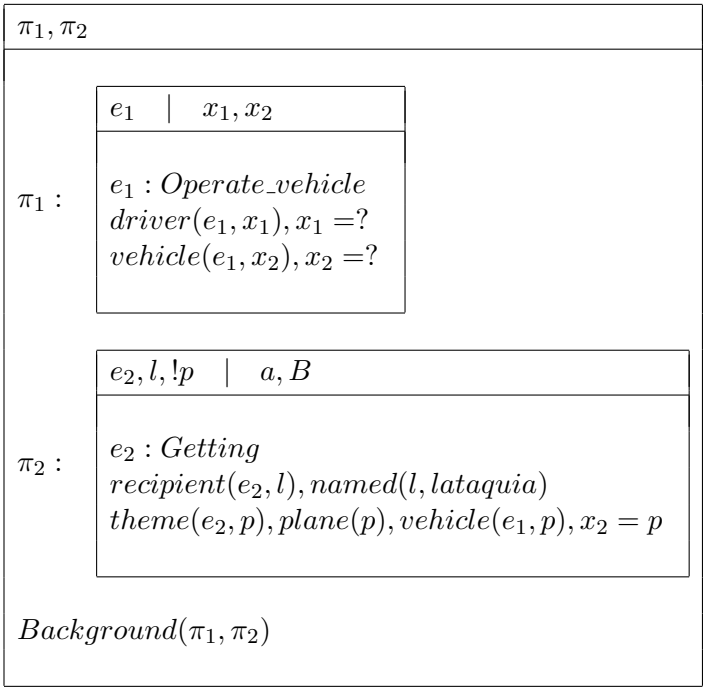
Table 8.5: Potential bridging relations in (8.56)

Let us first consider the mereological bridging relations. A subtype or part-of relation between the plane p and the driver x_1 in the activity instantiating *Operating_vehicle* is ruled out due to the sort incompatibility of the corresponding frame elements, *physical_entity*

and *sentient*, which cannot stand in a subtype relation. An (improper) part-of relationship between the vehicle x_2 and the plane p is not ruled out by a sort conflict. However, this interpretation does not resolve the underspecification of x_2 and is consequently not minimal. The third possibility, a mereological relation between the plane p and the eventuality e_1 leads to a sort conflict between *physical_entity* and *eventuality*.

As for the frame-related bridging relations, if $a = e_1$ is assumed then two options remain, one for each core FE of the *Operating_vehicle* instantiated by e_1. If $driver(e_1, p)$ is assumed to hold, p must be unified with x_1. This possibility is, like above, ruled out by a sort conflict. If we assume $vehicle(e_1, p)$, then we have no sort conflict, and a plane is a very plausible candidate for being a vehicle in a flying activity. Moreover, x_2 is specified as p. Thus, in this interpretation, the number of discourse referents and conditions is minimized and the number of resolved underspecifications is maximized. As a consequence, this discourse update is the preferred one, yielding a minimal model for the discourse. The enriched SDRS is shown in (8.57).

(8.57)



This was an example involving a BACKGROUND relation. But also other discourse relations are possible. Consider (8.58).

(8.58) a. Un híbrido nunca tiene problemas para repostar o recargar energía,

'A hybrid never has problems in refueling or recharging energy,'

b. pues la gasolina **la** venden en todas partes.

'because the petrol / they sell it everywhere.'
(http://ecoalternativa.blogspot.com/2009/07/mitos-y-realidades-sobre-los-coches.html)

Here, we have an EXPLANATION relation, indicated by the discourse marker "pues" ("for/since"). The verb "repostar" ("to refuel") evokes the frame *Filling*. The corresponding SDRS for

(8.58a) contains the following information (among other additional information, which is ignored here):

(8.59) π_1:

$e_1, h \quad \mid \quad x_1, x_2$
$e_1 : Filling$ $agent(e_1, x_1), x_1 =?$ $goal(e_1, h), hybrid(h)$ $theme(e_1, x_2), x_2 =?$

The FEs *Agent* and *Theme* are left unspecified (DNI). In (b), the frame *Commerce_sell* is evoked by "vender" ("to sell"), leading to an underspecified SDRS as follows:

(8.60) π_2:

$e_2, !g \quad \mid \quad a, B, y_1, y_2$
$e_2 : Commerce\_sell$ $buyer(e_2, y_1), y_1 =?$ $goods(e_2, g), petrol(g), B(a, g), B =?, a =?$ $seller(e_2, y_2), y_2 =?$

Here, the FEs *Buyer* and *Seller* are not instantiated. The bridging relation $B(a, g)$ triggered by the left dislocated DP "la gasolina" ("the petrol") can be specified as $theme(e_1, g)$ with $x_2 = g$, in a similar manner as in the last example.

Thus, it seems that whenever a frame contains a Definite Null Instantiation (DNI), it is likely to be filled by a subsequent bridging inference. It remains to be seen whether this hypothesis is still feasible if tested on a large corpus of texts annotated with frame information.

8.4 Conclusion

This chapter has indicated that the approach to bridging inferences taken in this thesis can be applied to language-specific phenomena. The resolution of bridging anaphora requires the interplay of various information sources: linguistic knowledge including syntactic and information structural properties of sentences, lexical semantic knowledge, and world knowledge. A very important ingredient is the discourse context. The phenomenon looked at in this chapter is, on the one hand, restricted by the surrounding discourse configuration and, on the other hand, imposes certain constraints on the latter.

The approach presented here tries to combine all these aspects in order to properly resolve bridging anaphora triggered by clitic Left Dislocations. A standard semantics is assigned to left dislocated constituents. The felicity of these constructions in a given discourse configuration is constrained by a topic constraint, which presupposes a topic constituent containing the bridging anchor. It was shown in detail how bridging anaphora triggered by CLLD are resolved, both for mereological and for frame-related bridging relations. Mereological bridging relations can be accounted for in an approach very close to standard SDRT. Information encoded in FrameNet can give important clues for resolving

these relations, albeit FrameNet data in its present state is not always sufficient for filling in the underspecified material. Further research in light of more exhaustive axiomatizations of frame knowledge may reveal more insights about the interaction of discourse structure and world knowledge in terms of frames. The proposed account works well for thematic and conceptual bridging relations, where possible bridging relations can be considerably restricted by frame information.

Summary

Subject of this thesis is the role bridging inferences play in discourse interpretation. Bridging inferences are a particular kind of pragmatic inferences, which must be drawn by the recipient of a text or a discourse in order to make sense of the linguistic input. Pragmatic inferences are in general characterized by two essential properties: they are defeasible and context-dependent.

As for the first property, it can happen that pragmatic inferences must be taken back as a discourse proceeds. The context is continuously changing and information showing up later can cancel conclusions already made. Formal methods to handle this type of reasoning are provided by nonmonotonic logics, such as Default Logic, Circumscription, Commonsense Entailment, and Abductive Reasoning. These approaches, albeit differing in formal details of the representation of defeasible inferences, have a comparable expressive power.

As for the second property, context-dependence, the explicitly expressed linguistic material often does not fully specify the intended meaning of utterances. In order to make sense of an utterance, hearers must draw additional inferences on the basis of contextual knowledge. The context is the assumedly shared knowledge among the discourse participants. It has been observed that speakers and hearers do not always examine in detail what is mutually believed. For most conversations it is sufficient to assume that the shared perspective of all participants is very close to the perspective of a particular participant. In general, the context of an utterance includes the particular situation in which an utterance is made, the discourse prior to the utterance, and general encyclopedic knowledge of the world. As a consequence, a formal theory of discourse interpretation cannot avoid including a representation of contextual knowledge.

The main characteristics of a text or a discourse are cohesion and coherence. On the one hand, cohesion emerges from anaphoric relationships between entities mentioned in the text. Although a considerable variety of anaphoric relations can be resolved by means of purely linguistic information encoded in a text, in many cases additional information is necessary in order to resolve anaphoric links. Two types of discourse anaphora can be distinguished. While a direct anaphora bears coreference of expressions, an indirect or bridging anaphor stands in a particular relationship, distinct from coreference, to an entity mentioned before, the bridging anchor. The hearer has to infer this implicit relation by means of pragmatic inferences in order to establish the anaphoric link and to make sense of the text. Anaphoric relations are established by default, but can be changed if new information contradicts a resolution already made.

On the other hand, coherence is established by virtue of discourse relations, which mark a relationship between chunks of texts, and not between single entities. Discourse relations can be expressed by cohesive means such as discourse connectors, but mostly they are not explicitly marked. In these cases, a discourse relation is established by default, albeit it may be retracted as a discourse proceeds. Thus, both establishing coherence and cohesion of a discourse involve pragmatic inferences. The two phenomena often occur intertwined:

establishing coherence depends on cohesion and vice versa.

Formal, cognitive, and psychological theories agree on the assumption that discourse interpretation involves the construction of a discourse model, which contains representations of objects and relations referred to in the discourse, but not the linguistic structure of the text itself. Such a discourse model is characterized by the following features: First, it is constructed by rich inferential means on the basis of underspecified linguistic material together with contextual knowledge. Second, it can remain underspecified: the knowledge available to a recipient is not always sufficient for completely resolving ambiguities or specifying anaphoric references. Third, it must be able to change dynamically as a discourse evolves.

Dynamic semantic theories were developed in order to capture the incremental process of how a discourse model emerges. Discourse Representation Theory (DRT) explicitly states how representations of discourse entities are obtained from the semantic content of sentences and explains why certain antecedents are not accessible for subsequent anaphoric references. However, this theory does not take into account the fact that not all discourse referents present in a discourse model are equally suited as antecedents of anaphora. Recently introduced discourse referents are more likely to be in the center of attention of discourse participants and are thus more probable candidates for antecedents of anaphora than referents introduced earlier. The dynamically changing availability of discourse referents has been accounted for in terms of focus, familiarity, givenness, accessibility, or salience. The focus of attention has been modelled by computational theories of anaphora resolution such as Focus Theory and Centering Theory. Pragmatic theories try to rely the resolution of anaphora on a few general constraints. In any case, none of the theories of anaphora resolution mentioned so far accounts for the complex hierarchical structure of discourses.

Coherent discourses are internally structured, bearing the following characteristics: Basic structural units are utterances or discourse segments, they are connected via discourse relations, and a hierarchical structure emerges from these connections. The number of assumed discourse relations, as well as their classification and properties, varies considerably across theories of discourse structure. Nevertheless, there seems to be a core inventory of discourse relations that different theories largely agree upon. While some theories advocate tree structure representations, others argue that graphs provide a more adequate representation of discourse structure. Another controversial notion is the discourse topic, the subject dealt with by a discourse segment. There are approaches assuming topics to be of diverse ontological types, stretching from propositions over questions to entities. In any case, the discourse topic has to be conceived of at a representational level, i.e. in the discourse model. One of the theories using the concept of discourse relations to build complex graph structures for coherent discourses is the Segmented Discourse Representation Theory (SDRT). This theory extends DRT by speech act discourse referents, which label the content of individual discourse segments, and discourse relations, which relate speech act discourse referents. In this thesis, SDRT has been taken as the basic model of discourse interpretation.

In sum, the first part of this thesis reviews the numerous existing approaches, which have been elaborated in a wide range of research contexts, to pragmatic inferences and defeasible reasoning in general, to the Common Ground and intentions of discourse participants, to cohesion and anaphora resolution, to coherence and discourse structure, and to discourse interpretation.

The second part of the thesis takes a closer look at bridging inferences. A new classification of bridging anaphora is made, based on existing corpus-based and psycholinguistic studies. Many theories argue that the relation between bridging anchor and anaphor is part of the lexical knowledge associated with the bridging anchor, especially if the anchor is expressed by a verb. However, recent psycholinguistic investigations have revealed that the bridging relation is mediated by the context in which the verb occurs and thus not likely to be stored in the lexicon. As a consequence, previous classifications of bridging anaphora often have difficulties in clearly determining the type of a given bridging anaphor. In fact, pragmatic inferences can be necessary in all types of bridging anaphora, since resolving them often requires non-linguistic knowledge. This thesis opts for a two-fold distinction of bridging anaphora: mereological and frame-related bridging relations.

Mereological bridging is characterized by some kind of part-of relationship between anaphor and anchor. The anchor is an already established discourse referent and its semantic type is either an entity or a set of entities.

Frame-related bridging bears on thematic and conceptual bridging relations. Here, the bridging anchor is an eventuality or a frame present in the discourse model, and the bridging relation is a thematic role which the entity denoted by the anaphor plays in the eventuality. This class subsumes also bridging anaphora involving more complex inferencing processes involving goals and intentions of communication participants. Although in these cases the bridging relation may be a conceptual relation of any type, it is often a causal, temporal, or spatial relation. All of these relations have to do with eventualities, even if they do not play as central a role as thematic relations.

Bridging is a challenge for accounts of anaphora resolution. While pragmatic accounts can explain the reasoning behind bridging in terms of intentions of discourse participants, they lack a precise formulation of these inferences. Computational and formal accounts can formally express these inferences with the help of nonmonotonic reasoning: minimal model generation can account for the preferences in anaphora resolution and automated anaphora resolution can be constrained by suitable default rules. The approach to bridging in SDRT provides the most extensive basis for drawing bridging inferences in an interplay of linguistic and contextual knowledge. In particular, this theory explains how bridging anaphora get anchored in the existing discourse structure and provides a basis for integrating world knowledge, albeit without spelling out in detail how the representations of world knowledge in the discourse model needed for drawing bridging inferences emerge in a systematic way.

This thesis develops an account of bridging inferences which is particularly aimed at covering both mereological and frame-related bridging anaphora. The powerful formal mechanisms of SDRT are integrated with encyclopedic knowledge provided by FrameNet, a cognitive network of prototypical scenarios or frames. FrameNet covers not only thematic information associated with lexical expressions, but also general cognitive concepts such as causal, temporal, and spatial relationships between entities. The present approach spells out how world knowledge, represented in frames, contributes to discourse interpretation, both for establishing discourse relations and for resolving bridging anaphora. Although some shortcomings in integrating the two lines of research still have to be resolved, the integration of FrameNet and SDRT works quite straightforwardly, assuming a neo-Davidson representation of eventualities and distinguishing regular and weak discourse referents.

In a nutshell, the following mechanism is proposed: Each eventuality talked about in an discourse evokes a particular frame and introduces some referents, which stand for frame elements, into the discourse model. Frame elements are entities that can participate in a

frame; they correspond roughly to thematic roles in the eventuality associated with that frame. The introduced referents can be weak and remain underspecified if the corresponding frame element is not explicitly expressed. Crucially, weak discourse referents provide plausible candidates for antecedents of subsequent bridging anaphora. In this way, the search space for suitable bridging antecedents can be restricted to available regular and weak discourse referents. Conditions on discourse referents which are already present in the discourse model are considered as preferred specifications of bridging relations. Additionally, it is indicated that the resolution of bridging anaphora can be explained in terms of a small number of general cognitive principles, which play a role in discourse interpretation. These include (i) a preference for minimal interpretations in which anaphoric conditions are resolved to already introduced weak or regular discourse referents, (ii) a preference for consistent and plausible interpretations, (iii) the determination of available attachment points for utterances and antecedents for anaphora by the discourse configuration, and (iv) maximal discourse coherence.

A particular construction, Clitic Left Dislocation (CLLD) in Spanish, has been analyzed in detail in the present framework. CLLD is a device used to link an utterance to the preceding discourse in a particular way, often involving bridging anaphora. Apart from discussing grammatical aspects and discourse functions of left dislocations, the focus is directed to the integration of bridging inferences triggered by this construction into the discourse model. Corpus studies have shown that, on the one hand, the felicity of this phenomenon is restricted by the surrounding discourse configuration and, on the other hand, it imposes certain constraints on the latter.

With regard to discourse topics, two fundamental discourse functions of CLLD can be distinguished: (i) CLLD can introduce a new discourse topic involving new discourse referents, and (ii) CLLD can progressively change or continue a given discourse topic, giving rise to a bridging inference. The first case is mainly used at the beginning of a text. In the second case, CLLD occurs in a subordinating environment, and the anchor of the bridging anaphor must be contained in or evoked by the superordinated constituent. A contrastive reading emerges if the anchor provides a suitable set of alternatives. These conditions on the discourse structure can be formally expressed by a topic constraint, which states that there must be a discourse topic constituent that contains the bridging anchor and that is superordinated to the utterance containing the bridging anaphor. It is shown in detail how bridging anaphora triggered by CLLD are resolved, both for mereological and for frame-related bridging relations.

This work tries to shed light on bridging inferences from different points of view. Insights from formal, computational, cognitive, and psychological research on the subject complement each other. Some steps have been taken in order to build bridges between these areas, resulting in a theory which provides a basis for future research in various directions.

The insights about the nature of bridging inferences may help to shape the role of pragmatic inferences in natural language understanding. In particular, the filling of underspecified material by default values for implicit information must be constrained in a principled way. It might be possible to put down the assumed constraints on anaphora resolution to an even smaller number of general principles of human cognition. An interesting question is whether the presented account of bridging inferences carries over to other kinds of phenomena at the borderline of semantics and discourse pragmatics.

An implementation of the proposed approach to bridging inferences using a logic programming paradigm that supports nonmonotonic reasoning is straightforward. Tech-

niques of model generation can be used to single out preferred interpretations of bridging anaphora. The information provided by FrameNet can easily be integrated in the reasoning process in a modular way. Since this resource is steadily growing and subject to changes, it remains to be tested if the approach is feasible for large amounts of naturally occurring data. A large-scale corpus analysis on a computational basis could provide new insights about the use, typical occurrences, and fine-grained distributional facts on bridging inferences.

List of Figures

List of Tables

List of Abbreviations

CDP	Continuing Discourse Patterns
CLLD	Clitic Left Dislocation
CNI	Constructional Null Instantiation
CS	Conceptual Structure
D-LTAG	Discourse - Lexicalized Tree Adjoining Grammar
DNI	Definite Null Instantiation
DOAP	Don't Overlook Anaphoric Possibilities
DPL	Dynamic Predicate Logic
DRS	Discourse Representation Structure
DRT	Discourse Representation Theory
ERP	Event-related Potential
FE	Frame Element
GCI	Generalized Conversational Implicature
HTLD	Hanging Topic Left Dislocation
INI	Indefinite Null Instantiation
LDM	Linguistic Discourse Model
MDC	Maximize Discourse Coherence
PCI	Particularized Conversational Implicature
PFS	Parameter-fixed Structure
QUD	Question under Discussion
RFC	Right Frontier Constraint
RST	Rhetorical Structure Theory
SDRS	Segmented Discourse Representation Structure
SDRT	Segmented Discourse Representation Theory
SF	Semantic Form
TAG	Tree Adjoining Grammar

Abbreviations in Glosses

acc	accusative
CL	clitic
dat	dative
f	feminine
FUT	future tense
INF	infinitive
IPF	imperfective
loc	locative
m	masculine
NEG	negation particle
nom	nominative
obj	object case
obl	oblique
part	partitive
PAST	past tense
pl	plural
poss	possessive
PTCP	participle
REFL	reflexive
sg	singular

Grammatical Categories

AdvP	adverbial phrase
AP	adjectival phrase
CP	complementizer phrase
DP	determiner phrase
IP	inflectional phrase
NP	noun phrase
PP	prepositional phrase
PRO	pronoun
Spec	specifier
VP	verb phrase
vP	verbal phrase (functional category)
XP	any phrase (maximal projection)

A Note on Used Corpora

The linguistic data used in this thesis is either taken from a variety of both written and spoken language corpora, or consists of attested data from native speakers, or is under discussion in the literature from the authors as cited. Used corpora include:

- a collection of fairytales of the brothers Jacob and Wilhelm Grimm ("Kinder- und Hausmärchen", first published in 1812) in a parallel corpus in several languages (http://www.grimmstories.com),
- the "Potsdam Commentary Corpus" (Stede, 2004a), a corpus of German newspaper articles from the regional newspaper "Märkische Allgemeine", with annotation of rhetorical relations,
- newspaper texts mainly from the Time Magazine (http://www.time.com), retrieved via "WebCorp" (http://www.webcorp.org.uk), a linguistic search engine to process large sections of the web,
- the "Corpus de referencia de la lengua española contemporanea (CRLEC)" (Marín, 1992) (http://www.lllf.uam.es/corpus/corpus_oral.html), a spoken language corpus,
- the "Corpus del Español" edited by Davies (2002) (http://www.corpusdelespanol.org),
- the "Corpus de referencia del español actual (CREA)" (Real Academia Española, 2008) (http://corpus.rae.es/creanet.html),
- the spoken language corpus "El habla de la ciudad de Madrid" compiled by Esgueva and Cantarero (1981),
- data from FrameNet (Baker et al., 1998) (http://framenet.icsi.berkeley.edu/) and
- from Spanish FrameNet (Subirats Rüggeberg, 2005) (http://gemini.uab.es/SFN).

Citation Index

Bibliography

Abbott, Barbara (1997): 'Models, truth and semantics.' *Linguistics and Philosophy*, **20**(2), 117–138.

Akkoyunlu, E. A., K. Ekanadham, and R. V. Huber (1975): 'Some constraints and tradeoffs in the design of network communications.' In *SOSP '75: Proceedings of the fifth ACM symposium on Operating systems principles*, pp. 67–74. ACM, New York.

Allott, Nicholas (2006): 'Game theory and communication.' In Benz et al. (2006), pp. 123–151.

Alonso i Alemany, Laura (2005): *Representing discourse for automatic text summarization via shallow NLP techniques*. Ph.D. thesis, Universitat de Barcelona.

Anagnostopoulou, Elena, Henk van Riemsdijk, and Frans Zwarts, eds. (1997): *Materials on Left Dislocation*. Benjamins, Amsterdam and Philadelphia.

Ariel, Mira (1990): *Accessing Noun-Phrase Antecedents*. Routledge, London.

Ariel, Mira (2001): 'Accessibility theory: an overview.' In Sanders et al. (2001), pp. 29–87.

Arregi, Karlos (2003): 'Clitic left dislocation is contrastive topicalization.' *U.Penn Working Papers in Linguistics*, **9**(1), 31–44.

Asher, Nicholas (1993): *Reference to Abstract Objects in Discourse*. Kluwer, Dordrecht.

Asher, Nicholas (2004a): 'Discourse topic.' *Theoretical Linguistics*, **30**(2-3), 163–201. doi:10.1515/thli.2004.30.2-3.163.

Asher, Nicholas (2004b): 'Troubles with topics: Comments on Kehler, Oberlander, Stede and Zeevat.' *Theoretical Linguistics*, **30**(2-3), 255–262. doi:10.1515/thli.2004.30.2-3.255.

Asher, Nicholas (2008): 'Troubles on the right frontier.' In Anton Benz and Peter Kühnlein, eds., *Constraints in Discourse*, pp. 29–52. Benjamins, Amsterdam.

Asher, Nicholas and Alex Lascarides (1998a): 'Bridging.' *Journal of Semantics*, **15**(1), 83–113.

Asher, Nicholas and Alex Lascarides (1998b): 'The semantics and pragmatics of presupposition.' *Journal of Semantics*, **15**(2), 239–299.

Asher, Nicholas and Alex Lascarides (2003): *Logics of Conversation*. Cambridge University Press, Cambridge.

Asher, Nicholas and Michael Morreau (1991): 'Commonsense entailment: A modal theory of non-monotonic reasoning.' In *IJCAI*, pp. 387–392.

Asher, Nicholas and James Pustejovsky (2005): 'Word meaning and commonsense metaphysics.' Published online at http://semanticsarchive.net.

Asher, Nicholas and Laure Vieu (2005): 'Subordinating and coordinating discourse relations.' *Lingua*, **115**, 591–610.

Averintseva-Klisch, Maria (2008): *Rechte Satzperipherie im Diskurs: Die NP-Rechtsversetzung im Deutschen*. Ph.D. thesis, University of Tübingen.

Bach, Kent (1994): 'Conversational impliciture.' *Mind and Language*, **9**, 124–162.

Baker, Collin, Charles Fillmore, and Beau Cronin (2003): 'The structure of the FrameNet database.' *International Journal of Lexicography*, **16**(3), 281–296.

Baker, Collin, Charles Fillmore, and John Lowe (1998): 'The Berkeley FrameNet project.' In Christian Boitet and Pete Whitelock, eds., *Proceedings of the 36th ACL and 17th COLING*, pp. 86–90. San Francisco, CA.

Barwise, Jon (1988): 'Three views of common knowledge.' In M. Y. Vardi, ed., *Proceedings of the 2nd Conference on Theoretical Aspects of Reasoning About Knowledge: Theoretical aspects of rationality and knowledge*, pp. 365–379. Morgan Kaufmann Publishers, San Francisco, CA.

Baumgartner, Peter and Michael Kühn (2000): 'Abducing Coreference by Model Construction.' *Journal of Language and Computation*, **1**(2), 175–190.

Beaver, David (2004): 'The optimization of discourse anaphora.' *Linguistics and Philosophy*, **27**(1), 3–56.

Benz, Anton, Gerhard Jäger, and Robert van Rooij (2006): *Game Theory and Pragmatics*. Palgrave MacMillan.

Bierwisch, Manfred (1983): 'Semantische und konzeptuelle Repräsentation lexikalischer Einheiten.' In R. Růžička and W. Motsch, eds., *Untersuchungen zur Semantik*, pp. 61–99. Akademie Verlag, Berlin.

Bierwisch, Manfred and Ewald Lang (1989): *Dimensional Adjectives : Grammatical Structure and Conceptual Interpretation*. Springer-Verlag, Berlin.

Blackburn, Patrick and Johan Bos (2003): 'Computational semantics.' *Theoria*, **18**(1), 27–45.

Blackburn, Patrick, Johan Bos, Michael Kohlhase, and Hans de Neville (1999): 'Inference and computational semantics.' In Harry Bunt and Elias Thijsse, eds., *IWCS-3*. Tilburg, The Netherlands.

Blackburn, Patrick, Johan Bos, and Kristina Striegnitz (2006): *Learn Prolog Now!*, volume 7 of *Texts in Computing*. College Publications.

Blackwell, Sarah E. (2003): *Implicatures in Discourse: The Case of Spanish NP anaphora*. Benjamins, Amsterdam.

Blutner, Reinhardt (1998): 'Lexical pragmatics.' *Journal of Semantics*, **15**, 115–162.

Blutner, Reinhardt (2000): 'Some aspects of optimality in natural language interpretation.' *Journal of Semantics*, **17**, 189–216.

Blutner, Reinhardt, A. Lessmöllmann, and Rob van der Sandt (1996): 'Conversational implicature and lexical pragmatics.' In *Proceedings of the AAAI Spring Symposium on Conversational Implicature*. Stanford.

Blutner, Reinhardt and Henk Zeevat (2004): *Optimality Theory and Pragmatics*. Palgrave/Macmillan.

Bos, Johan (2001): *Underspecification and Resolution in Discourse Semantics*. Ph.D. thesis, Universität des Saarlandes.

Bos, Johan, Paul Buitelaar, and Anne-Marie Mineur (1995): 'Bridging as coercive accommodation.' In *Proceedings to the Workshop on Computational Logic for Natural Language Processing (CLNLP)*. Edinburgh.

Bos, Johan and Malvina Nissim (2008): 'Combining Discourse Representation Theory with FrameNet.' In R. Rossini Favretti, ed., *Frames, Corpora, and Knowledge Representation*, pp. 169–183. Bononia University Press, Bologna.

Bras, Myriam (2007): 'French adverb d'abord and discourse structure.' In M. Aurnague, J. Larrazabal, and K. Korta, eds., *Language, Representation and Reasoning. Memorial Volume to Isabel Gomez Txurruka*, pp. 77–102. Universidad del País Vasco, Bilbao.

Brennan, Jonathan and Liina Pylkkänen (2008): 'Processing events: Behavioral and neuromagnetic correlates of Aspectual Coercion.' *Brain and Language*, **106**, 132–143.

Brennan, Susan, Marilyn Friedman, and Carl Pollard (1987): 'A centering approach to pronouns.' In *Proceedings of 25th conference of the Association for Computational Linguistics*, pp. 155–162. Stanford, CA.

Brewka, Gerhard, Jürgen Dix, and Kurt Konolige (1997): *Nonmonotonic Reasoning: An Overview*. CSLI Publications, Stanford.

Brewka, Gerhard and Thomas Eiter (1999): 'Preferred answer sets for extended logic programs.' *Artificial Intelligence*, **109**, 297–356.

Britton, Bruce K. and Arthur C. Graesser, eds. (1996): *Models of Understanding Text*. Erlbaum, Hillsdale, NJ.

Brunetti, Lisa (2007): 'Left dislocation in romance and contrast effects.' In *Workshop on Contrast*. ZAS Berlin.

Büring, Daniel (1997): *The Meaning of Topic and Focus — The 59th Street Bridge Accent*. Routledge, London.

Büring, Daniel (2005): *Binding Theory*. Cambridge Textbooks in Linguistics. Cambridge University Press, Cambridge.

Burkhardt, Petra (2006): 'Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials.' *Brain and Language*, **98**(2), 159–168.

Carston, Robyn (1988): 'Explicature, implicature and truth-theoretic semantics.' In Ruth Kempson, ed., *Mental representations: the interface between language and reality*, pp. 155–181. Cambridge University Press, Cambridge.

Carston, Robyn (1999): 'The semantics/pragmatics distinction: A view from Relevance Theory.' In Ken Turner, ed., *The Semantics/Pragmatics Interface from Different Points of View*, pp. 85–125. Elsevier, Amsterdam.

Carston, Robyn (2004): 'Truth-conditional content and conversational implicature.' In C. Bianchi, ed., *The Semantics/Pragmatics Distinction*. CSLI Publications, Stanford.

Chafe, Wallace L. (1976): 'Givenness, contrastiveness, definiteness, subjects, topics and points of view.' In Charles N. Li, ed., *Subject and Topic*, pp. 27–55. Academic Press.

Charniak, E. and S. Shimony (1994): 'Cost-based abduction and map explanation.' *Artificial Intelligence*, **66**, 345–374.

Chierchia, Gennaro (1995): *Dynamics of Meaning: Anaphora, Presuppositions, and the Theory of Grammar*. University of Chicago Press, Chicago.

Chierchia, Gennaro (2004): 'Scalar Implicatures, Polarity Phenomena, and the Syntax / Pragmatics Interface.' In Adriana Belletti, ed., *Structure and Beyond. The Cartography of Syntactic Structures*, pp. 39–103. Oxford University Press, Berlin and Oxford and New York.

Chomsky, Noam (1973): 'Conditions on transformations.' In S. R. Anderson and P. Kiparsky, eds., *A festschrift for Morris Halle*. Holt, Rinehart & Winston, New York.

Chomsky, Noam (1981): *Lectures on Government and Binding*. Foris, Dordrecht.

Cinque, Guglielmo (1990): *Types of A-bar-dependencies*. MIT Press, Cambridge, MA.

Clark, Herbert (1975): 'Bridging.' In *Proceedings of the 1975 workshop on Theoretical issues in natural language processing*, pp. 169–174. Association for Computational Linguistics.

Clark, Herbert (1977): 'Bridging.' In P. N. Johnson-Laird and P. Wason, eds., *Thinking: Readings in Cognitive Science*, pp. 411–420. Cambridge.

Clark, Herbert (1996): *Using Language*. Cambridge University Press, Cambridge.

Clark, Herbert and Susan Brennan (1991): 'Grounding in communication.' In L. et al. Resnick, ed., *Perspectives on socially shared cognition*, pp. 127–149. Washington, DC.

Clark, Herbert and Catherine Marshall (1981): 'Definite reference and mutual knowledge.' In Aravind Joshi, Bonnie Webber, and Ivan Sag, eds., *Elements of discourse understanding*. Cambridge University Press, Cambridge.

Clark, Herbert and Edward Schaefer (1989): 'Contributing to discourse.' *Cognitve Science*, **13**, 259–294.

Cohen, Ariel (2006): 'Anaphora resolution and minimal models.' In J. Bos and A. Koller, eds., *ICoS-5*, pp. 7–16. Buxton, England.

Cohen, Ariel (2007): 'Anaphora resolution by default.' In J. Geertzen, E. Thijsse, H. Bunt, and A. Schiffrin, eds., *IWCS-7*, pp. 53–64. Tilburg, The Netherlands.

Consten, Manfred (2004): *Anaphorisch oder deiktisch? - Zu einem integrativen Modell domänengebundener Referenz*. Niemeyer, Tübingen.

Copestake, Ann, Dan Flickinger, Ivan Sag, and Carl Pollard (2005): 'Minimal Recursion Semantics: An introduction.' *Journal of Research on Language and Computation*, **3**(2–3), 281–332.

Cornish, Francis (1999): *Anaphora, Discourse, and Understanding. Evidence from English and French*. Clarendon Press, Oxford.

Cristea, Dan, Nancy Ide, and Laurent Romary (1998): 'Veins theory: a model of global discourse cohesion and coherence.' In *Proceedings of the 17th international conference on Computational linguistics*, pp. 281–285. Association for Computational Linguistics.

Curran, James, Stephen Clark, , and Johan Bos (2007): 'Linguistically motivated large-scale NLP with C&C and Boxer.' In *Proceedings of the ACL 2007 Demonstrations Session (ACL-07 demo)*, pp. 29–32.

Dalrymple, Mary, Stuart M. Shieber, and Fernando C. N. Pereira (1991): 'Ellipsis and higher-order unification.' *Linguistics and Philosophy*, **14**(4), 399–452.

Danlos, Laurence (2001): 'Event coreference between two sentences.' In Harry Bunt, Reinhard Muskens, and Elias Thijsse, eds., *Computing Meaning Vol. 2*. Kluwer, Tilburg.

Danlos, Laurence (2006): 'Anaphoric expressions and discourse structures.' In *Proceedings des Journées Sémantique et Modélisation*. Bordeaux.

Danlos, Laurence (2007): 'D-STAG : A discourse formalism using synchronous TAG.' In K. Korta M. Aunargue and J. Lazzarabal, eds., *Language, Representation and Reasoning*. University of the Basque country Press.

Davidson, Donald (1967): 'The logical form of action sentences.' In N. Rescher, ed., *The Logic of Decision and Action*. University of Pittsburgh Press, Pittsburgh.

Davies, Mark (2002): *Corpus del español (100 millones de palabras, siglo XIII–siglo XX)*. http://www.corpusdelespanol.org.

de Beaugrande, Robert-Alain and Wolfgang Ulrich Dressler (1981): *Einführung in die Textlinguistik*. Niemeyer, Tübingen.

Doetjes, Jenny, Elisabeth Delais-Roussarie, and Petra Sleeman (2002): 'The prosody of left detached constituents in french.' In Bernard Bel and Isabelle Marlien, eds., *Speech Prosody 2002*. Aix-en-Provence, France.

Dölling, Johannes (1997): 'Semantic form and abductive fixation of parameters.' In Rob van der Sandt, Reinhardt Blutner, and Manfred Bierwisch, eds., *From Underspecification to Interpretation. Working Papers of the Institute of Logic and Linguistics*, pp. 113–138. IBM Deutschland, Heidelberg.

Dölling, Johannes (2001): 'Systematische Bedeutungsvariationen: Semantische Form und kontextuelle Interpretation.' *Linguistische Arbeitsberichte*, **78**. Institut für Linguistik, University of Leipzig.

Dölling, Johannes (2003): 'Flexibility in adverbal modification: Reinterpretation as contextual enrichment.' In E. Lang, C. Maienborn, and C. Fabricius-Hansen, eds., *Modifying Adjuncts*, pp. 511–552. Mouton de Gruyter, Berlin.

Dölling, Johannes (2005): 'Semantische Form und pragmatische Anreicherung: Situationsausdrücke in der Äusserungsinterpretation.' *Zeitschrift für Sprachwissenschaft*, **24**, 159–225.

Dowty, David, Robert Wall, and Stanley Peters (1981): *Introduction to Montague Semantics*. Reidel, Dordrecht.

Eco, Umberto (1985): 'Hörner, Hufe, Sohlen. Einige Hypothesen zu drei Abduktionstypen.' In Umberto Eco, ed., *Der Zirkel oder im Zeichen der Drei*, pp. 288–320. Munich.

Egg, Markus (2005): *Flexible Semantics for Reinterpretation Phenomena*. CSLI Publications, Stanford.

Egg, Markus and Gisela Redeker (2008): 'Underspecified discourse representation.' In Anton Benz and Peter Kühnlein, eds., *Constraints in Discourse*, pp. 117–138. Benjamins, Amsterdam.

Erkü, Feride and Jeanette Gundel (1987): 'The pragmatics of indirect anaphors.' In J. Verschueren and M. Bertuccelli-Papi, eds., *The pragmatic perspective*, pp. 533–545. Benjamins, Amsterdam.

Esgueva, Manuel and Margarita Cantarero (1981): *El habla de la ciudad de Madrid. Materiales para su estudio*. Instituto Miguel de Cervantes, Madrid.

Farkas, Donka F. and Henriëtte De Swart (2003): *The Semantics of Incorporation*. Stanford Monographs in Linguistics. CSLI Publications, Stanford.

Fauconnier, Gilles (1994): *Mental Spaces*. Cambridge University Press, New York.

Fellbaum, Christiane, ed. (1998): *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, MA.

Fillmore, Charles (1968): 'The case for the case.' In E. Bach and R. Harms, eds., *Universals in Linguistic Theory*. Rinehart and Winston, New York.

Fillmore, Charles (1976): 'Frame semantics and the nature of language.' In *Annals of the NY Academy of Sciences: Conf. on the Origin and Development of Language and Speech 280*, pp. 20–32.

Fillmore, Charles, Christopher Johnson, and Miriam Petruck (2003): 'Background to FrameNet.' *International Journal of Lexicography*, **16**(3), 235–250.

Forbes, Katherine, Eleni Miltsakaki, Rashmi Prasad, Anoop Sarkar, Aravind Joshi, and Bonnie Webber (2001): 'D-ltag system: Discourse parsing with a lexicalized tree adjoining grammar.' In *Proceedings of ESSLLI 2001: Workshop on Information Structure, Discourse Structure and Discourse Semantics*.

Fraurud, Kari (1990): 'Definiteness and the processing of NPs in natural discourse.' *Journal of Semantics*, **7**, 395–433.

Freitas, Sergio Antônio Andrade de (2005): *Interpretação automatizada de textos: processamento de anáforas*. Ph.D. thesis, Universidade Federal do Espírito Santo, Vitória, Brasil.

Frey, Werner (2005): 'Pragmatic properties of certain german and english left peripheral constructions.' *Linguistics*, **43**(1), 89–129.

Frisson, Steven (2009): 'Semantic underspecification in language processing.' *Language and Linguistics Compass*, **3**(1), 111–127.

Gaddy, Michelle, Paul van den Broek, and Yung-Chi Sung (2001): 'The influence of text cues on the allocation of attention during reading.' In Sanders et al. (2001), pp. 89–110.

Gardent, Claire and Karsten Konrad (2000): 'Interpreting definites using model generation.' *Journal of Language and Computation*, **1**(2), 193–209.

Gardent, Claire, Hélène Manuélian, and Eric Kow (2003): 'Which bridges for bridging definite descriptions?' In *Proceedings of LINC'03*. Budapest, Hungary.

Gardent, Claire and Bonnie Webber (2001): 'Towards the use of automated reasoning in discourse disambiguation.' *Journal of Logic, Language, and Information*, **10**(4), 487–509.

Garnham, Alan and Jane Oakhill (1990): 'Mental models as contexts for interpreting texts: Implications from studies of anaphora.' *Journal of Semantics*, **7**, 379–393.

Garnham, Alan and Jane Oakhill (1996): 'The mental models theory of language comprehension.' In Britton and Graesser (1996), pp. 313–339.

Garrod, Simon and A. J. Sanford (1982): 'The mental representation of discourse in a focussed memory system: implications for the interpretation of anaphoric noun phrases.' *Journal of Semantics*, **1**(1), 21–41.

Garrod, Simon and Melody Terras (2000): 'The contribution of lexical and situational knowledge to resolving discourse roles: Bonding and resolution.' *Journal of Memory and Language*, **42**, 526–544.

Gernsbacher, Morton A. (1996): 'The structure-building framework: what it is, what it might also be, and why.' In Britton and Graesser (1996), pp. 289–311.

Ginzburg, Jonathan (1996): 'Interrogatives: questions, facts and dialogue.' In Shalom Lappin, ed., *The handbook of contemporary semantic theory*, pp. 385–422. Blackwell, Oxford.

Ginzburg, Jonathan (2009): *Semantics for Conversation*. Studies in Computational Linguistics. CSLI Publications, Stanford.

Gómez Txurruka, Isabel (2003): 'El significado de la conjunción y.' *Teorema*, **XXII**(1-2), 55–84.

Grice, Paul (1975): 'Logic and conversation.' In Peter Cole and Jerry L. Morgan, eds., *Syntax and semantics*, volume 3. Academic Press, New York.

Groenendijk, Jeroen and Martin Stokhof (1984): *Studies on the semantics of questions and the pragmatics of answers*. Ph.D. thesis, Universiteit van Amsterdam.

Groenendijk, Jeroen and Martin Stokhof (1991): 'Dynamic predicate logic.' *Linguistics and Philosophy*, **14**(1), 39–100.

Grosz, Barbara, Aravind Joshi, and Scott Weinstein (1995): 'Centering: A framework for modelling the local coherence of discourse.' *Computational Linguistics*, **21**(2), 203–226.

Grosz, Barbara and Candace Sidner (1986): 'Attention, intentions, and the structure of discourse.' *Computational Linguistics*, **12**(3), 175–204.

Gundel, Jeanette, Nancy Hedberg, and Ron Zacharski (1993): 'Cognitive status and the form of referring expressions in discourse.' *Language*, **69**(2), 274–307.

Hajičová, Eva, Barbara Partee, and Petr Sgall (1998): *Topic-Focus Articulation, Tripartite Structures, and Semantic Content*. Kluwer Academic Publishers, Amsterdam, Netherlands.

Hankamer, Jorge and Ivan Sag (1976): 'Deep and surface anaphora.' *Linguistic Inquiry*, **7**, 391–428.

Harman, Gilbert (1977): 'Review of Linguistic Behaviour by Jonathan Bennett.' *Language*, **53**, 417–424.

Haviland, Susan E. and Herbert Clark (1974): 'What's new? acquiring new information as a process in comprehension.' *Journal of Verbal Learning and Verbal Behavior*, **13**(5), 512–521.

Hawkins, John (1978): *Definiteness and Indefiniteness. A Study in Reference and Grammaticality Prediction*. Croom Helm, London.

Heim, Irene (1982): *On the Semantics of Definite and Indefinite Noun Phrases*. Ph.D. thesis, University of Massachusetts at Amherst.

Heim, Irene and Angelika Kratzer (1998): *Semantics in Generative Grammar*. Blackwell, Oxford.

Heinemann, Margot and Wolfgang Heinemann (2002): *Grundlagen zur Textlinguistik. Interaktion - Text - Diskurs*. Niemeyer, Tübingen.

Hendriks, Petra (2004): 'Coherence relations, ellipsis, and contrastive topics.' *Journal of Semantics*, **21**, 132–154.

Hidalgo Downing, Raquel (2001): *La Dislocación a la izquierda en el español hablado. Funciones discursivas: Estudio de corpus*. Ph.D. thesis, Universidad Complutense de Madrid.

Hirschberg, Julia (1991): *A Theory of Scalar Implicature*. Ph.D. thesis, University of Pennsylvania.

Hirschberg, Julia and Christine H. Nakatani (1996): 'A prosodic analysis of discourse segments in direction-giving monologues.' In *ACL*, pp. 286–293.

Hobbs, Jerry (1979): 'Coherence and coreference.' *Cognitive Science*, **3**(1), 67–90.

Hobbs, Jerry (1985a): 'On the coherence and structure of discourse.' Technical Report CSLI-85-37, CSLI, Stanford.

Hobbs, Jerry (1985b): 'Ontological promiscuity.' In *Proceedings of the 23rd ACL meeting*, pp. 61–69.

Hobbs, Jerry (1990): *Literature and Cognition*. CSLI Publications, Stanford.

Hobbs, Jerry (1996): 'On the relation between the informational and intentional perspectives on discourse.' In E. Hovy and D. Scott, eds., *Computational and Conversational Discourse: Burning Issues–An Interdisciplinary Account*, pp. 139–157. Springer, Berlin.

Hobbs, Jerry (2001): 'Syntax and metonymy.' In P. Bouillon and F. Busa, eds., *The Language of Word Meaning*, pp. 290–311. Cambridge University Press, Cambridge.

Hobbs, Jerry, Mark Stickel, Douglas Appelt, and Paul Martin (1993): 'Interpretation as abduction.' *Artificial Intelligence*, **63**, 69–142.

Horn, Larry (1984): 'Towards an new taxonomy for pragmatic inference: Q-based and R-based implicatures.' In D. Schiffrin, ed., *Meaning, form and use in context: Linguistic Applications (GURT '84)*, pp. 11–42. Georgetown University Press, Washington.

Horton, William and Boaz Keysar (1996): 'When do speakers take into account common ground?' *Cognition*, **59**, 91–117.

Huang, Yan (1994): *The Syntax and Pragmatics of Anaphora: A Study with Special Reference to Chinese*. Cambridge University Press, Cambridge.

Hume, David (1748): *An Enquiry Concerning Human Understanding*. London.

Iris, M. A., B. E. Litowitz, and M. Evens (1988): 'Problems of the part-whole relation.' In M. W. Evens, ed., *Relational Models of the Lexicon: Representing Knowledge in Semantic Networks*, pp. 261–288. Cambridge University Press, Cambridge.

Jackendoff, Ray (1983): *Semantics and Cognition*. MIT Press, Cambridge, MA.

Jackendoff, Ray (1987): 'The status of thematic relations in linguistic theory.' *Linguistic Inquiry*, **18**, 369–412.

Jacobs, Joachim (2001): 'The dimensions of topic-comment.' *Linguistics*, **39**(4), 641–681.

Jäger, Gerhard (2007): 'Game dynamics connects semantics and pragmatics.' In Ahti-Veikko Pietarinen, ed., *Game Theory and Linguistic Meaning*, pp. 89–102. Elsevier, Amsterdam.

Jasinskaja, Katja (2006): 'Non-canonical applications of topic continuity: Restatement and elaboration.' In Sidner, Harpur, Benz, and Kühnlein, eds., *Proceedings of the 2nd Workshop on Constraints in Discourse*, pp. 107–115. Maynooth, Ireland.

Johnson-Laird, P. N. (1983): *Mental Models*. Cambridge University Press, Cambridge.

Johnson-Laird, P. N. (2005): 'Mental models in thought.' In *The Cambridge Handbook of Thinking and Reasoning*, pp. 179–212. Cambridge University Press, Cambridge.

Joshi, Aravind (1985): 'Tree-adjoining grammars.' In D. Dowty, L. Karttunen, and A. Zwicky, eds., *Natural language parsing*, pp. 206–250. Cambridge University Press, Cambridge.

Joshi, Aravind, Laura Kallmeyer, and Maribel Romero (2007): 'Flexible composition in LTAG: Quantifier scope and inverse linking.' In Reinhard Muskens and Harry Bunt, eds., *Computing Meaning*, volume Volume 3. Kluwer, Dordrecht.

Kadmon, Nirit (1990): 'Uniqueness.' *Linguistics and Philosophy*, **13**, 273–324.

Kamp, Hans (1981): 'A theory of truth and semantic representation.' In J. Groenendijk, T. Janssen, and M. Stokhof, eds., *Formal Methods in the Study of Language*. Mathematisch Centrum, Amsterdam.

Kamp, Hans and Uwe Reyle (1993): *From Discourse to Logic. Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Kluwer, Dordrecht.

Kamp, Hans and Antje Rossdeutscher (1994): 'Remarks on lexical structure and DRS construction.' *Theoretical Linguistics*, **20**, 97–164.

Kamp, Hans, Josef van Genabith, and Uwe Reyle (2005): 'Discourse representation theory.' In Dov Gabbay and Franz Guenthner, eds., *Handbook of Philosophical Logic*. Kluwer.

Kaplan, David (1977): 'Demonstratives.' In J. Almog, J. Perry, and H. Wettstein, eds., *Themes from Kaplan*, pp. 481–563. Oxford University Press, Oxford, 1989 edition.

Karttunen, Lauri (1976): 'Discourse referents.' In James McCawley, ed., *Syntax and Semantics 7, Notes from the Linguistic Underground*, pp. 363–385. Academic Press, New York.

Keenan, Elinor Ochs and Bambi Schieffelin (1976a): 'Foregrounding referents: a reconsideration of left dislocation in discourse.' *Berkeley Linguistics Society*, **2**, 240–257.

Keenan, Elinor Ochs and Bambi Schieffelin (1976b): 'Topic as a discourse notion: A study of topic in the conversations of children and adults.' In Charles Li, ed., *Subject and topic*, pp. 336–384. Academic Press, New York.

Kehler, Andrew (2002): *Coherence, Reference, and the Theory of Grammar*. CSLI publications, Stanford, CA.

Kehler, Andrew (2004): 'Discourse topics, sentence topics, and coherence.' *Theoretical Linguistics*, **30**(2-3), 227–240. doi:10.1515/thli.2004.30.2-3.227.

Keysar, Boaz, D. Barr, J. Balin, and J. Brauner (2000): 'Taking perspective in conversation: The role of mutual knowledge in comprehension.' *Psychological Science*, **11**(1), 32–38.

Kleiber, Georges (1997): 'Des anaphores associatives méronymiques aux anaphores associatives locatives.' *Verbum (Presses Universitaires de Nancy)*, **1-2**, 25–66.

Kleiber, Georges (1999): 'Associative anaphora and part-whole relationship: The condition of alienation and the principle of ontological congruence.' *Journal of Pragmatics*, **31**(3), 339–362.

Klein, Wolfgang and Christiane von Stutterheim (1987): 'Quaestio und referentielle Bewegung in Erzählungen.' *Linguistische Berichte*, **109**, 163–183.

Knott, Alistair (1996): *A Data-Driven Methodology for Motivating a Set of Coherence Relations*. Ph.D. thesis, University of Edinburgh.

Knott, Alistair and Robert Dale (1994): 'Using linguistic phenomena to motivate a set of coherence relations.' *Discourse Processes*, **18**(1), 35–62.

Knott, Alistair, Jon Oberlander, Michael O'Donnell, and Chris Mellish (2001): 'Beyond elaboration: The interaction of relations and focus in coherent text.' In Sanders et al. (2001), pp. 181–196.

Knott, Alistair and Ted Sanders (1998): 'The classification of coherence relations and their linguistic markers: An exploration of two languages.' *Journal of Pragmatics*, **30**(2), 135–175.

Koenig, Jean-Pierre and Gail Mauner (1999): 'A-definites and the discource status of implicit arguments.' *Journal of Semantics*, **16**(3), 207–236.

Kohlhase, Michael (2000): 'Model generation for discourse representation theory.' In Werner Horn, ed., *ECAI-2000*. Berlin.

Konrad, Karsten (2000): *Model Generation for Natural Language Interpretation and Analysis*. Ph.D. thesis, Universität des Saarlandes, Saarbrücken.

Krifka, Manfred (1992): 'A compositional semantics for multiple focus constructions.' In Joachim Jacobs, ed., *Informationsstruktur und Grammatik*, pp. 17–53. Westdeutscher Verlag, Opladen.

Krifka, Manfred (1998): 'Additive particles under stress.' In Devon Strolovitch and Aaron Lawson, eds., *Proceedings of SALT VIII*, pp. 111–129. CLC Publications, Ithaca, NY.

Kruijff-Korbayová, I. and Eva Hajičová (1997): 'Topics and centers: A comparison of the salience-based approach and the centering theory.' *Prague Bulletin of Mathematical Linguistics*, **67**, 25–50.

Lakoff, George (1987): *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. University of Chicago Press, Chicago.

Langacker, Ron (1987): *Foundations of Cognitive Grammar, vol. 1: Theoretical Prerequisites*. Stanford University Press, Stanford.

Langacker, Ron (2008): *Cognitive Grammar: A Basic Introduction*. Oxford University Press, New York.

Lappin, Shalom (1996): 'The interpretation of ellipsis.' In Shalom Lappin, ed., *The Handbook of Contemporary Semantic Theory*, pp. 145–175. Blackwell, Oxford.

Lascarides, Alex and Nicholas Asher (1991): 'Discourse relations and defeasible knowledge.' In *Proceedings to the 29th Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 55–62.

Lascarides, Alex and Nicholas Asher (2007): 'SDRT: Dynamic semantics with discourse structure.' In H. Bunt and R. Muskens, eds., *Computing Meaning: Volume 3*. Kluwer, Tilburg.

Levelt, Willem (1989): *Speaking: From Intention to Articulation*. MIT Press, Cambridge, MA.

Levinson, Stephen (2000): *Presumptive Meanings*. MIT Press, Cambridge.

Lewis, David (1969): *Convention: A Philosophical Study*. Havard University Press, Cambridge, MA.

Lewis, David (1973): *Counterfactuals*. Havard University Press, Cambridge, MA.

Lewis, David (1979): 'Scorekeeping in a language game.' *Journal of Philosophical Logic*, **8**, 339–359.

Lifschitz, Vlamidir (1989): 'Circumscriptive theories: A logic-based framework for knowledge representation.' In R. H. Thomason, ed., *Philosophical Logic and Artificial Intelligence*, pp. 109–159. Kluwer, Dordrecht.

Link, Godehard (1983): 'The logical analysis of plurals and mass terms: A lattice-theoretical approach.' In Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow, eds., *Meaning, Use, and Interpretation of Language*, pp. 302–323. de Gruyter, Berlin.

López, Luis (2006): 'Is there a topic in this sentence?' Talk given at University of Leipzig, Germany, 03-05-2006.

López, Luis (2009): *A Derivational Syntax for Information Structure*. Oxford University Press, Oxford.

Maienborn, Claudia (2001): 'On the position and interpretation of locative modifiers.' *Natural Language Semantics*, **9**(2), 191–240.

Mann, William and Sandra Thompson (1988): 'Rhetorical Structure Theory: toward a functional theory of text organization.' *Text*, **8**(3), 243–281.

Marcu, Daniel (1997): *The Rhetorical Parsing, Summarization, and Generation of Natural Language Texts*. Ph.D. thesis, University of Toronto.

Marcu, Daniel (2000): *The Theory and Practice of Discourse Parsing and Summarization*. MIT Press, Cambridge, MA.

Marín, Francisco Marcos (1992): *CRLEC: Corpus de referencia de la lengua española contemporanea*. Universidad Autónoma de Madrid, Spain. Available at http://www.lllf.uam.es/corpus/corpus_oral.html.

Matsui, Tomoko (2000): *Bridging and Relevance*. John Benjamins, Amsterdam.

Mauner, Gail, Michael Tanenhaus, and Greg Carlson (1995): 'Implicit arguments in sentence processing.' *Journal of Memory and Language*, **34**(3), 357–382.

McCarthy, John (1980): 'Circumscription: a form of non-monotonic reasoning.' *Artificial Intelligence*, **13**, 27–39.

Minsky, Marvin (1975): 'A framework for representing knowledge.' In P. H. Winston, ed., *The Psychology of Computer Vision*, pp. 211–277. McGraw-Hill, New York.

Montague, Richard (1973): 'The proper treatment of quantification in ordinary English.' In Jaakko Hintikka, Julius Moravcsik, and Patrick Suppes, eds., *Approaches to Natural Language*, pp. 221–242. Dordrecht.

Moore, Robert C. (1985): 'Semantical considerations on nonmonotonic logic.' *Artifical Intelligence*, **25**(1), 75–94.

Muskens, Reinhard (1996): 'Combining montague semantics and discourse representation.' *Linguistics and Philosophy*, **19**, 143–186.

Nissim, Malvina (2001): *Bridging Definites and Possessives: Distribution of Determiners in Anaphoric Noun Phrases*. Ph.D. thesis, Department of Linguistics, University of Pavia, Italy.

Nissim, Malvina, Andrea Sansò, and Claudia Soria (1999): 'Towards a compositional frame semantics.' In *Proceedings of the European Conference on Cognitive Science (ECCS99)*. Siena, Italy.

Nunberg, Geoffrey (1995): 'Transfers of meaning.' *Journal of Semantics*, **12**, 109–132.

Oberlander, Jon (2004): 'On the reduction of discourse topic.' *Theoretical Linguistics*, **30**(2-3), 213–225. doi:10.1515/thli.2004.30.2-3.213.

Padilla García, Xose (2001): *El orden de palabras en el español coloquial*. Ph.D. thesis, Universitat de València.

Parikh, Prashant (2001): *The Use of Language*. CSLI Publications, Stanford, CA.

Parsons, Terence (1990): *Events in the Semantics of English*. MIT Press, Cambridge, MA.

Peirce, Charles Sanders (1903): *Pragmatism – The Logic of Abduction*. The Harvard Lectures. Harvard University, Cambridge, MA. Reprinted in Peirce (1934).

Peirce, Charles Sanders (1934): 'Pragmatism and pragmaticism.' In C. Hartshorne and P. Weiss, eds., *Collected Papers of Charles Sanders Peirce.*, volume V. Harvard University Press, Cambridge, MA.

Pickering, Martin and Simon Garrod (2004): 'Toward a mechanistic psychology of dialogue.' *Behavioral and Brain Sciences*, **27**(2), 169–226.

Piwek, Paul and Emiel Krahmer (2000): 'Presuppositions in context: Constructing bridges.' In Pierre Bonzon, Marcos Cavalcanti, and Rolf Nossum, eds., *Formal Aspects of Context*, pp. 85–106. Kluwer Academic Publishers, Dordrecht.

Poesio, Massimo and David Traum (1997): 'Conversational actions and discourse situations.' *Computational Intelligence*, **13**(3), 309–347.

Poesio, Massimo and Renata Vieira (1998): 'A corpus-based investigation of definite descriptions use.' *Computational Linguistics*, **24**(2), 183–216.

Polanyi, Livia (1988): 'A formal model of the structure of discourse.' *Journal of Pragmatics*, **12**(5-6), 601–638.

Polanyi, Livia, Martin Van Den Berg, and David Ahn (2003): 'Discourse structure and sentential information structure.an initial proposal.' *Journal of Logic, Language and Information*, **12**(3), 337–350.

Polanyi, Livia and Remko Scha (1984): 'A syntactic approach to discourse semantics.' In *Proceedings of the 10th international conference on Computational linguistics*, pp. 413–419. Association for Computational Linguistics.

Prince, Alan and Paul Smolensky (1993): 'Optimality theory: Constraint interaction in generative grammar.' Technical report, Rutgers University.

Prince, Ellen (1981): 'Towards a taxonomy of given-new information.' In Peter Cole, ed., *Radical Pragmatics*, pp. 223–256. Academic Press, New York.

Prince, Ellen (1992): 'The ZPG letter: Subjects, definiteness, and information-status.' In Sandra Thompson and William Mann, eds., *Discourse description: diverse analyses of a fund raising text*, pp. 295–325. John Benjamins, Amsterdam.

Prince, Ellen (1997): 'On the functions of Left-Dislocation in English discourse.' In A. Kamio, ed., *Directions in functional linguistics*, pp. 117–144. John Benjamins, Amsterdam.

Pustejovsky, James (1995): *The Generative Lexicon*. MIT Press.

Pylkkänen, Liina and Brian McElree (2006): 'The syntax-semantics interface: On-line composition of sentence meaning.' In M. J. Traxler and M. A. Gernsbacher, eds., *Handbook of Psycholinguistics*, pp. 537–577. Elsevier, New York.

Rappaport, Malka and Beth Levin (1988): 'What to do with theta-roles.' In Wendy Wilkins, ed., *Thematic Relations*, volume 21 of *Syntax and Semantics*, pp. 7–36. Academic Press, San Diego.

Real Academia Española (2008): *Banco de datos (CREA). Corpus de referencia del español actual*. http://www.rae.es.

Recanati, François (2003): 'Embedded implicatures.' *Philosophical Perspectives*, **17**(1), 299–332.

Recanati, François (2004): *Literal Meaning*. Cambridge University Press, Cambridge.

Reese, B., P. Denis, Nicholas Asher, J. Baldridge, and J. Hunter (2007): 'Reference manual for the analysis and annotation of rhetorical structure.' Technical report, University of Texas at Austin.

Reinhart, Tanya (1982): 'Pragmatics and Linguistics: An Analysis of Sentence Topics.' *Philosophica*, **27**, 53–94.

Reiter, Raymond (1980): 'A logic for default reasoning.' *Artificial Intelligence*, **13**, 81–132.

Rizzi, Luigi (1997): 'The fine structure of the left periphery.' In Liliane Haegeman, ed., *Elements of Grammar*, pp. 289–330. Kluwer, Dortrecht.

Roberts, Craige (1996): 'Information structure in discourse: Towards an integrated formal theory of pragmatics.' In Jae-Hak Yoon and Andreas Kathol, eds., *Papers in Semantics*, volume 49 of *OSU Working Papers in Linguistics*. Ohio State University.

Roberts, Craige (2003): 'Uniqueness in definite noun phrases.' *Linguistics and Philosophy*, **26**, 287–350.

Rooth, Mats (1985): *Association with Focus*. Ph.D. thesis, University of Massachusetts at Amherst.

Ross, John Robert (1967): *Constraints on Variables in Syntax*. PhD dissertation, MIT, Cambridge, MA.

Ruppenhofer, Josef, Michael Ellsworth, Miriam Petruck, Christopher Johnson, and Jan Scheffczyk (2005): 'FrameNet II: Extended theory and practice.' Technical report, ICSI, Berkeley.

Russell, Bertrand (1905): 'On Denoting.' *Mind*, **14**, 479–493.

Sag, Ivan (1981): 'Formal semantics and extralinguistic context.' In Pter Cole, ed., *Radical Pragmatics*. Academic Press, New York.

Sanders, Ted (1997): 'Semantic and pragmatic sources of coherence: On the categorization of coherence relations in context.' *Discourse Processes*, **24**, 119–147.

Sanders, Ted, Joost Schilperoord, and Wilbert Spooren, eds. (2001): *Text Representation: Linguistic and Ppsycholinguistic Aspects*. John Benjamins, Amsterdam.

Sanders, Ted and Wilbert Spooren (2001): 'Text representation as an interface between language and its uses.' In Sanders et al. (2001), pp. 1–25.

Sanders, Ted, Wilbert Spooren, and Leo Noordman (1992): 'Toward a taxonomy of coherence relations.' *Discourse Processes*, **15**(1), 1–35.

Sanford, A. J., Simon Garrod, A. Lucas, and R. Henderson (1993): 'Pronouns without explicit antecedents?' *Journal of Semantics*, **2**(3/4), 303–318.

Schank, Roger and Robert Abelson (1977): *Scripts, Plans, Goals, and Understanding*. Earlbaum, Hillsdale, NJ.

Schwarz, Monika (2000): *Indirekte Anaphern in Texten: Studien zur domänengebundenen Referenz und Kohärenz im Deutschen*. Niemeyer, Tuebingen.

Schwarz-Friesel, Monika (2007): 'Indirect anaphora in text. A cognitive account.' In Monika Schwarz-Friesel, Manfred Consten, and Mareile Hillevi Knees, eds., *Anaphors in Text*, Studies in Language Companion Series 86. John Benjamins, Philadelphia.

Searle, John (1969): *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press., Cambridge.

Sidner, Candace (1981): 'Focusing for interpretation of pronouns.' *American Journal of Computational Linguistics*, **7**, 217–231.

Simons, Patrik, Ilkka Niemelá, and Timo Soininen (2002): 'Extending and implementing the stable model semantics.' *Artificial Intelligence*, **138**(1-2), 181–234.

Singer, Murray (1979): 'Processes of inference during sentence encoding.' *Memory & Cognition*, **7**(3), 192–200.

Sperber, Dan and Deirdre Wilson (1986): *Relevance: Communication and Cognition*. Blackwell, Oxford.

Stalnaker, Robert (1968): 'A theory of conditionals.' In B. Rescher, ed., *Studies in Logical Theory*. Oxford.

Stalnaker, Robert (1974): 'Pragmatic presuppositions.' In M. Munitz and P. Unger, eds., *Semantics and Philosophy*. New York University Press, New York.

Stalnaker, Robert (1978): 'Assertion.' *Syntax and Semantics*, **9**, 315–332.

Stede, Manfred (2004a): 'Does discourse processing need discourse topics?' *Theoretical Linguistics*, **30**(2-3), 241–253. doi:10.1515/thli.2004.30.2-3.241.

Stede, Manfred (2004b): 'The Potsdam commentary corpus.' In *Proceedings of the Workshop on Discourse Annotation, 42nd Meeting of the Association for Computational Linguistics*. Barcelona, Spain.

Steedman, Mark (2000): 'Information structure and the syntax-phonology interface.' *Linguistic Inquiry*, **31**, 649–689.

Steube, Anita (2003): 'Bridge contours in german assertive main clauses.' *Folia Linguistica*, **37**(1/2), 215–249.

Stickel, Mark (1991): 'A Prolog-like inference system for computing minimum-cost abductive explanations in natural-language interpretation.' *Annals of Mathematics and Artificial Intelligence*, **4**, 89–106.

Stone, Matthew and Richard Thomason (2002): 'Context in abductive interpretation.' In Bos, Foster, and Matheson, eds., *Proceedings to EDILOG*. Edinburgh.

Strube, Michael and Udo Hahn (1999): 'Functional centering: Grounding referential coherence in information structure.' *Computational Linguistics*, **25**, 309–344.

Subirats Rüggeberg, Carlos (2005): 'FrameNet español. Una red semántica de marcos conceptuales.' In E. Serra y Gerd Wotjak, ed., *Cognición y percepción lingüísticas. Comunicaciones presentadas al VI Congreso Internacional de Lingüística Hispánica, Leipzig, 2003*, pp. 182–196. Universidad de Valencia y Universidad de Leipzig, Valencia.

Taboada, Maite and William Mann (2006a): 'Applications of rhetorical structure theory.' *Discourse Studies*, **8**(4), 567–588.

Taboada, Maite and William Mann (2006b): 'Rhetorical structure theory: Looking back and moving ahead.' *Discourse Studies*, **8**(3), 423–459.

Thomason, Richard (1997): 'Nonmonotonicity in linguistics.' In Johan van Benthem and Alice ter Meulen, eds., *The Handbook of Logic and Language*. Elsevier, Amsterdam.

Traum, David and James Allen (1992): 'A Speech Acts approach to grounding in conversation.' In *Proceedings to the International Conference on Spoken Language Processing (ICSLP'92)*, pp. 137–140.

Uriagereka, Juan (1995): 'Aspects of the syntax of clitic placement in western romance.' *Linguistic Inquiry*, **26**(1), 79–123.

Vallduví, Enric (1992): *The Informational Component*. Garland Publishing, New York.

van den Broek, Paul, K. Risden, C. Fletcher, and R. Thurlow (1996): 'A 'landscape' view of reading: Fluctuating patterns of activation and the construction of a stable memory representation.' In Britton and Graesser (1996), pp. 165–187.

van der Sandt, Rob (1992): 'Presupposition projection as anaphora resolution.' *Journal of Semantics*, **9**, 333–377.

van Dijk, Teun A. (2008): *Discourse and Context: A Socio-Cognitive Approach.* Cambridge University Press, Cambridge.

van Dijk, Teun A. and Walter Kintsch (1983): *Strategies of Discourse Comprehension.* Academic Press, New York.

van Eijck, Jan and Hans Kamp (1997): 'Representing discourse in context.' In J. van Benthem and A. ter Meulen, eds., *Handbook of Logic and Linguistics*, pp. 179–237. Elsevier, Amsterdam.

van Kuppevelt, Jan (1995): 'Discourse structure, topicality and questioning.' *Journal of Linguistics*, **31**, 109–147.

van Riemsdijk, Henk (1997): 'Left dislocation.' In Anagnostopoulou et al. (1997), pp. 1–10.

Vendler, Zeno (1957): 'Verbs and times.' *Philosophical Review*, **66**(2), 143–160.

Vieu, Laure and Michel Aurnague (2007): 'Part-of relations, functionality, dependence.' In M. Aurnague, M. Hickmann, and L. Vieu, eds., *The Categorization of Spatial Entities in Language and Cognition*, pp. 307–336. Benjamins, Amsterdam.

Vieu, Laure and Laurent Prévot (2004): 'Background in SDRT.' In *Workshop SDRT, Traitement Automatique du Langage Naturel, Fez, Morocco*, pp. 485–494.

Vieu, Laure and Laurent Prévot (2008): 'The moving right frontier.' In Anton Benz and Peter Kühnlein, eds., *Constraints in Discourse*, pp. 53–66. Benjamins, Amsterdam.

Villalba, Xavier (2000): *The syntax of sentence periphery.* Ph.D. thesis, Department of Catalan Philology, Universitat Autònoma de Barcelona.

von Stutterheim, Christiane and Wolfgang Klein (2002): 'Quaestio and l-perspectivation.' In Carl Friedrich Graumann and Werner Kallmeyer, eds., *Perspective and Perspectivation in Discourse.* John Benjamins, Amsterdam.

Walker, Marilyn A. (2000): 'Toward a model of the interaction of centering with global discourse structure.' *Verbum.*

Walker, Marilyn A., Aravind Joshi, and Ellen Prince, eds. (1998): *Centering Theory in Discourse.* Academic Press, New York.

Webber, Bonnie (1988): 'Discourse deixis: reference to discourse segments.' In *Proceedings of the 26th ACL meeting*, pp. 113–122. doi:http://dx.doi.org/10.3115/982023.982037.

Webber, Bonnie (2004): 'D-LTAG: extending lexicalized TAG to discourse.' *Cognitive Science*, **28**(5), 751–779.

Webber, Bonnie, Matthew Stone, Aravind Joshi, and Alistair Knott (2003): 'Anaphora and discourse structure.' *Computational Linguistics*, **29**(4), 545–587.

Williams, Edwin (1997): 'Blocking and anaphora.' *Linguistic Inquiry*, **28**, 577–628.

Wilson, Deirdre and Tomoko Matsui (1998): 'Recent approaches to bridging: Truth, coherence, relevance.' *UCL Working Papers in Linguistics*, **10**.

Winston, M., R. Chaffin, and D. Herrmann (1987): 'A taxonomy of part-whole relations.' *Cognitive Science*, **11**, 417–444.

Wolf, Florian and Edward Gibson (2005): 'Representing discourse coherence: A corpus-based analysis.' *Computational Linguistics*, **31**(2), 249–287.

Wolf, Florian and Edward Gibson (2006): *Coherence in Natural Language: Data Structures and Applications*. MIT Press, Cambridge, MA.

Wolf, Florian, Edward Gibson, and Timothy Desmet (2004): 'Discourse coherence and pronoun resolution.' *Language and Cognitive Processes*, **19**(6), 665–675.

Zawisławska, Magdalena, Magdalena Derwojedowa, and Jadwiga Linde-Usiekniewicz (2008): 'A FrameNet for Polish.' In Doris Schönefeld, Matthias Irmer, and Christina Witt, eds., *Converging Evidence: Proceedings to the Third International Conference of the German Cognitive Linguistics Association (GCLA'08)*, pp. 116–117. University of Leipzig, Leipzig.

Zeevat, Henk (1989): 'A compositional approach to discourse representation theory.' *Linguistics and Philosophy*, **12**, 95–131.

Zeevat, Henk (2004): 'Asher on discourse topic.' *Theoretical Linguistics*, **30**(2-3), 203–211. doi:10.1515/thli.2004.30.2-3.203.

Zeevat, Henk (2006): 'Discourse structure in optimality theoretic pragmatics.' In John Harpur, Anton Benz, Candace Sidner, and Peter Kühnlein, eds., *Proceedings of Constraints in Discourse II (CID-06), Maynooth, Ireland*, pp. 155–161.

Zeevat, Henk (2009): 'Optimal interpretation as an alternative to Gricean pragmatics.' In Behrens and Fabricius-Hansen, eds., *Structuring Information in Discourse: The Explicit/Implicit Dimension*, Oslo Studies in Language, pp. 191–216. University of Oslo, Norway.

Zeevat, Henk and Elena Karagjosova (2007): 'The history and grammaticalisation of 'doch'/'toch'.' Electronic proceedings of FSIM2.

Ziv, Yael (1994): 'Left and right dislocations: Discourse functions and anaphora.' *Journal of Pragmatics*, **22**(6), 629–645.