Semantic Underspecification and Ellipsis Resolution

Diplomarbeit
Angefertigt unter der Leitung von Prof. Dr. Manfred Pinkal

vorgelegt von:
Fei-Yu Xu
Hiermit erkläre ich, daß ich diese Arbeit selbständig verfaßt und keine anderen als die angegebenen Hilfsmittel verwendet habe.

Saarbrücken, den 12ten März 1998

Fei-Yu Xu
Abstract

This thesis deals with the underspecified semantic representation and resolution of ellipsis. We focus on the interaction of Verb Phrase Ellipsis (henceforth VPE) and anaphora. In particular, we address the distribution of strict and sloppy readings in these phenomena and present an analysis to explain it. This analysis is integrated into a framework that enables a uniform treatment of semantically underspecified information and parallelism phenomena in discourse.

The analysis put forward in this thesis makes correct predictions about the strict and sloppy readings of a series of problematic cases in the literature. It overcomes undergeneration or overgeneration problems. We use linking relations between anaphoric pronouns and their antecedents and claim that the interaction between linking relations and the parallelism requirement in VPE results in strict and sloppy readings. The analysis presented in this thesis agrees with the proposals in (Kehler 1995; Hobbs and Kehler 1997). We integrated our analysis into an extension of the approach of Niehren et al. (1997), called Constraint Language for Lambda-link Structures (CLLS) (Egg et al. 1998). CLLS provides a simple means of handling anaphora in VPE. CLLS can express dominance constraints, binding constraints, linking constraints and parallelism constraints over lambda-link structures. Lambda-link structures are tree structures with additional relations to express lambda-binding and anaphoric links. We use dominance constraints and binding constraints to express semantically underspecified information such as scope ambiguities. The linking constraints are applied to describe the linking relations of pronouns and their antecedents. The parallelism constraints are used to express the parallelism between semantic representations. The advantage of CLLS is that it gives a simple, uniform treatment of scope underspecification, parallelism and anaphora without over- and undergeneration.

Although recently a lot of approaches to semantic underspecification as well as ellipsis resolution have emerged in computational linguistics, few of them are designed to treat both problems uniformly. CLLS shows how to integrate the treatment of ellipsis resolution with semantic underspecification in a uniform and elegant way.
Acknowledgments

I wish to thank all those who have helped and encouraged me during my studies at the University of the Saarland.

Above all, I am grateful to my supervisor Manfred Pinkal - not only for numerous fruitful discussions during the development of my thesis, but also for his constant encouragement and support during my entire studies. I also thank Hans Uszkoreit for his substantial proposals for my studies and writing thesis.

I am especially indebted to Markus Egg for his valuable comments and suggestions on the subject of this thesis. Special thanks are due to Jan Lerner. This thesis has benefitted very much from his suggestions and comments. I am also grateful to Peter Ruhrberg and Joachim Niehren for their discussions and suggestions. I would like to thank Claire Gardent for her comments on my thesis.

A special debt of gratitude to Johan Bos who was always willing to help me by improving my English writing. I am grateful to C. J. Rupp. As a English native speaker, C. J. has helped a lot by the revision of this thesis.

Many thanks are due to those who gave comments on the drafts of this thesis: Judith Baur, Markus Becker, Petra Dünges, Daniela Kurz, Wojciech Skut, Susanna Kuschert, Stephan Oepen, Claudia Verburg and Verena Winter. Thanks and apologies to all those important people I forgot.

Finally, extra special thanks to my husband, who kept supporting and encouraging me during my studies. I am grateful to my parents for their understanding and encouragement.
Contents

1 Introduction ................................................. 2
   1.1 VPE and anaphora ........................................ 3
   1.2 Uniform Treatment of Semantic Underspecification and Parallelism ........ 5
   1.3 CHORUS .................................................. 8
   1.4 Overview ............................................... 9

2 Approaches to Verb Phrase Ellipsis and Anaphora ........... 10
   2.1 VPE and Anaphora ........................................ 11
   2.2 Williams' Analysis ...................................... 13
   2.3 The Analysis of Dalrymple, Shieber, and Pereira ............ 16
   2.4 Kehler's Analysis ...................................... 20
   2.5 The Parallelism Theory of Hobbs and Kehler ................ 26
   2.6 Conclusion ............................................. 31

3 Underspecified Treatment of Ambiguities .................. 33
   3.1 Introduction ............................................ 33
   3.2 Scope Ambiguities and Verb Phrase Ellipsis ................ 34
   3.3 Hole Semantics ......................................... 35
   3.4 The Framework of Context Unification ..................... 38
   3.5 Conclusion ............................................. 43

4 Underspecified Treatment of Verb Phrase Ellipsis and its Interaction with Anaphora and Quantifiers ........... 45
   4.1 A Constraint Language for Lambda-link Structures .......... 45
CONTENTS

4.2 Underspecified Representation and Resolution of Verb Phrase Ellipsis and Anaphora ................................................. 50
4.3 Applications ................................................................. 63
4.4 Extensions of the Approach ............................................... 67
4.5 Conclusion ................................................................. 71

5 Conclusion and Outlook .................................................... 72

Bibliography ........................................................................ 74
Chapter 1

Introduction

This thesis deals with the underspecified semantic representation and resolution of ellipsis. We focus on the interaction of Verb Phrase Ellipsis (henceforth VPE) and anaphora. In particular, we address the distribution of strict and sloppy readings in these phenomena and present an analysis to explain it. This analysis is integrated into a framework that focuses on a uniform treatment of semantically underspecified information and parallelism phenomena in discourse.

Natural language is known for the ambiguity of its expressions. Ambiguity can arise from various different sources: it may be lexical (as in the cases of lexical homonymy and polysemy), structural (as in the cases of syntactic and scope ambiguities) or contextual (as in anaphoric ambiguities). Combinations of such different sources occur as well.

Traditional theories of natural language semantics involve mapping natural language utterances into expressions of an unambiguous representation language (logical formulas). Within these theories, interpreting an ambiguous utterance is typically modeled by choosing the intended interpretation from a plurality of possible ones. As discussed in Pinkal (1996a), this generate-and-test method is neither computationally feasible nor cognitively plausible. Therefore, underspecified semantic representation has received increasing attention in the last years. The aim of underspecified semantic representation is to produce compact representations of the set of possible readings of a discourse. Several formalisms for underspecified semantic representation have emerged in the last few years. Pinkal (1996b) summarizes that their common basis is the description of a range of possible readings in terms of a constraint set; and the resolution process is modeled as a stepwise, monotonic extension of this constraint set, which corresponds to the subsequent reduction of the range of possible readings.

When anaphora and quantifiers are involved in VPE, the problem of ambiguities can be very complex. To understand VPE, the meaning of the elided verb phrase in the elliptical clause must be recovered from the context, while obeying a certain parallelism between the antecedent clause and the elliptical expression. Therefore, the semantic underspecification formalism must consider the parallelism requirement in ellipsis in order to deal with various elliptical phenomena.
1.1 VPE and anaphora

An illustration of VPE is sentence (1):

(1) Dan likes Mary, and Bill does too.

In this example, ‘Dan likes Mary’, is the antecedent clause (henceforth, the source clause) whereas ‘Bill does too’ is the elliptical one (henceforth, the target clause). The stranded auxiliary in the target clause indicates the deletion of a Verb Phrase (VP). ‘Dan’ and ‘Bill’ are the so-called parallel elements. The parallel elements are the constituents in the source clause and their overt parallel counterparts in the target clause. Although its VP is unexpressed, the target clause in this discourse certainly means that Bill likes Mary too. The parallelism holding between the two clauses enforces this reading. The interpretation of VPE must be constrained by parallelism, which is an essential feature of ellipsis. However, parallelism constraints go beyond VP-identity. Anaphoric elements in the source clause, such as pronouns, cause additional readings for the target clause. For instance, consider sentence (2):

(2) John revised his paper before the teacher did.

The interesting point about this example is the interpretation of the elided pronoun. Assuming that the overt pronoun ‘his’ refers to the source parallel element ‘John’, there can be two readings for the target clause, namely one in which the teacher revised John’s paper (the so-called strict reading), and one in which the teacher revised the teacher’s paper (known as the sloppy reading). The terms strict and sloppy refer to the mode of interpretation of an elided pronoun in the target clause. If its interpretation is the same as that of its parallel pronoun in the source clause, it is called strict; otherwise sloppy.

In our analysis, the strict and sloppy readings arise from the interaction between linking relations in the source clause and the parallelism requirement in ellipsis. A linking relation describes the referential relation between a pronoun and its antecedent. The parallelism requirement is given as follows.

If the semantic interpretation of a pronoun α in the source VP of VPE is linked to the semantic interpretation of a constituent β, then the semantic interpretation of its parallel pronoun α′ in the target VP can be either

(i) linked to the semantic interpretation of α (henceforth, referential parallelism)

or

(ii) linked to the semantic interpretation of β′ to establish a parallel link, with β and β′ being structurally parallel (henceforth, structural parallelism).\(^2\)

---

\(^1\)We represent co-reference by co-indexing.

\(^2\)We adopt the terms referential parallelism and structural parallelism from Fox (1996). But the definition of parallelism in our approach is different from that in Fox (1996). Parallelism in Fox (1996) is a relation between the syntactic representations of the source clause and the target clause (see section 2.4).
1.1 VPE and anaphora

\[ \text{John}_1 \text{ revised his}_1 \text{ paper} \]

\[ \text{the teacher}_2 \text{ revised his paper.} \]

Figure 1.1: Reconstruction of Links for Sentence (2)

For example, there exists a linking relation in the source clause of sentence (2). That is, the interpretation of the pronoun ‘his’ is linked to the interpretation of the source parallel element ‘John’. In Figure 1.1, we represent this link as an arrow pointing from ‘his’ to ‘John’. According to the parallelism requirement, the link in the source clause gives rise to two possible ways to reconstruct a new linking relation for the elided pronoun in the target clause. We represent the two linking options in Figure 1.1 as two dashed arrows pointing from the elided pronoun ‘his’. This means that the interpretation of the elided pronoun is either linked to that of its parallel pronoun (referential parallelism) or linked to that of the target parallel element ‘the teacher’ (structural parallelism). In the case of referential parallelism, the elided pronoun refers to the source pronoun which refers to ‘John’, yielding the strict reading such as

(3) John revised John’s paper before the teacher revised John’s paper.

In the case of structural parallelism, the elided pronoun refers to the target parallel element ‘the teacher’, resulting in the sloppy reading:

(4) John revised John’s paper before the teacher revised the teacher’s paper.

Not all sentences are as simple as (2). There are a series of problematic cases of strict and sloppy ambiguities mentioned in the literature. In this introductory chapter we present one of them. It is the well-known many-pronouns-puzzle, given in (5).

(5) John\(_1\) said that he\(_1\) liked his\(_1\) mother, and Bill did too.

One might expect that there are \(2^n\) readings for the target clause for examples in which there are \(n\) pronouns in the source clause which are coreferent with the source parallel element. However, Dahl (1974) noticed that example (5) has only three readings, rather than four (\(2^2\)). They are enumerated in (6).

(6) 1. John said that John liked John’s mother, and
     Bill said that John liked John’s mother.
     (strict-strict)
2. John said that John liked John’s mother, and
   Bill said that Bill liked John’s mother.
   (sloppy-strict)

3. John said that John liked John’s mother, and
   Bill said that John liked Bill’s mother.
   (sloppy-sloppy)

The fourth reading, which is unavailable is the strict-sloppy reading, given in (7).

(7) *John said that John liked John’s mother, and
    Bill said that John liked Bill’s mother.
    (strict-sloppy)

Sentences like (5) are often used as benchmarks to test the adequacy of an ellipsis analysis for interaction of VPE and anaphora. Our analysis accounts for the unavailable reading in (5) too: There exists two linking relations in the source clause. One is between ‘John’ and ‘he’ and the other is between ‘he’ and ‘his’. The interaction of the linking relations in the source clause with the parallelism requirement eliminates (7).

Our analysis makes correct predictions about strict and sloppy readings of a series of problematic cases in the literature and overcomes undergeneration or overgeneration problems. The idea underlying our analysis agrees with the proposals made in (Kehler 1995; Hobbs and Kehler 1997). However, the frameworks used there cannot treat the ellipsis problem in an underspecified way and they cannot be extended straightforwardly to deal with other semantic underspecification phenomena, such as scope ambiguities. Our approach integrates the VPE analysis into a general underspecification mechanism which covers the interaction of scope, anaphora and parallelism. In the following section, we will describe the framework into which our analysis is integrated.

1.2 Uniform Treatment of Semantic Underspecification and Parallelism

Semantic underspecification allows a compact representation of the set of possible readings of a natural language expression. Recently, a lot of theories for semantic underspecification have emerged in computational linguistics. Most of them concentrate on lexical or scope ambiguities. Niehren et al. (1997) propose a framework of Context Unification (CU) which deals with phenomena of quantifier scope, VPE and their interaction in a uniform and underspecified way. However, CU does not cope with anaphora. The problem is that the constraint language proposed in CU cannot express the anaphoric link, which is crucial for the ellipsis analysis. The Constraint Language for Lambda-link Structures (CLLS) (Egg et al. 1998) extends the approach of CU and integrates our ellipsis analysis in a straightforward way.
The Approach of Context Unification

In Niehren et al. (1997), the constraint language of CU is used as a semantic description language (meta language) for semantic underspecification which takes the typed higher order language (HOL) (Dowty et al. 1981) or some other logical language as its object language. It can express context constraints over finite trees. The trees represent formulas of some object language. For example, the tree in (8), illustrated in Figure 1.2, represents the HOL formula, given in (9).

(8) \( \text{john} @ \text{lam}_x ((\text{his} @ \text{paper}) @ \text{lam}_y ((\text{revise} @ \text{var}_y) @ \text{var}_x)) \)

(9) \( \text{john}(\lambda x.((\text{his}(\text{paper}))(\lambda y.(((\text{revise}(y))(x))))) ) \)

(8) consists of binary function symbols such as @, unary function symbols (\text{lam}_x), and symbols which denote variables and constants of the object language. The binary function symbol @ describes the functional application of HOL. The unary function symbols \text{lam}_x or \text{lam}_y describe \( \lambda \)-abstraction over the variables \( x \) or \( y \) in HOL. Other symbols in (8) stand for the corresponding constants of the object language. Naturally, a tree consists of subtrees. For instance, the above tree has \text{his}@\text{paper} as one of its subtrees.

The constraint language can express equality, subtree and parallelism constraints over trees. The equality and subtree constraints are used to describe semantically underspecified information (scope ambiguities). The parallelism constraints are applied to express the parallelism between the semantics of the source clause and the semantics of the target clause. A constraint is an underspecified description of a tree structure. (8) is the representation for the source clause of (2). It can be fully described by means of the constraints for the equality and subtree relations. However, the linking relation between \text{his} (the semantics of the pronoun 'his') and \text{john} (the semantics of the proper name 'John') is not expressible in CU. It follows
that the constraint language of CU must be extended and modified to deal with anaphora in VPE additionally.

A Constraint Language for Lambda-link structures (CLLS)

CLLS is an extension of the constraint language of CU. It not only deals with the interaction of VPE and quantifier scope but also with the interaction of VPE and anaphora. CLLS expresses constraints over lambda-link structures. Lambda-link structures are tree structures with additional linking relations (for anaphoric links) and binding relations (for lambda bindings). They extend and modify tree structures in the following respects:

1. Each node in the tree structures is given a unique name. This gives us a direct handle on occurrences of portions of semantic information.

2. The lambda binding of a variable is expressed directly as a binding relation between two nodes, labelled as var and lam. A binding relation expresses that an object language variable z is bound by a λ-binder of the object language like λx inside of the scope of the λ-binder. Because the binding relation is represented explicitly, there is no need to give the variable and the binder an index name as done for CU (e.g., lam_z). Thus, it avoids the problem of variable capturing, which CU and other underspecification approaches suffer from. CLLS makes possible to express lambda binding across underspecified dominance relations. Binding relations cannot be destroyed by other material in the structure.

3. The anaphoric link of a pronoun and its antecedent is modeled as a linking relation between two nodes. This means that different occurrences of anaphora can be kept separate and their dependencies can be controlled. This is crucial for our analysis.
For instance, the lambda-link structure, displayed in Figure 1.3, extends and modifies the tree in Figure 1.2 with additional relations to express lambda bindings and anaphoric links. Its nodes are named by $X_1$ to $X_{13}$. There are dominance relations between nodes that relate every node to at least one other node. For example, node $X_1$ dominates node $X_2$ and node $X_3$. The two binding relations are illustrated as dotted arrows pointing from the nodes labeled as $\text{law}$ to the nodes labeled as $\text{var}$. The linking relation is depicted by a dashed arrow pointing from node $X_7$ to node $X_2$. It expresses the linking relation between the semantic representations of the pronoun ‘his’ and its antecedent ‘John’.

Now we have gained some intuitions about a lambda-link structure. The constraint language can express constraints over lambda-link structures, such as dominance constraints, binding constraints, linking constraints and parallelism constraints. The lambda-link structure in Figure 1.3 can be described by means of constraints. Constraints can be represented graphically too. For example, the graph in Figure 1.3 can also be regarded as a constraint for the lambda-link structure which it describes.

Our analysis of the interaction between VPE and anaphora is integrated into this framework in the following way: Linking relations between the semantics of pronouns and their antecedents are expressed by linking constraints over nodes; the options for anaphoric linking as expressed in the parallelism requirement are spelled out in the resolution algorithm for solving the interaction of linking constraints and parallelism constraints. Parallelism constraints are used to describe the parallelism between semantic representations. Actually, the basic idea of dealing with the underspecified representation and resolution of VPE underlying our account stems from the approach of CU. The semantic representation of VPE is a union of the constraint sets obtained by interpreting source clause and target clause independently of each other plus the parallelism constraint. Its resolution involves solving the constraints. CLLS preserves the expressive power of CU and is defined in such a way that we can straightforwardly integrate our analysis into it. We will demonstrate the use of this framework by examples of the interaction between VPE, quantifier scope and anaphora.

1.3 CHORUS

This thesis is written within the research project “Constraint-based Higher-Order Representation for Underspecified Semantics (CHORUS)”. CHORUS is part of the Sonderforschungsbereich 3 (SFB) 378 at the University of the Saarland in Germany. The central concern of CHORUS is to develop formalisms and processing methods for computational semantics that are suitable for coping with various kinds of underspecification. One of the research topics is the development of a logical language for underspecified semantics employing meta-variables and higher-order constraints. CU belongs to the research results in the first phase in CHORUS. CLLS is the further development of CU.

The project CHORUS has developed a natural language processing system CHOLI, together with another SFB 378 project LISA (Linguistische Inference für die Semantische Auswertung). CHOLI consists of a HPSG grammar component whose syntax is close to the propos-

---

3 Special Research Division
4 Linguistic Inference for Semantic Evaluation
als presented in Pollard and Sag (1994), an interface for the determination of parallelism in discourse and a resolution component\(^5\). The grammar component covers a fragment of English grammar. The examples treated by the grammar include VPE, anaphora, quantifier scope, and their interaction. LISA and CHORUS share the syntax of the grammar. CHORUS uses the constraint language of CU as the semantic representation language for the grammar. The output of the semantic construction is a set of constraints. The set of constraints is solved by the resolution component. The resolution component is an implementation of the resolution algorithm of CU (Koller 1997). Parallel to this thesis, I have worked on the implementation of the grammar component and the interface component of CHOLI. CLLS and its resolution algorithm is still under development. They will be integrated into CHOLI in the future.

CHOLI is implemented in the programming language Oz (Saraswat et al. 1991; Smolka 1994; Smolka 1995). Oz is a higher-level programming language combining constraint inference with concurrency. It supports the implementation of the constraint-based grammar formalisms like HPSG. In particular, it is suitable for the implementation of the representation and the resolution of the constraint-based semantic formalisms such as CU and CLLS.

1.4 Overview

In Chapter 2, we discuss four approaches to the interaction of VPE and anaphora. We begin with a representative of the so-called identity-of-relations analyses, Williams (1977), followed by the ellipsis theory proposed by Dalrymple et al. (1991), which is based on higher-order unification, the proposal made by Kehler (1995) and the parallelism theory provided by Hobbs and Kehler (1997). All these approaches discuss problems that arise in the interaction of VPE and anaphora. Together, they indicate the problems in this research area and its development in last two decades.

In Chapter 3, we will discuss the underspecified treatment of ambiguities. We will describe Hole Semantics proposed by Bos (1996) as a general introduction to semantic underspecification. Hole Semantics supports underspecified semantic representations. It is regarded as a generalization of the formalism of Underspecified Discourse Representation Structures (Reyle 1993). Then we will describe the framework of context unification (Nehren et al. 1997) from which our approach stems.

In Chapter 4, we will present our analysis of the strict/sloppy ambiguity with the help of a series of problematic cases found in the literature. In addition, we will describe the framework of constraints over \(\lambda\)-link structures. We will demonstrate the appropriateness of this framework by examples of the interaction between VPE, anaphora and quantifier scopes. In addition, we will outline some extensions to our account, including stripping, gapping and extended parallelism in ellipsis.

In Chapter 5, the last chapter, we sum up the whole thesis and point out directions for future research.

\(^5\)The parser used in CHOLI is developed by the SFB 378 project NEGRA (Nebenläufige Grammatische Verarbeitung, in English, concurrent grammatical processing).
Chapter 2

Approaches to Verb Phrase Ellipsis and Anaphora

Discussions about anaphora in VPE usually begin with the observation that there is an ambiguity in the interpretation of the elided pronoun in elliptical constructions like (1):

(1) Max saw his mother, and Oscar did, too.

The stranded auxiliary in the second clause (the target clause) indicates the elision of a VP. A representation for the elided VP must be recovered from the representation of another clause, in this case, the first clause (the source clause). The two subjects ‘Max’ and ‘Oscar’ are parallel elements. The parallel elements are constituents in the source clause and their overt parallel counterparts in the target clause. Under the reading in which the pronoun ‘his’ refers to the source parallel element ‘Max’, the target clause can be understood either as Oscar having seen Max’s mother - the strict reading - or Oscar having seen Oscar’s mother - the sloppy reading.

The problem of resolving VPE is to obtain a correct interpretation of the target clause which contains an elided VP. As shown in example (1), if the source clause contains anaphoric elements like pronouns, an ambiguity arises in the target clause. A proper analysis of ellipsis must therefore explain the ambiguity in the target clause.

In the last two decades, a lot of attention has been paid to the problem of ellipsis resolution in theoretical and computational linguistics. Following Dalrymple et al. (1991) (DSP, henceforth), we divide the ellipsis analyses into two categories, identity-of-relations analyses (e.g., Fiengo and May 1994; Gawron and Peters 1990; Williams 1977; Sag 1976) and non-identity analyses (e.g., Niehren et al. 1997; Hobbs and Kehler 1997; Fox 1996; Gardent and Kohlhase 1996; Kehler 1995; Dalrymple et al. 1991). The identity-of-relations analyses require that the representation for the elided VP in the target clause should be identical with the representation for the source VP. Therefore, ambiguity of the target clause comes about because the source clause is ambiguous. The identity-of-relations analyses place the ambiguity in the source clause. Sag (1976) and Williams (1977) belong to the most prominent and influential analyses in the identity-of-relations analyses. The two approaches require that the reconstruction of the syntactic tree in the target clause must obey the constraint that the semantic
interpretations of the source VP and the target VP are identical. Sag (1976) uses deletion in the resolution process whereas Williams (1977) employs copying. Non-identity analyses do not insist on the identical representations for the two VPs in the source and the target clause. They allow that ambiguity of the target clause arises in the resolution process. In this chapter, we will concentrate on the semantic accounts among the non-identity analyses. In the semantic accounts, the target clause is resolved at a semantic level of representation. Before we discuss the different approaches, we will at first describe different phenomena of the interaction of VPE and anaphora.

2.1 VPE and Anaphora

The problem area of the interaction of VPE and anaphora is characterized by a number of examples for the strict/sloppy ambiguity in which readings are mysteriously missing and small changes lead to different intuitive judgments. These examples are usually used as benchmarks to judge whether an ellipsis analysis is general and adequate. Three of them are listed below.

**Cascaded Ellipsis**  Sentence (2), due to Dahl (1972), contains two elliptical constructions: The first clause is the source clause of the second clause whereas the second clause is the source clause of the third clause.

(2)  John\(_1\) realizes that he\(_1\) is a fool, but Bill\(_2\) does not, even though his\(_2\) wife does.

Assuming that the pronoun 'he' in the first clause refers to 'John' and the pronoun 'his' in the third clause refers to 'Bill', Dahl (1972) notes that this sentence has not only the two readings like (3-1) and (3-2) in which the two elided pronouns in the second and the third clause have a strict or a sloppy reading simultaneously, but also a third reading, such as (3-3), in which the elided pronoun in the third clause has a strict reading when the elided pronoun in the second clause has a sloppy reading.

(3)  1. John realizes that John is a fool, but
    Bill does not realize that John is a fool, even though
    Bill's wife realizes that John is a fool.
    (strict, strict)

    2. John realizes that John is a fool, but
    Bill does not realize that Bill is a fool, even though
    Bill's wife realizes that Bill's wife is a fool.
    (sloppy, sloppy)

    3. John realizes that John is a fool, but
    Bill does not realize that Bill is a fool, even though
    Bill's wife realizes that Bill is a fool.
    (sloppy, strict)
2.1 VPE and Anaphora

Although Sag (1976) claims that the third reading is not available, many other ellipsis theories (e.g., Dalrymple et al. 1991; Fiengo and May 1994; Kehler 1995) agree with Dahl (1972). We find the reading acceptable too. However, Fiengo and May (1994) argue that the third reading is only available, if the subject in the third clause contains a pronoun which is taken as referring to ‘Bill’. Following the DSP analysis, we claim that the third reading is also available in a certain context, if the subject in the third clause does not contain a pronoun referring to the subject in the second clause, like sentence (4).

(4) John realises that he is a fool, but Bill does not, even though Oscar does.

**Many-Pronouns-Puzzle** We have already presented the many-pronouns-puzzle (5) in Chapter 1. For convenience, we will not repeat its readings again.

(5) John said that he liked his mother, and Bill did too.

**Five-Reading Sentence** Sentence (6), given in Gawron and Peters (1990), is a complex case of the interaction of strict and sloppy readings.

(6) John revised his paper before the teacher did, and Bill did too.

(6) contains sentence (7), which itself contains an elliptical VP, as its source clause. (7) has a strict and a sloppy reading according to the interpretation of its elided pronoun.

(7) John revised his paper before the teacher did.

DSP claim that (6) has five readings. Various ellipsis analyses generate either too many or too few readings for it. The challenge of dealing with sentences like (6) is the question of how to reconstruct the target clause from a source clause which itself is a case of the interaction of VPE and anaphora. For (6), the overt pronoun and the elided pronoun in the source clause give rise to two parallel elided pronouns in the target clause. Let’s pay attention to the interpretation of the two elided pronouns. The five correct readings are enumerated in (8). To distinguish the readings from each other, we name these readings in accordance with the interpretations of the four pronouns in this sentence. For instance, if the four pronouns in a reading are interpreted as John, then the reading is called JJJJ.

(8) 1. John revised John’s paper before the teacher revised John’s paper, and Bill revised John’s paper before the teacher revised John’s paper. (JJJJ)
   2. John revised John’s paper before the teacher revised John’s paper, and Bill revised Bill’s paper before the teacher revised John’s paper. (JJBJ)
   3. John revised John’s paper before the teacher revised John’s paper, and Bill revised Bill’s paper before the teacher revised Bill’s paper. (JJBB)
   4. John revised John’s paper before the teacher revised the Teacher’s paper, and Bill revised John’s paper before the teacher revised the Teacher’s paper. (JTJT)
5. John revised John's paper before the teacher revised the Teacher's paper, and
Bill revised Bill's paper before the teacher revised the Teacher's paper. (JTB

JJBB in (9) is the unavailable reading.

(9) *John revised John's paper before the teacher revised John's paper, and
Bill revised John's paper before the teacher revised Bill's paper. (JJJB)

JJBJ is regarded as an unavailable reading by some analyses. DSP argue that although
JJBJ is difficult, it is not impossible. We follow DSP's analysis. Kehler (1995) and Hobbs
and Kehler (1997) accept the reading too.

Given the above examples, we start in the next section with the analysis of Williams (1977), as
an example of the identity-of-relations analyses, and then, discuss three more important non-
identity analyses in the literature, namely, the higher-order unification approach proposed by
DSP, the proposal made by Kehler (1995), and the parallelism theory provided by Hobbs and
Kehler (1997).

2.2 Williams' Analysis

In Williams' theory of VPE resolution, the meaning of the target VP is recovered by copying
the representation of the source VP to the VP position in the target clause. The representation
of the source VP is a lambda expression.

Williams uses the Derived VP Rule (DVPR) to convert VPs in surface structure (Chomsky
1973; Chomsky 1975) into properties in lambda notation. The DVPR has been proposed by
Partee (1975). It prefixes a VP with a lambda and a variable, and places a variable bound
by the lambda in the position of the logical subject of the verb of the VP. The operation of
the DVPR is illustrated below, where the surface structure (10-1) is converted into the logical
form (10-2).

(10) 1. Max$_1$ [saw his$_1$ mother]$_{VP}$
     2. Max$_1$ [\$x.x saw his$_1$ mother]$_{VP}$

The second $x$ in (10-2) is in the position of the logical subject of the verb saw. (10-2) can be
read as 'Max$_1$ has the property of having seen his$_1$ mother'.

The representation for the sloppy reading in VPE is generated by using the so-called Pronoun
Rule. This rule converts a subject-coreferent pronoun in a VP into a variable bound by the
lambda operator of that VP. Applying the Pronoun Rule to (10-2) results in the following
logical form.

(11) Max [\$x.x saw x's mother]$_{VP}$

Williams proposed to use the Pronoun Rule as an optional rule to account for the strict/sloppy
ambiguity. The source VP in example (12) has therefore two representations shown in (13).
(12) Max$_1$ saw his$_1$ mother, and Oscar did too.

(13) 1. $\lambda x. x$ saw his$_1$ mother
2. $\lambda x. x$ saw $x$’s mother

The first representation (13-1) is derived by using the DVPR only.

$$\text{Max}_1 \ [\text{saw his}_1 \ \text{mother}]_{VP}$$

**DVPR**

$$\text{Max}_1 \ [\lambda x. x \ \text{saw his}_1 \ \text{mother}]_{VP}$$

The second representation (13-2), is derived by first applying the DVPR and then the Pronoun Rule to the source VP:

$$\text{Max}_1 \ [\text{saw his}_1 \ \text{mother}]_{VP}$$

**DVPR**

$$\text{Max}_1 \ [\lambda x. x \ \text{saw his}_1 \ \text{mother}]_{VP}$$

**Pronoun Rule**

$$\text{Max}_1 \ [\lambda x. x \ \text{saw } x’\text{’s mother}]_{VP}$$

The above two representations of the source VP determine the readings for the target clause: Copying (13-1) to the target VP position yields the strict reading (14-1), and copying (13-2) gives rise to the sloppy reading (14-2).

(14) 1. Oscar$_2$ $[\lambda x. x \ \text{saw his}_1 \ \text{mother}]_{VP}$
2. Oscar$_2$ $[\lambda x. x \ \text{saw } x’\text{’s mother}]_{VP}$

Therefore, the key point of Williams’ analysis of the strict/sloppy ambiguity is the ambiguous interpretation of the pronouns in the source VP. Williams’ analysis has a very simple and attractive resolution strategy. However, the ambiguous interpretation of pronouns in the source VP seems unmotivated if one considers the source clause in isolation, since the various derived properties of the source VP result in the same meaning for the whole source clause. Furthermore, it fails to make correct predictions in more complicated cases. Especially, the cascaded ellipsis (2), repeated below as example (15), is problematic for this analysis.

(15) John$_1$ realizes that he$_1$ is a fool, but Bill$_2$ does not, even though his$_2$ wife does.

Williams’ analysis predicts only those two readings in which the second and the third clause have the same (strict or sloppy) type of reading. The reading, given in (16), where the second clause has a sloppy reading while the third clause has a strict, is undervisible.
(16) John realizes that *John is a fool, but
      Bill does not realize that *Bill is a fool, even though
      Bill’s wife realizes that *Bill is a fool.
      (sloppy, strict)

If the second clause has a sloppy reading, the logical form of the source VP (the VP in the
first clause), must be represented as follows:

(17) \( \lambda x. x \) realizes \( x \) is a fool

That is, the pronoun ‘he’ must be represented as a lambda-bound variable. After copying (17)
to the VP in the second clause, the VP representation of the second clause is therefore (17)
as well, which leads to a sloppy reading for the third clause – there is no way to get a strict
reading for the third clause. Thus, Williams’ identity-of-relations analysis undergenerates in
the case of the cascaded ellipsis.

Another problem of Williams’ analysis is the optional use of the Pronoun Rule which plays
an important role for the reconstruction of the strict and sloppy readings. It leads to over-
geneneration for sentences like the many-pronouns-puzzle (5), repeated here as (18). Williams’
analysis predicts four readings (recall that this sentence has only three readings).

(18) \( \text{John}_1 \) said that \( \text{he}_1 \) liked \( \text{his}_1 \) mother, and \( \text{Bill}_2 \) did too.

In the source VP, there are two pronouns referring to the subject. Each of them can be inter-
preted as either a lambda bound variable or a referent. Therefore, there are four possibilities
of the source VP representation. The fourth is the unwanted one:

1. \( \lambda x. x \) said that \( \text{he}_1 \) liked \( \text{his}_1 \) mother
2. \( \lambda x. x \) said that \( x \) liked \( \text{his}_1 \) mother
3. \( \lambda x. x \) said that \( x \) liked \( x \)’s mother
4. \( \ast \lambda x. x \) said that \( \text{he}_1 \) liked \( x \)’s mother

In summary, the key point of Williams’ analysis of the strict/sloppy ambiguity is the am-
biguous interpretation of the pronouns in the source VP. Williams’ analysis has a very simple
and attractive resolution strategy. However, the ambiguous interpretation of pronouns in the
source VP seems unmotivated if one considers the source clause in isolation, since the various
derived properties of the source VP result in the same meaning for the whole source clause.
Moreover, Williams’ analysis fails to give correct predictions of some complicated cases of the
strict/sloppy ambiguity. The idea of copying under identity eliminates an accepted reading
in the cascaded ellipsis. The optional use of the Pronoun Rule results in overgeneration in
the case of the many-pronoun-puzzle.
2.3 The Analysis of Dalrymple, Shieber, and Pereira

DSP offer a parallelism theory for ellipsis resolution that does not require an ambiguous interpretation of source pronouns. The basic idea underlying their theory of ellipsis resolution is to recover a property common to the source clause and the target clause by using Higher-Order Unification (HOU). The level of representation at which HOU operates is a type theoretic representation after beta reduction.

More precisely, the theory of DSP is based on the fact that the source clause and the target clause share the same representations modulo the semantic representations of the parallel elements. Let the common semantic representation of the source clause and the target clause be represented by a second-order free variable $P$, the semantic representation of the source clause be represented by $s$ and the semantic representations of the source parallel elements be represented by $s_1, \ldots, s_n$, $P$ can be derived by solving the equation like (19).

\[(19) \ P(s_1, s_2, \ldots, s_n) = s\]

DSP utilize Huet’s higher-order unification algorithm (Huet 1975) to solve equations like (19). The aim of the algorithm is to find a substitution of terms for free variables which makes an equation true. Applying $P$ to the semantic representation of target parallel elements $t_1, \ldots, t_n$, results in the semantic representation of the target clause, $t$ in (20).

\[(20) \ P(t_1, t_2, \ldots, t_n) = t\]

We go through DSP’s analysis of example (21) step by step.

(21) Max saw his mother, and Oscar did, too.

In (21), the parallel elements are the two subjects ‘Max’ and ‘Oscar’. The semantic representation of the source clause is shown in (22), in which the semantic representation of the source parallel element ‘Max’ is underlined.

\[(22) \ saw(max, mother_of(max))\]

In DSP’s analysis, the semantic representations of source parallel elements in the source representation are called primary occurrences. They are underlined in the source representation in DSP’s analysis. The semantic representations of pronouns referring to source parallel elements are called secondary occurrence. The secondary occurrences are not underlined in the source representation. In (22), the nonunderlined term $max$ is the secondary occurrence. Given the source representation and the representation of the source parallel element, the equation to be solved is (23).

\[(23) \ P(max) = saw(max, mother_of(max))\]

Actually, there are four solutions for $P$ in equation (23) after the application of Huet’s algorithm.
(24) 1. \( \lambda x. \text{saw}(\text{max}, \text{mother}_o f(\text{max})) \)
    
   2. \( \lambda x. \text{saw}(\text{max}, \text{mother}_o f(x)) \)

   3. \( \lambda x. \text{saw}(x, \text{mother}_o f(\text{max})) \)

   4. \( \lambda x. \text{saw}(x, \text{mother}_o f(x)) \)

In DSP’s analysis, each solution for \( P \) must obey an additional constraint, namely the primary occurrence restriction, which requires that an admissible solution for \( P \) should not contain any primary occurrences.

According to the primary occurrence restriction, the first two solutions in (24) are not admissible. Therefore, the accepted ones are (24-3) and (24-4) where the primary occurrence is abstracted over. Note that the ambiguity of the property \( P \) is caused by the optional abstraction over the secondary occurrence. To generate a reading for the target clause, either one of the solutions of \( P \) may be applied to the semantic representation of the target parallel element, \( \text{o} \text{scar} \). Applying (24-3) to \( \text{o} \text{scar} \) yields the strict reading in (25-1); applying (24-4) results in the sloppy reading given in (25-2).

(25) 1. \( \text{saw}(\text{o} \text{scar}, \text{mother}_o f(\text{max})) \)
    
   2. \( \text{saw}(\text{o} \text{scar}, \text{mother}_o f(\text{o} \text{scar})) \)

In DSP’s analysis, there is no requirement of assuming an ambiguity of the pronouns in the source VP. Instead, the ambiguity of strict and sloppy readings arises in the resolution process, i.e., the optional abstraction over the secondary occurrence.

DSP’s analysis proves its advantages in handling the cascaded ellipsis. As discussed earlier, the reading of the cascaded ellipsis (2) (= (26)), namely (27), is undervariable in Williams’ analysis. In (27), the second clause has a sloppy reading whereas the third clause has a strict reading. DSP do not have any difficulties in generating this reading.

(26) John\(_1\) realizes that he\(_1\) is a fool, but Bill\(_2\) does not, even though his\(_2\) \text{wife} does.

(27) John realizes that \text{John} is a fool, but
    
    Bill does not realize that \text{Bill} is a fool, even though
    
    Bill’s \text{wife} realizes that \text{Bill} is a fool.
    
    (sloppy, strict)

According to DSP, the representation of the first clause in (26) is represented in (28).

(28) \( \text{realize}(\underline{\text{john}}, \text{fool}(\underline{\text{john}})) \)

To recover the representation of the second clause, we need to solve the following equation:

(29) \( P(\text{john}) = \text{realize}(\underline{\text{john}}, \text{fool}(\underline{\text{john}})) \)

One of the solutions for \( P \) in equation (29) is (30). It leads to the sloppy reading of the second clause.
(30) \( \lambda x. \text{realize}(x, \text{fool}(x)) \)

Applying this solution to \textit{bill} which is the semantic representation of the target parallel element ‘Bill’ plus a negation, we get the representation corresponding to the sloppy reading of the second clause (after beta reduction).

(31) \( \neg\text{realize}(\text{bill}, \text{fool(bill)}) \)

The parallel elements of the second and the third clause are ‘Bill’ and ‘his wife’, where \( \text{realize}(\text{bill}, \text{fool(bill)}) \) is the source representation. The equation to be solved is given in (32).

(32) \( P(\text{bill}) = \text{realize}(\underline{\text{bill}}, \underline{\text{fool(bill)}}) \)

The solution (33) for \( P \), in which the secondary occurrence is not abstracted over, yields the strict reading of the third clause.

(33) \( \lambda x. \text{realize}(x, \text{fool(bill)}) \)

The resolution of the cascaded ellipsis demonstrates the characteristic of DSP’s analysis: the ambiguity of the target clause arises in the resolution process of HOU.

Nonetheless, DSP’s analysis is known to overgenerate: while anaphora cause the strict/sloppy ambiguity, proper names do not. However, proper names cannot be distinguished from pronouns that refer to them in semantic representations in DSP’s analysis. Therefore, if the source clause contains proper names which have the same semantic representations as the source parallel elements, DSP’s analysis derives sloppy readings too. For example, there are two occurrences of the proper name ‘Felix’ in sentence (34). One of them is the source parallel element.

(34) Felix’s mother thinks Felix is a genius, and Siegfried’s mother does too.

This example has only one reading, such as

(35) Felix’s mother thinks \textit{Felix} is a genius, and Siegfried’s mother thinks \textit{Felix} is a genius too.

But DSP’s approach predicts two readings. We assume that the parallel elements are ‘Felix’ and ‘Siegfried’. The source representation is (36).

(36) \( \text{think}(\underline{\text{mother_of(felix)}}, \underline{\text{genius(felix)}}) \)

The equation to be solved is given as follows.

(37) \( P(\text{felix}) = \text{think}(\underline{\text{mother_of(felix)}}, \underline{\text{genius(felix)}}) \)
The second occurrence of *felix* should not be abstracted over, since it is not the semantic representation of a pronoun. However, DSP make no difference between a representation of a pronoun and a representation of a proper name. Therefore, there are two admissible solutions for *P* in the above equation: (38-1) and the unwanted reading (38-2).

(38) 1. \( \lambda x. \text{think}(\text{mother} \_of(x), \text{genius}(\text{felix})) \)

2. \( \lambda x. \text{think}(\text{mother} \_of(x), \text{genius}(x)) \)

Another case of overgeneration of DSP's analysis is caused by the optional abstraction over secondary occurrences. This method is similar to the optional use of the pronoun rule in Williams' analysis. As mentioned above, Williams' analysis overgenerates in the case of the many-pronouns-puzzle. DSP has the same overgeneration problem: They predict four readings instead of three for the many-pronouns-puzzle (39) (= (5)).

(39) John\(_1\) said that he\(_1\) liked his\(_1\) mother, and Bill\(_2\) did too.

The source representation is (40).

(40) \( \text{say}(\underline{\text{john}}, \text{like}(\text{john}, \text{mother} \_of(\text{john}))) \)

There are four possible solution for *P* in (41). They are given in (42), where the fourth is the unwanted one.

(41) \( P(\text{john}) = \text{say}(\underline{\text{john}}, \text{like}(\text{john}, \text{mother} \_of(\text{john}))) \)

(42) 1. \( \lambda x. \text{say}(x, \text{like}(\text{john}, \text{mother} \_of(\text{john}))) \)

2. \( \lambda x. \text{say}(x, \text{like}(x, \text{mother} \_of(\text{john}))) \)

3. \( \lambda x. \text{say}(x, \text{like}(x, \text{mother} \_of(x))) \)

4. \( *\lambda x. \text{say}(x, \text{like}(\text{john}, \text{mother} \_of(x))) \)

To sum up, DSP's analysis provides a simple computational method for ellipsis interpretation. Compared with the identity-of-relations analyses, it does not require an ambiguity of pronouns in the source clause to explain the strict/sloppy ambiguity in the target clause. The ambiguity arises in the resolution process. While the analysis treats the cascaded ellipsis in an elegant way, it overgenerates: First, it uses a type theoretic representation in which the proper names and pronouns referring to them cannot be distinguished from each other in the semantic representations, this causes overgeneration in cases like (34). The problem indicates that the level of semantic representation at which an ellipsis resolution process operates must be able to distinguish the semantics of pronouns from the semantics of full NPs. Second, the optional abstraction over secondary occurrences cannot, by itself, cope with the many-pronouns-puzzle. Therefore, a more fine-grained analysis is needed to explain the strict/sloppy ambiguity in the case of the many-pronouns-puzzle.
2.4 Kehler’s Analysis

Kehler (1995) provides a general theory to explain the strict/sloppy ambiguity. It gives correct predictions for examples which are problematic for the above approaches, such as the many-pronouns-puzzle. In addition, it generates five correct readings for the five-reading sentence (6). As noted by DSP, most of the ellipsis analyses generate either too many or too few readings for (6). The accounts of Williams (1977) and Sag (1976) can be shown to derive only two readings. The analysis of Gawron and Peters (1990) predicts three readings. DSP themselves derive six readings\(^1\).

Kehler claims that there is a certain referential dependence, termed as \textit{linking relation}, between the semantic representations of pronouns and their antecedents. The linking relations in the source clause determine the linking relations in the target clause, and the interpretation of the elided pronouns depends on the interpretation of their antecedents. The linking relation in the semantic representations corresponds to the binding relation in syntactic representations (Chomsky 1981). According to Chomsky (1981), \textit{A} is defined to \textit{bind} \textit{B} in the syntactic representation iff \textit{A} and \textit{B} are co-indexed, and \textit{A} c-commands \textit{B}. \textit{A} c-commands \textit{B} iff \textit{A} does not dominate \textit{B}, and \textit{B} does not dominate \textit{A}, and the first branching node dominating \textit{A} also dominates \textit{B}\(^2\). The correspondence is defined in the so-called \textit{role linking rule}:

A referential element is linked to the most immediate coreferential element that c-commands it in the syntax. If there are no c-commanding elements, then it is linked to the most salient discourse referent. (Kehler 1995, page 68)

That is, the linking relation between the term representing a pronoun and the term representing its antecedent holds if the antecedent binds this pronoun most immediately (i.e. the most immediate coreferential element c-commands the pronoun). If there is no binding possible, the linking relation can still be determined by discourse information.

Kehler represents all entities and relations as unique terms and predicates in the logical representation. All unique terms have an index (identifying them as such), although these will only be displayed when necessary in order to improve readability. Given such a logical language, the linking relation between two terms in the semantic representation can be expressed explicitly. A term representing a pronoun has a \textit{link} associated with it which establishes the (one-way) relation between this pronoun and its antecedent. The link property is associated with the index of the term to which it refers. For instance, the semantic representation for the source clause of sentence (43) is shown in representation (44).

(43) Max\(_{1}\) saw his\(_{1}\) mother, Oscar did too.

(44) \textit{saw}(\textit{max}_{s0}, \textit{mother}_of(\textit{him}_{s1\rightarrow s0}))

---

\(^1\)DSP argue that the unwanted reading can be eliminated by using an additional constraint such as the anaphoric dependence proposed in Kehler’s analysis.

\(^2\)Syntactic representations are usually trees. Constituents are nodes in a syntactic tree.
The term him\textsubscript{1}\rightarrow\textsubscript{s0} identifies a term him having the index \textsubscript{s1}, which is linked to the term representing the antecedent of the pronoun having the index \textsubscript{s0}, in this case max\textsubscript{s0}. This linking relation corresponds to the binding relation between the subject ‘Max’ and the pronoun ‘his’. (44) corresponds to the reading Max saw Max’s mother.

Kehler uses a copying algorithm to resolve the semantic representation for the target clause. We are interested in how the strict/sloppy ambiguity is derived. Kehler uses operations referring and copying to derive such an ambiguity. In the case of referring, the link is established between the term and its parallel term in the source. For example, we assume that (44) is the source representation. The target representation (45-1) is derived according to referring. The link \textsubscript{t1}\rightarrow\textsubscript{s1} is established between the target term him\textsubscript{t1} and its parallel term him\textsubscript{s1} in (44). In the case of copying, a link in the target clause is created between two target terms. The link is only admissible if there is a link between their parallel source terms in the source representation. For instance, in (45-2) the link \textsubscript{t1}\rightarrow\textsubscript{s0} is created between the term him\textsubscript{t1} and the term oscar\textsubscript{s0}, and this link is parallel to the link \textsubscript{s1}\rightarrow\textsubscript{s0} in (44). The term him\textsubscript{s1} and max\textsubscript{s0} are the parallel terms of him\textsubscript{t1} and oscar\textsubscript{t0}, respectively. The representation (45-1) corresponds to the strict reading of the target clause of (43), i.e., Oscar saw Max’s mother. (45-2) represents the sloppy reading, namely Oscar saw Oscar’s mother.

\begin{equation}
(45) \begin{align*}
1. & \quad \text{saw(oscar}_{t0}, \text{mother}_{of}(\text{him}_{t1\rightarrow s1})) \\
2. & \quad \text{saw(oscar}_{t0}, \text{mother}_{of}(\text{him}_{t1\rightarrow t0}))
\end{align*}
\end{equation}

In Kehler’s analysis, a linking relation in the source clause causes two options of a new linking relation in the target clause. Kehler makes use of either referring or copying to reconstruct linking relations in the target representation. The new linking relations establish the strict and sloppy readings for the target clause. Although Kehler’s analysis is simple, it gives correct predictions about the many-pronouns-puzzle and the five-reading sentence.

Consider the source clause of the many-pronouns-puzzle, given in (46). It contains two most immediate bindings in its syntactic representation: The subject ‘John’ binds the first pronoun ‘he’ most immediately while the first pronoun ‘he’ binds the second pronoun ‘his’ most immediately. According to the role linking rule, the two immediate bindings in the syntax give rise to two corresponding links in the semantic representation: the representation of the first pronoun is linked to the representation of the subject and the representation of the second pronoun is linked to the representation of the first pronoun. Hence, the semantic representation of the source clause is (47), containing two links: \textsubscript{s1}\rightarrow\textsubscript{s0} and \textsubscript{s2}\rightarrow\textsubscript{s1}.

\begin{equation}
(46) \quad \text{John}_{t1} \text{id\ he}_{t1} \text{ liked his}_{t1} \text{ mother}, \text{ and Bill did too.}
\end{equation}

\begin{equation}
(47) \quad \text{say(john}_{s0}, \text{like(he}_{s1\rightarrow s0}, \text{mother}_{of}(\text{him}_{s2\rightarrow s1}))}
\end{equation}

Given (47) as the source representation, the option of applying either referring or copying to the pronominal terms in the target leads to four representations for the target clause, shown in (48). The term he\textsubscript{t1} represents the semantics of the first elided pronoun, while him\textsubscript{t2} stands for the second one.
2.4 Kehler’s Analysis

(48) 1. say(bill₁₀, like(he₁₁→₁₂, mother₉₀(him₂₂→₁₃)))
      (refer, refer)

2. say(bill₁₀, like(he₁₁→₁₂, mother₉₀(him₂₂→₁₃)))
      (refer, copy)

3. say(bill₁₀, like(he₁₁→₁₀, mother₉₀(him₂₂→₁₃)))
      (copy, refer)

4. say(bill₁₀, like(he₁₁→₁₀, mother₉₀(him₂₂→₁₃)))
      (copy, copy)

Unlike the analyses of Williams and DSP, in which four distinct representations correspond to four different readings, in Kehler’s approach the four representations lead to three readings. The reason is that (48-1) and (48-2) represent the same reading, namely, the strict-strict reading given in (49). In (48-1), he₁₁ is linked to he₁₂ to receive a strict reading, while him₂₂ is linked to him₁₂, resulting in the strict reading too. Therefore, (48-1) represents the strict-strict reading. In (48-2), he₁₁ is linked to he₁₂₁, yielding the strict reading. The term him₂₂ is linked to he₁ which is interpreted as strict. Therefore, him₁₂ in (48-2) has a strict reading too. Hence it follows that (48-2) also represents the strict-strict reading.

(49) Bill said John liked John’s mother.
      (strict, strict)

In (48-3), he₁₁ is linked to bill₁₀ to have a sloppy reading, and the link assigned to him₁₂ provides a strict reading. Clearly, (48-3) represents the sloppy-strict reading:

(50) Bill said Bill liked John’s mother.
      (sloppy, strict)

(48-4) stands for the sloppy-sloppy reading. The term he₁₁ is linked to bill₁₀ to receive a sloppy reading. him₁₂ is linked to he₁₁ and is therefore interpreted as sloppy too.

(51) Bill said Bill liked Bill’s mother.
      (sloppy, sloppy)

Altogether, the four representations result in three wanted readings: strict-strict, sloppy-strict and sloppy-sloppy. The representation of the unwanted strict-sloppy reading would be (52). It is not derivable in this analysis since there is no available parallel link for₁₂→₁₀ in the source representation.

(52) say(bill₁₀, like(he₁₁→₁₂, mother₉₀(him₂₂→₁₃)))

In the following, we will show that Kehler’s analysis generates the five correct readings of the five-reading sentence (53), which are repeated in (54).

(53) John₁ revised his₁ paper before the teacher₂ did, and Bill₃ did too.
1. John revised John's paper before the teacher revised John's paper, and Bill revised John's paper before the teacher revised John's paper. (JJJJ)

2. John revised John's paper before the teacher revised John's paper, and Bill revised Bill's paper before the teacher revised John's paper. (JJBJ)

3. John revised John's paper before the teacher revised John's paper, and Bill revised Bill's paper before the teacher revised Bill's paper. (JJBB)

4. John revised John's paper before the teacher revised the Teacher's paper, and Bill revised John's paper before the teacher revised the Teacher's paper. (JTJT)

5. John revised John's paper before the teacher revised the Teacher's paper, and Bill revised Bill's paper before the teacher revised the Teacher's paper. (JTBT)

The sixth reading (55) derivable by DSP is unavailable.

(55) *John revised John's paper before the teacher revised John's paper, and Bill revised John's paper before the teacher revised Bill's paper.

(JJJJ)

We go through the derivation of the five readings, showing that a representation for the reading JJBJ is not available.

The source clause has a strict and a sloppy reading. The following two representations correspond to the two readings. In (56-1), him_{b1} is linked to him_{a1} (by referring), yielding the strict reading. In (56-2), him_{b1\to b0} is received by using copying, resulting in the sloppy reading.

(56) 1. before(revise(john_{a0}, paper_{of(him_{a1\to a0})}),
       revise(teacher_{b0}, paper_{of(him_{b1\to a1})}))

2. before(revise(john_{a0}, paper_{of(him_{a1\to a0})}),
       revise(teacher_{b0}, paper_{of(him_{b1\to b0})}))

We now consider the representations for the target clause. First, we take the representation for the strict reading (56-1) as the source representation. Since there are two pronominal terms him_{a1} and him_{b1} in the source representation, there are four options of links of the target terms him_{c1} and him_{d1}, enumerated in (57).

(57) 1. before(revise(bill_{c0}, paper_{of(him_{c1\to a1})}),
       revise(teacher_{a0}, paper_{of(him_{d1\to b1})}))
       (refer, refer)

2. before(revise(bill_{c0}, paper_{of(him_{c1\to a1})}),
       revise(teacher_{a0}, paper_{of(him_{d1\to c1})}))
       (refer, copy)

3. before(revise(bill_{c0}, paper_{of(him_{c1\to c0})}),
       revise(teacher_{a0}, paper_{of(him_{d1\to b1})}))
       (copy, refer)
2.4 Kehler's Analysis

4. before(revise(bill₀, paper_of(him₁→e₀)),
   revise(teacher₀, paper_of(him₂→c₁)))
   (copy, copy)

Note that in the case of the strict reading the two pronouns in the source refer to John. First, let us look at (57-1). The term him₁→a₁ and him₂→b₁ say that the elided pronouns are linked to their parallel pronouns, respectively. Therefore, the two elided pronouns have the same reference as their parallel pronouns, namely John. (57-1) represents the reading JJJJ.

In (57-2), him₁→a₁ means that the first elided pronoun refers to its parallel pronoun. Therefore, it refers to John. Another nominal term him₂→b₁ says that the second elided pronoun has the same reference as the first elided pronoun, namely John as well. Clearly, (57-2) represents JJJJ too.

In (57-3), him₁→c₀ says that the first elided pronoun refers to Bill. The term him₂→b₁ let the second elided pronoun have the same reading as its parallel pronoun, namely John. The representation results in the reading JJB.

(57-4) is the representation for the third reading, namely JBBB. The term him₂→c₁ means that the second elided pronoun receives the same reading as the first elided pronoun him₁→e₀ which refers to Bill.

Kehler's analysis reconstructs three readings JJJJ, JBJB and JBJB in the case of the strict reading of the source clause. We now consider the case in which the source clause has a sloppy reading, given in (56-2). The target clause is also assigned four possible representations, shown in (58).

(58) 1. before(revise(bill₀, paper_of(him₁→a₁)),
   revise(teacher₀, paper_of(him₂→b₁)))
   (refer, refer)

2. before(revise(bill₀, paper_of(him₁→a₁)),
   revise(teacher₀, paper_of(him₂→a₀)))
   (refer, copy)

3. before(revise(bill₀, paper_of(him₁→e₀)),
   revise(teacher₀, paper_of(him₂→b₁)))
   (copy, refer)

4. before(revise(bill₀, paper_of(him₁→e₀)),
   revise(teacher₀, paper_of(him₂→a₀)))
   (copy, copy)

Let us check how many readings can be derived in the sloppy reading case. Note that in the case of the sloppy reading the two pronouns in the source are interpreted as John and the teacher, respectively. First, let us consider (58-1). The term him₁→a₁ and him₂→b₁ say that the elided pronouns are linked to their parallel pronouns in the source, respectively. Therefore, the two elided pronouns have the same readings as their parallel pronouns, namely John and the teacher, respectively. Hence, (58-1) represents the reading JTJT.
In (58-2), \( \text{him}_{c1 \rightarrow d1} \) means that the first elided pronoun has the same reading as its parallel pronoun, namely, \( \text{John} \). The second pronominal term \( \text{him}_{d1} \) is linked to \( \text{teacher}_{d0} \). Therefore, the second pronoun refers to the \( \text{teacher} \). (58-2) is the representation for \( \text{JTJT} \) too.

In (58-3), \( \text{him}_{c1 \rightarrow c0} \) says that the first elided pronoun refers to \( \text{Bill} \). The term \( \text{him}_{d1 \rightarrow b1} \) means that the second elided pronoun has the same reading as its parallel pronoun, namely, \( \text{the teacher} \). (58-3) represents the reading \( \text{JTBT} \).

The reading \( \text{JTJT} \) is represented by (58-4) too. The first elided pronoun refers to \( \text{Bill} \) according to the link in \( \text{him}_{c1 \rightarrow c0} \). The term \( \text{him}_{d1 \rightarrow d0} \) says the second elided pronoun refers to \( \text{the teacher} \).

The two readings \( \text{JTJT} \) and \( \text{JTBT} \), along with the three readings derived earlier for the strict case, are the five readings shown in (54). The inadmissible reading given in (55) is not derivable by Kehler's analysis.

Fox (1996) provides an approach similar to Kehler's analysis. Fox's analysis accounts for the many-pronouns-puzzle. But it fails to deal with the five-reading sentence. The reason is that the resolution process in Fox (1996) operates at the syntactic level\(^3\). In Fox (1996), the many-pronouns-puzzle is explained by the interaction between immediate bindings in the syntactic representation and a so-called theorem of Parallelism. Immediate bindings are obtained according to the approach of Heim (1982). The theorem of Parallelism requires that the NPs in the elided and antecedent VP must either (a) have the same referential value (referential parallelism) or (b) have parallel immediate bindings (structural parallelism). If we ignore the difference between syntactic binding relations and linking relations defined by Kehler, Fox's referential parallelism and structural parallelism are similar to Kehler's referring and copying, respectively. In the case of the many-pronouns-puzzle, the source representation according to Fox (1996) is (59). We represent the bindings as arrows. The two immediate bindings correspond to the two linking relations in Kehler's analysis. The interaction of the two immediate bindings and the theorem of Parallelism results in the three readings for the many-pronouns-puzzle. They can be derived in a similar way as Kehler has done for the many-pronouns-puzzle.

\begin{equation}
\text{(59)} \quad \text{John}_1 \text{ said he}_1 \text{ liked his}_1 \text{ mother}
\end{equation}

According to the theorem of Parallelism, the representation for the strict reading of the source clause of the five-reading sentence is (60). No new binding relation between the two parallel pronouns is produced, since the source pronoun does not c-command the reconstructed pronoun. (60) causes two readings for the target clause only. As discussed above, in the case of the strict reading for the source clause, the target clause has three readings. It indicates that Fox's analysis undergenerates. Otherwise, if we treat the \( \text{before} \) clause as a VP modification, then 'John' binds the reconstructed pronoun, since 'John' c-commands 'his' in this case. This means that the source representation is (61). In (61), the second pronoun is not bound by the first pronoun, and the sloppy reading of the second elided pronoun is therefore independent on the sloppy reading of the first elided pronoun. Applying the theorem of Parallelism to (61)

\[^3\text{Many approaches have discussed problems of syntactic accounts for ellipsis resolution (e.g., Dalrymple et al. 1991; Kehler 1995; Gardent et al. 1997). In this work, we focus on semantic accounts.}\]
leads to four readings for the target clauses. One reading, namely JJJB, is unwanted. Thus, either way, Fox's analysis does not cover the five-reading sentence.

\[(60) \quad \text{John}_1 \text{ revise his}_1 \text{ paper before the teacher}_2 \text{ revise his}_1 \text{ paper.} \]

\[(61) \quad \text{John}_1 \text{ revise his}_1 \text{ paper before the teacher}_2 \text{ revise his}_1 \text{ paper.} \]

In contrast to Fox, the referring operation in Kehler's analysis does derive a new linking relation between the semantic representation of the source pronoun and the semantic representation of the reconstructed pronoun, see (56-1). The new linking relation is independent on the syntactic information. In addition, Kehler's analysis is more fine grained than other analyses discussed above. First, it regards the referential dependences between the pronominal anaphora and their antecedents. Second, the elided pronouns are not simply interpreted as strict or sloppy. Their interpretations follow from the reconstructed linking relations. The reconstructed linking relations are not constrained by syntactic information. However, Kehler's ellipsis resolution approach is purely procedural, and parallelism between the source clause and the target clause is not expressed in a declarative way.

### 2.5 The Parallelism Theory of Hobbs and Kehler

Hobbs and Kehler (1997) (H&K) present a general account of parallelism in discourse and apply it to the resolution of VPE and anaphora. Within H&K's analysis, ellipsis resolution is a special case of proving the parallelism between two fragments of discourse. The approach of H&K can be regarded as an elaboration of Kehler (1995) (Kehler) discussed above. Therefore, it has the same coverage as far as the interaction of VPE and anaphora ascertained. In this section, we will apply it to the five-reading sentence.

H&K use a special concept of logical form proposed by Hobbs (1985) as semantic representation format. The representation is a variant of Davidsonian-style representation (Davidson 1967). A crucial piece of their treatment of VPE is the explicit representation of coreference relations, denoted by the predicate \textit{Coref}. Terms in a coreference relation can be distinguished from each other by the events under which they are described. For example, \textit{Coref}(x, e_1, y, e_2) says that x under the description associated with the event e_1 is coreferential with y under the description associated with the event e_2. From this, we can infer \(x = y\) but not \(e_1 = e_2\). This is important for them to control the referential dependence between two terms.

Within the logical concept used by H&K, a property is defined to consist of a predicate applied to a number of arguments, and the interpretation of a clause is represented as a combination of a property having a number of arguments and the properties of the arguments. For instance, the semantic representation of the source clause of (62) is given in (63).

\[(62) \quad \text{John}_1 \text{ revised his}_1 \text{ paper before the teacher did.} \]
(63) `\text{revise}'(e_{12}, j, p_1)\\ j : \text{John}'(e_{11}, j)\\ p_1 : \text{paper}'(e_{15}, p_1)\\ \text{Poss}'(e_{14}, x_1, p_1)\\ x_1 : \text{he}'(e_{13}, x_1)\\ \text{Coref}(x_1, e_{13}, j, e_{11})\\

The property `\text{revise}'(e_{12}, j, p_1) means that the event $e_{12}$ is a revising event by $j$ of $p_1$. The two arguments $j$ and $p_1$ are associated with some other properties. For instance, the argument $j$ has a property $\text{John}'(e_{11}, j)$ which is the $e_{11}$ of $j$'s being named $\text{John}'$. The linking relation between the representation of the pronoun ‘he’ and the representation of its antecedent ‘John’ is expressed by a coreference relation: the $\text{Coref}$ relation links $x_1$, the variable corresponding to ‘he’ (eventuality $e_{13}$), to its antecedent $j$, the entity described by ‘John’ (eventuality $e_{11}$).

For H&K, ellipsis resolution is done by proving the parallelism between the semantic representations of the source clause and the target clause. Formally, parallelism is characterized by the similarity between two properties. Two properties are similar if two corresponding properties can be inferred from them in which the predicates are the same and the corresponding pairs of arguments are either coreferential or similar, as illustrated in (64).

\begin{align*}
(64) \quad \text{Similar}[p_1'(e_1, x_1, \ldots, z_1), p_2'(e_2, x_2, \ldots, z_2)] :\\ & \quad p_1'(e_1, x_1, \ldots, z_1) \text{ infers } p'(e_1, x_1, \ldots, z_1) \text{ and }\\ & \quad p_2'(e_2, x_2, \ldots, z_2) \text{ infers } p'(e_2, x_2, \ldots, z_2), \text{ where }\\ & \quad \text{Coref}(x_1, \ldots, x_2, \ldots) \text{ or Similar}[x_1, x_2],\\ & \quad \ldots\\ & \quad \text{Coref}(z_1, \ldots, z_2, \ldots) \text{ or Similar}[z_1, z_2]\\
\end{align*}

(65) \quad \text{Similar}[x_1, x_2] :\\ & \quad \text{Similar}[p_1''(\ldots, x_1, \ldots), p_2''(\ldots, x_2, \ldots)],\\ & \quad \ldots\\ & \quad \text{Similar}[q_1''(\ldots, x_1, \ldots), q_2''(\ldots, x_2, \ldots)]

\text{Similar}[x_1, x_2] holds if other inferentially independent properties of $x_1$ and $x_2$ are similar, as illustrated in (65). Two properties are inferentially independent if neither can be derived from the other. For example, given a knowledge base $K$ representing mutual knowledge of the participants in the discourse, properties $p_1''(\ldots, x_1, \ldots)$ and $q_2''(\ldots, x_1, \ldots)$ are inferentially independent, if neither $K$, $p_1''(\ldots, x_1, \ldots) \vdash q_2''(\ldots, x_1, \ldots)$ nor $K$, $q_1''(\ldots, x_1, \ldots) \vdash p_1''(\ldots, x_1, \ldots)$.

In fact, similarity is a matter of degree. H&K claim that the more corresponding pairs of inferentially independent properties are found, the stronger the similarity. Interpretations should seek to maximize similarity. The ideal case for proving similarity of two properties is that their predicates are the same and their arguments can be coreferential. To explain the strict/sloppy ambiguity in VPE, H&K make use of the option that two entities can be either coreferential or similar.
We illustrate this approach with (62), a simple case of VPE. Take (63) as the source representation. We know there is an elided eventuality \( e_{22} \) of unknown type \( P \) in the target clause, of which the logical subject is the teacher \( t \).

\[
(66) \quad P(e_{22}, t) \\
t : \text{teacher}'(e_{21}, t)
\]

Since ellipsis presupposes parallelism, \( e_{22} \) must stand in a parallel relation to the source clause eventuality, namely John’s revising his paper \( (e_{21}) \). To establish \( \text{Similar}(e_{12}, e_{22}) \), we need to show that their corresponding arguments are similar: \( j \) and \( t \) are similar by virtue of being persons. The corresponding objects \( p_1 \) and \( p_2 \) are similar if we take \( p_2 \) to be a paper and to have a \( \text{Poss} \) property similar to that of \( p_1 \). The latter is true if corresponding to the possessor \( x_1 \), there is an \( x_2 \) that is similar to \( x_1 \). Note that \( x_1 \) corresponds to the pronoun ‘he’ in the source clause, and the variable \( x_2 \) corresponds to the elided pronoun in the target clause. Therefore, the construction of the similarity between \( x_2 \) and \( x_1 \) is the reconstruction of the semantic representation of the elided pronoun. \( x_1 \) and \( x_2 \) can be proved to be either coreferential (case (a) in (67)) or similar by having similar properties, e.g., having similar dependences established by \( \text{Coref} \) (case (b) in (67)). In the former case, \( x_2 \) is coreferential with \( x_1 \) which in turn is coreferential with \( j \), yielding the strict reading. In the latter case, \( x_2 \) has a similar coreference relation to that of \( x_1 \), and \( x_2 \) is linked to the teacher \( t \), yielding the sloppy reading.

\[
(67) \quad \begin{align*}
&\text{before}'(e_{12}, e_{22}) \\
&\text{revise}'(e_{12}, j, p_1) \\
&\quad j : \text{John}'(e_{11}, j) \\
&\quad p_1 : \text{paper}'(e_{15}, p_1) \\
&\quad \text{Poss}'(e_{14}, x_1, p_1) \\
&\quad x_1 : \text{he}'(e_{13}, x_1) \\
&\quad \text{Coref}(x_1, e_{13}, j, e_{11}) \\
&\text{revise}'(e_{22}, t, p_2) \\
&\quad t : \text{teacher}'(e_{21}, t) \\
&\quad p_2 : \text{paper}'(e_{25}, p_2) \\
&\quad \text{Poss}'(e_{24}, x_2, p_2) \\
&\quad x_2 : \text{he}'(e_{23}, x_2) \\
&[\text{Coref}(x_2, e_{23}, x_1, e_{13}) \ (a)] \\
&[\text{Coref}(x_2, e_{23}, t, e_{21}) \ (b)]
\end{align*}
\]

On the whole, the strict/sloppy ambiguity of the target clause results from proving similarity of the representations of the parallel pronouns: they are either coreferential or similar by having similar coreference relations. The reconstructed coreference relations in the target clause must be determined by the coreference relations in the source clause. In the following, we will show that this approach generates five correct readings in the five-reading sentence (68). The five readings are repeated in (69).

\[
(68) \quad \text{John}_1 \ \text{revised} \ \text{his}_1 \ \text{paper} \ \text{before} \ \text{the} \ \text{teacher} \ \text{did}, \ \text{and} \ \text{Bill} \ \text{did} \ \text{too}.
\]

\[
(69) \quad \begin{align*}
1. \ &\text{John} \ \text{revised} \ \text{John}_'s \ \text{paper} \ \text{before} \ \text{the} \ \text{teacher} \ \text{revised} \ \text{John}_'s \ \text{paper}, \ \text{and} \\
&\text{Bill} \ \text{revised} \ \text{John}_'s \ \text{paper} \ \text{before} \ \text{the} \ \text{teacher} \ \text{revised} \ \text{John}_'s \ \text{paper}. \ (\text{JJJJ}) \\
2. \ &\text{John} \ \text{revised} \ \text{John}_'s \ \text{paper} \ \text{before} \ \text{the} \ \text{teacher} \ \text{revised} \ \text{John}_'s \ \text{paper}, \ \text{and} \\
&\text{Bill} \ \text{revised} \ \text{Bill}_'s \ \text{paper} \ \text{before} \ \text{the} \ \text{teacher} \ \text{revised} \ \text{John}_'s \ \text{paper}. \ (\text{JJBJ})
\end{align*}
\]
3. John revised John's paper before the teacher revised John's paper, and
Bill revised Bill's paper before the teacher revised Bill's paper. (JBBB)
4. John revised John's paper before the teacher revised the Teacher's paper, and
Bill revised John's paper before the teacher revised the Teacher's paper. (JTJT)
5. John revised John's paper before the teacher revised the Teacher's paper, and
Bill revised Bill's paper before the teacher revised the Teacher's paper. (JTBT)

Taking the case (*a) in (67) as the source representation (the strict reading), the target
representation should be (70) in which each elided pronoun has two options for its coreference
relation. Let us consider various kinds of coreference relations in the target representation. In
cases (*a1) and (*a3), the pronouns are taken to be coreferential with their parallel pronouns.
In cases (*a2) and (*a4), they are proved to have a similar coreference relation to their parallel
pronouns.

\[
\begin{align*}
&\text{before}'(e_{32}, e_{42}) \\
&\text{revise}'(e_{32}, b, p_3) \\
b : &\text{Bill}'(e_{31}, b) \\
p_3 : &\text{paper}'(e_{35}, p_3) \\
&\text{Poss}'(e_{34}, x_3, p_3) \\
x_3 : &he'(e_{33}, x_3) \\
&[\text{Coref}(x_3, e_{33}, x_1, e_{13}) \ (\ast a1)] \\
&[\text{Coref}(x_3, e_{33}, b, e_{31}) \ (\ast a2)] \\
&\text{revise}'(e_{42}, t, p_4) \\
t : &\text{teacher}'(e_{41}, t) \\
p_2 : &\text{paper}'(e_{45}, p_4) \\
&\text{Poss}'(e_{44}, x_4, p_4) \\
x_4 : &he'(e_{43}, x_4) \\
&[\text{Coref}(x_4, e_{43}, x_2, e_{23}) \ (\ast a3)] \\
&[\text{Coref}(x_4, e_{43}, x_3, e_{33}) \ (\ast a4)]
\end{align*}
\]

The four possible combinations of coreference relations are enumerated below.

1. (*a1) with (*a3),
2. (*a1) with (*a4),
3. (*a2) with (*a3),
4. (*a2) with (*a4).

Let us go through these four combinations and show how many readings can be derived from
them.

First, if the combination of (*a1) and (*a3) is chosen, the first pronoun in the target clause
x_3 is coreferential with x_1 which is coreferential with j. The second pronoun in the target
clause x_4 is coreferential with x_2 which is also coreferential with x_1. Therefore, the two elided
pronouns are coreferential with j. It is known that under the strict reading the two pronouns
in the source clause refer to John. Therefore, this combination yields the reading JJJJ.

The second combination is (*a1) with (*a4). The first elided pronoun x_3 is coreferential with
j. In (*a4), the second elided pronoun x_4 is coreferential with x_3 which is coreferential with
j. Therefore, the second elided pronoun refers to John as well. This combination leads also
to the reading JJJJ.
In the third combination (*a2) with (*a3), the first elided pronoun \(x_3\) is coreferential with \(b\). The second elided pronoun \(x_4\) is coreferential with \(x_2\) which is coreferential with \(j\). As a result, the third case corresponds to the reading **JJBJ**.

Finally, if the combination of (*a2) and (*a4) is chosen, the second elided pronoun \(x_4\) is coreferential with the first elided pronoun \(x_3\) which is coreferential with \(b\). Therefore, both elided pronouns refer to **Bill**. This combination results in the reading **JJBB**.

We have seen that H&K’s analysis derives three correct readings, i.e., **JJJJ**, **JJBJ** and **JJBB**, on the base of the strict reading of the source clause.

Now we consider the sloppy reading of the source clause. Taking case (*b) in (67) as our source representation, the target representation is (71). Cases (*b1) and (*b3) are obtained by proving the pronouns to be coreferential, and (*b2) and (*b4) are the results of proving them to have similar coreference relations.

\[
\begin{align*}
\text{before}'(e_{32}, e_{42}) & \quad \text{revise}'(e_{32}, b, p_3) \\
\text{revise}'(e_{32}, b, p_3) & \quad \text{revise}'(e_{42}, t, p_4) \\
\text{\(b: \text{Bill}'(e_{31}, b)\)} & \quad \text{\(t: \text{teacher}'(e_{41}, t)\)} \\
\text{\(p_3: \text{paper}'(e_{35}, p_3)\)} & \quad \text{\(p_2: \text{paper}'(e_{45}, p_4)\)} \\
\text{Poss}'(e_{34}, x_3, p_3) & \quad \text{Poss}'(e_{44}, x_4, p_4) \\
x_3 \text{ : he}'(e_{33}, x_3) & \quad x_4 \text{ : he}'(e_{43}, x_4) \\
[\text{Coref}(x_3, e_{33}, x_1, e_{13})] & \quad [\text{Coref}(x_4, e_{43}, x_2, e_{23})] \\
[\text{Coref}(x_3, e_{33}, b, e_{31})] & \quad [\text{Coref}(x_4, e_{43}, t, e_{41})]
\end{align*}
\]

There are four combinations of the coreference relations in this case as well:

1. (*b1) with (*b3),
2. (*b1) with (*b4),
3. (*b2) with (*b3),
4. (*b2) with (*b4).

Let us go through these four combinations and show how many readings can be derived from them.

First, if the combination of (*b1) and (*b3) is chosen, the first elided pronoun \(x_3\) is coreferential with \(x_1\) which is coreferential with \(j\) whereas the second elided pronoun \(x_4\) is coreferential with \(x_2\) which is coreferential with \(t\). Therefore, the first elided pronoun refers to **John** while the second pronoun refers to **the teacher**. It is known that under the sloppy reading the two pronouns in the source clause refer to **John** and **the teacher**, respectively. Therefore, this combination yields the reading **JTJT**.

The second combination is (*b1) with (*b4). The first elided pronoun \(x_3\) refers to **John**. According to (*b4), the second elided pronoun \(x_4\) is coreferential with \(t\) again. Therefore, the second elided pronoun refers to **the teacher**. This combination leads to the reading **JTJT** as well.
In the third combination (*b2) with (*b3), the first elided pronoun \( x_3 \) is coreferential with \( b \). Therefore, it refers to Bill. The second elided pronoun \( x_4 \) is coreferential with \( x_2 \) which is coreferential with \( t \). As a result, the third case corresponds to the reading \textbf{JTBT}.

If the fourth combination (*b2) with (*b4) is chosen, we have the equations \( x_3 = b \) and \( x_4 = t \) from which the reading \textbf{JTBT} is derived.

The sloppy reading case of the source clause leads to two readings for the target clause: \textbf{JTJT} and \textbf{JTBT}. These and the three readings in the strict reading case give us the five wanted readings.

Although Kehler and H&K use different frameworks, their basic ideas of dealing with the interaction of VPE and anaphora are the same with respect to the following two points: First, the linking relation between the pronouns and their antecedents are expressed explicitly in the semantic representation. In Kehler, they are represented as linking. In H&K, they are represented by coreference relations. Kehler uses indices to distinguish the terms standing in a linking relation. H&K make use of distinct events under which the terms in a coreference relation are described, although Hobbs admits that their logical form with respect to the events in the coreference relations is ontologically inadequate. Second, the reconstruction of the two linking options for the elided pronouns in the target clause is integrated into the resolution process. Kehler generates the two options by using referring and copying defined in his resolution algorithm. In H&K, the two options are obtained by proving the similarity between the representations of the parallel pronouns. They are either coreferential or have similar coreference relations. H&K can be regarded as an elaboration of Kehler, since there VPE resolution is embedded in a general framework of proving parallelism in a discourse, and parallelism is modeled in a declarative way.

2.6 Conclusion

The discussion of the above analyses highlighted some stages in the development of the research area of VPE and anaphora in the last two decades.

Williams’ analysis is one of the prototypes of the identity-of-relations analyses. In Williams’ analysis, the strict/sloppy ambiguity is caused by the ambiguous interpretation of the pronouns in the source VP by optional use of the Pronoun Rule. We have pointed out several problems arising in the analysis: First, the requirement of the ambiguous interpretation of source pronouns is unmotivated if we consider the meaning of the source clause in isolation. Second, the VP-identity constraint leads to the undergeneration in the cascaded ellipsis (2). Third, the optional use of the Pronoun Rule is insufficient to deal with the many-pronouns-puzzle (5).

The theory proposed by DSP claims that the ambiguity between strict and sloppy readings arises during the resolution process. DSP’s analysis treats sentences like the cascaded ellipsis (2) in an elegant way. It does not undergenerate. However, in this analysis, semantic representations of proper names and pronouns referring to them cannot be distinguished from each other. Therefore, this leads to overgeneration for sentences like (34) which contain multiple occurrences of proper names. In DSP’s analysis, the strict and sloppy readings are generated by the optional abstraction over the secondary occurrence. This method has overgeneration
2.6 Conclusion

problem in the cases of the many-pronouns-puzzle and the five-reading sentence.

Kehler provides a general theory to explain the strict/sloppy ambiguity of the interaction between VPE and anaphora. The elided pronouns are not simply interpreted as strict or sloppy. Their interpretations must be determined according to the reconstructed linking relations. Kehler makes use of the explicit representation of linking relation in the semantic representation. Two options of either referring or copying defined in the resolution algorithm leads to resolution of linking relations in the target clause. However, Kehler's analysis is a procedural approach.

The approach proposed by Hobbs and Kehler is an elaboration of Kehler's analysis. The analysis of VPE and anaphora is integrated into a general framework of dealing with parallelism in discourse. The resolution problem of VPE is explained as a special case of proving the parallelism between the semantic representations of two fragments in the discourse.

We argue that an adequate treatment of VPE cannot be resolved by simply identifying a target VP with its source VP, and the resolution process must establish a parallelism between larger units (clauses) that VP occurs in. In particular, the semantic representation on which the resolution process operates should be able to express the linking relations between pronouns and their antecedents.
Chapter 3

Underspecified Treatment of Ambiguities

The central concern of this thesis is the underspecified treatment of ambiguities in ellipsis. In this chapter, we will give a brief description of the motivation and advantages of underspecified semantic representation. In particular, we will present two approaches to underspecified treatment of ambiguities: Hole Semantics of Bos (1996) and the framework of Context Unification proposed by Ntehren et al. (1997). Both of them can be applied to cases of scope ambiguities. The framework of context unification in addition covers the interaction between quantifier scope and ellipsis.

3.1 Introduction

Traditional theories of natural language semantics like Montague Grammar (Montague 1974) use the generate-and-test method to deal with ambiguous expressions. They generate first representations for all possible readings, and then select the intended one from them. As discussed in Pinkal (1996b), this generate-and-test method is neither computationally feasible nor cognitively plausible. For instance, if an expression has multiple readings, generating and storing all possible readings consumes much resource of the natural language processing system, and leads to a drastic increase of run time. Moreover, it has often been acknowledged that humans do not have great trouble dealing with ambiguous input, and communications among humans do not always require complete disambiguation.

An alternative way of dealing with ambiguity has been intensively investigated in the last years in computational semantics. The aim of semantic underspecification is to produce compact representations of the set of possible readings of an expression. Several semantic underspecification theories have emerged in the last years. They include Quasi Logical Form (QLF) (Alshawi and Crouch 1992), the Underspecified Discourse Representation Theory (UDRT) (Reyle 1993), Minimal Recursion Semantics (Copestake et al. 1995; Egg and Lebath 1995), Hole Semantics (Bos 1996), and the Underspecified Semantic Description Language (USDL) (Pinkal 1996a), the framework of Context Unification (CU) (Ntehren et al. 1997), etc. These approaches have been motivated almost exclusively by computational considerations about
processing natural language semantics. As mentioned in Pinkal (1996a), they are similar in
their basic approaches, describing the range of possible readings in terms of a constraint set,
and modeling the resolution process as a stepwise, monotonic extension of this constraint set,
which corresponds to the subsequent reduction of the range of possible readings.

The advantage of the underspecification mechanism is the improvement of computational
efficiency. In particular, it has proved to be very useful in the area of machine translation, when
the source language and the target language share ambiguities: The underspecified representa-
tion of the source language expression can be directly translated into the underspecified
representation of the target language expression without disambiguation.

Above all, the underspecification approach is cognitively more adequate, since it comes closer
to the way in which humans deal with ambiguous information: either leave the ambiguity
unresolved if it does not disturb communication or wait until more information is available.

3.2 Scope Ambiguities and Verb Phrase Ellipsis

The phenomenon of quantifier scope ambiguity is a challenge for the representation of semantic
ambiguities. Consider sentence (1):

(1) Every man likes a woman.

This example is two ways ambiguous with respect to the scope relation between the quantifying
noun phrases ‘every man’ and ‘a woman’. One reading is that each man likes some woman.
Another reading is that there exists a woman whom every man likes. The first-order logic
formula (2-1) represents the first reading whereas (2-2) corresponds to the second reading.

(2) 1. \( \forall x (\text{man}(x) \rightarrow (\exists y (\text{woman}(y) \land \text{likes}(x, y)))) \)

2. \( \exists y (\text{woman}(y) \land (\forall x (\text{man}(x) \rightarrow \text{likes}(x, y)))) \)

Semantic representations of natural language expressions are conventionally constructed on
the basis of their syntactic analysis, while obeying the compositional principle. The composi-
tional principle requires that the representation of a compound expression depends on
meaning of its parts and the way in which they are combined in the syntax. Montague (1974)
proposes a one to one mapping from a syntactic analysis to an unambiguous type theoretic
representation. He solves the scope ambiguity problem by assuming a syntactic difference
as the base of the different scope orders. Example (1) is assumed to have two syntactic
analyses motivated by the two different scope orderings. But most syntacticians argue that
syntactic analyses should be independent on semantic information. Cooper (1983) and Keller
(1988) propose a storage technique to derive more than one semantic representation from one
syntactic analysis. Their approaches can be regarded as the first step to scope underspeci-
fication. However, their methods are procedural. To overcome the problems discussed above,
various formalisms to scope underspecification have emerged (e.g., QLF, UDRT, USDL, Hole
Semantics, CU, etc.). The basic idea underlying those approaches is to produce one repre-
sentation for all readings. In this chapter, we will discuss two of them: Hole Semantics in
section 3.3 and CU in section 3.4. In Hole Semantics, Bos utilizes constraints with the help of meta-variables to describe underspecified scopes of quantifiers. CU handles scope underspecification in a similar way as done in Hole Semantics. The advantage of CU is that its constraint language can express parallelism between semantic representations too. Therefore, it can be applied to treat the interaction of quantifier scope and ellipsis, e.g., sentences like (3), in an underspecified way. (3), given in Hirschbühler (1982), is called the Hirschbühler sentence.

(3) A Canadian flag was hanging in front of most windows, and an American one was too.

Since the source clause is two ways ambiguous according to the scope relation between the quantifying noun phrases 'a Canadian flag' and 'most windows', the second clause (the target clause) will have two readings as well. However, the parallelism imposed by ellipsis requires that the scope relation in the target clause must be analogous to the scope relation in the source clause. Thus, there are only two readings in (3), instead of four: either the interpretation is the wide scope existential one in both cases (a single Canadian flag as well as a single American one was hanging in front of most windows), or it is the narrow scope existential one (in front of most windows, there was some Canadian flag and in front of most windows, there was some American flag too).

Early ellipsis analyses of Williams (1977) and Sag (1976) can only derive that reading in which both subject noun phrases, namely the existential quantifiers, take the widest scope. DSP use a special method for interpreting quantifier scoping developed by Pereira (1990) to integrate quantifiers to their approach. The method is close to the proposal of Cooper (1983). Therefore, the ambiguity of (3) is a result of the two possible orderings between the different processes of scoping the quantifier and resolving the ellipsis in DSP’s analysis. It begs the question of whether DSP can integrate the possible readings in a compact representation. In contrast to DSP, CU provides a uniform and underspecified treatment of quantifier scope, ellipsis and their interaction.

### 3.3 Hole Semantics

Bos (1996) presents a description language for underspecified semantic representation. It provides a straightforward way for the treatment of scope ambiguities. The basic idea of the description language is formulated by Bos as follows: formulas of the object language (an arbitrary logical language) are first broken into their smallest possible pieces. All of these pieces carry a unique label. The scopes of logical operators are represented by holes as place holders. The various pieces of formulas, together with a set of constraints that tell how the different pieces of structures can be combined, constitute the underspecified semantic representation.

Labels are constants, written as $l_0$, $l_1$, ..., $l_n$. Holes are variables ranging over formulas, written as $h_0$, $h_1$, ..., $h_n$. $\leq$ is a two-place relation between formulas. For example, $l \leq h$ says that a formula with label $l$ is in the scope\(^1\) of an operator, which is represented as the hole

\(^1\) Note that there are different uses of the term scope. For clarity, we call the scope of an logical operator scope, the scope of a quantifier nuclear scope.
3.3 Hole Semantics

$h$. There is always one hole that does not appear anywhere in the labeled formulas: $h_0$. $h_0$ stands for the widest scope — there is no label or hole that outscopes $h_0$. An underspecified semantic representation is an ordered pair $\langle L, C \rangle$, in which $L$ is a set of labeled formulas of the form $l : \varphi$ ($l$ is the label and $\varphi$ is a formula), and $C$ the set of constraints of the form $l \leq h$ for some label $l$ and hole $h$.

In the following, we will apply Hole Semantics to (4) (= (1)).

(4) Every man likes a woman.

Let (5) be the syntactic analysis for (4) on which we can build the underspecified representation. We translate the constituents into pieces of formulas with a unique label, where the logical operators have holes as their arguments. The two determiners ‘every’ and ‘a’ are translated into two quantified formulas with two labels, respectively, where the arguments of logical operators $\rightarrow$ and $\land$ are represented as holes: $l_1 : \forall x(h_1 \rightarrow h_2)$ and $l_2 : \exists y(h_3 \land h_4)$.

The two common nouns ‘man’ and ‘woman’ are represented as $l_3 : \text{man}(x)$ and $l_4 : \text{woman}(y)$. $l_5 : \text{likes}(x, y)$ is taken as the translation of the verb ‘likes’. The labeled formulas form the set $L$.

The set of constraints $C$ is constructed as follows. Either the universal quantifier or the existential quantifier, which are contained in the formulas labeled by $l_1$ and $l_2$, respectively, can have widest scope in the whole formula, described by the two constraints $l_1 \leq h_0$ and $l_2 \leq h_0$. The formulas labeled by $l_3$ and $l_4$ are the restrictors of the quantifiers. Therefore, $l_3$ must be in the scope of $h_1$, represented as $l_3 \leq h_1$, and $l_4$ must be in the scope of $h_3$ ($l_4 \leq h_3$). Since the formula labeled by $l_5$ must be contained in the nuclear scopes of the two quantifiers, there are two more constraints $l_5 \leq h_2$ and $l_5 \leq h_4$.

$L$ and $C$ together form the underspecified representation (6), visualized in Figure 3.1 (the constraints are represented as arrows).

$$
\begin{align*}
(6) \quad & \left\{ \begin{array}{l}
\{ l_1 : \forall x(h_1 \rightarrow h_2) \\
\{ l_2 : \exists y(h_2 \land h_3) \\
\{ l_3 : \text{man}(x) \\
\{ l_4 : \text{woman}(y) \\
\{ l_5 : \text{likes}(x, y) \\
\end{array} \right\}, \quad \left\{ \begin{array}{l}
l_1 \leq h_0 \\
l_2 \leq h_0 \\
l_3 \leq h_1 \\
l_4 \leq h_3 \\
l_5 \leq h_2 \\
l_5 \leq h_4 \\
\end{array} \right\}
\end{align*}
$$
Although the scopes of the quantifiers in (6) are underspecified, the constraints ensure that the quantifiers take their admissible scopes.

Resolution of the underspecified representation in Hole Semantics is done by *plugging* the holes. A plugging is a one-to-one correspondence from holes to labels, where all the constraints are satisfied.

For (6), there are exactly two pluggings, which are listed in table 3.1. Plugging 1 results in the formula (2-1) in which the universal quantifier has the widest scope. Plugging 2 corresponds to the formula (2-2) in which the existential quantifier has a wider scope over the universal one.

<table>
<thead>
<tr>
<th>Pluggings</th>
<th>$h_0$</th>
<th>$h_1$</th>
<th>$h_2$</th>
<th>$h_3$</th>
<th>$h_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$l_1$</td>
<td>$l_3$</td>
<td>$l_2$</td>
<td>$l_4$</td>
<td>$l_5$</td>
</tr>
<tr>
<td>2</td>
<td>$l_2$</td>
<td>$l_4$</td>
<td>$l_1$</td>
<td>$l_3$</td>
<td>$l_5$</td>
</tr>
</tbody>
</table>

The idea of breaking down formulas, labeling them, and putting $\leq$-constraints on them, underlying Hole Semantics, stems from UDRT (Reyle 1993). A notable difference is that Reyle uses labels for Discourse Representation Structures (DRSs), whereas Bos uses them for smaller logic units. Since Bos’s description language is not designed for a specific logical language, it is thereby a generalized version of Reyle’s approach. However, both UDRT and Hole Semantics mix object and meta language symbols in their description language. They can become quite complex in appearance when more phenomena are integrated.

Hole semantics does not attempt to handle parallelism between the semantic representations. Therefore, it does not cover the cases of ambiguities that result from the interaction between quantifier scope and ellipsis.
3.4 The Framework of Context Unification

Nehren et al. (1997) propose a uniform framework in which quantifier scope, ellipsis and their interaction are treated in an underspecified way. The constraint language of CU can express constraints over trees, such as equality, subtree and parallelism constraints. It is used as a description language for underspecified semantic representations. It is a restricted version of the Underspecified Semantic Description Language (USDL) (Pinkal 1996a). For space limitation, we will not describe the difference between the two languages.

Like Hole Semantics, the constraint language of CU is a meta language which specifies the structures of a chosen object language formulas. One of the differences between the constraint language of CU and Hole Semantics is that the object language symbols do not appear in the representation of CU. The clear separation of the meta and the object level facilitates the study of the precise expressive power that is required for the treatment of different phenomena of underspecification, and the computational complexity involved in resolving the representations.

The basic idea of the approach of CU is as follows: Trees are used to represent the object language formulas. The equality and subtree constraints are applied to describe how the pieces of formulas can be combined. Parallelism constraints are applied to represent parallelism between semantic representations. The underspecified representation is a set of constraints.

Let us consider a simple case to illustrate the formalism. We assume that the constraint language of CU takes typed higher-order logic (HOL) as its object language (Dowty, Wall, and Peters 1981). For example, the tree (7), illustrated as the graph (8), represents the HOL formula (9).

(7) \text{walk}@j

(8)
\begin{center}
\begin{tikzpicture}
  \node (a) {\texttt{\textit{@}}};
  \node (b) [below left of=a] {\texttt{\textit{walk}}};
  \node (c) [below right of=a] {\texttt{\textit{j}}};
  \draw (a) -- (b);
  \draw (a) -- (c);
\end{tikzpicture}
\end{center}

(9) \textit{walk}(j)

The binary function symbol @ describes the function application of HOL, taking object language functor and argument as its left and right argument. The constants \text{walk} and j stand for the object language constants \text{walk} and j respectively. A tree consists of subtrees. For instance, \text{walk} as well as j are subtrees of \text{walk}@j in (7).

In the constraint language, the first order meta-variables are defined to denote trees, written as $X_0$, $X_1$, \ldots, $X_n$. The second order meta-variables, called \textit{context variables}, denote one-place functions (context functions) from trees to trees, written as $C_1$, \ldots, $C_n$. The subtree relation is expressed by using meta-variables. Let $X_0$ denote the tree \text{walk}@j, and $X_1$ the tree \text{walk}. The subtree relation between \text{walk}@j and \text{walk} is then represented by the equation (10). $C$ in (10) stands for a context function applied to $X_1$. (10) says that the tree denoted by $X_0$ is equal to an tree which has the tree denoted by $X_1$ as its subtree.
(10) \( X_0 = C(X_1) \)

Now consider (11), depicted in (12).

(11) \text{walk} \at \text{m}

(12) \begin{tikzpicture}
  \node (walk) {walk};
  \node (m) [below right of=walk] {m};
  \draw [->] (walk) -- (m);
\end{tikzpicture}

(11) is structurally parallel to (7), since their structures are equal up-to their subtrees \( m \) and \( j \), respectively. Assuming that \( X_1 \) and \( X_3 \) denote (7) and (11), respectively, the equality up-to relation between the two trees can indirectly expressed in 
CU by the constraint pair in (13). According to Niehren et al. (1997), constraint pairs like (13) are used to represent parallelism between semantic representations in discourse, for example, in ellipsis. We call constraint pairs like (13) \textit{parallelism constraints}.

(13) \[
\begin{align*}
X_1 &= C_1(j) \\
X_3 &= C_1(m)
\end{align*}
\]

Briefly, the set of \textit{well-formed formulas} in the constraint language of CU consists of

(i) A set of non-logical individual constants denoting the object language constants (e.g., \text{walk} for \text{walk}', \text{a} for \text{a}', etc.), and object variables (e.g., \( \text{var}_v \) for every object language variable \( v \)), and sets of \( n \)-place function constants denoting object-language operations (e.g., \( \@ \) for functional application, \( \text{lam}_v \) for lambda abstraction over a variable \( v \) in the object language).

(ii) A set of first-order variables (written as \( X_0, X_1, \ldots, X_n \)), as well as a set of one-place second-order variables (written as \( C_0, C_1, \ldots, C_n \)).

(iii) A set of terms. A term \( t \) is either a first-order variable \( X \), a construction \( f(t_1, \ldots, t_n) \) (\( f \) is either a non-logical individual constant or a \( n \)-place function constant, and \( t_i \) (\( 0 \leq i \leq n \)) is a term too.), or an application of a second-order variable \( C \) to a term \( t' \), i.e., \( C(t') \).

(iv) A set of context constraints. A context constraint \( \varphi \) is either an equation of two second-order terms \( t = t' \) or a conjunction of two context constraints \( \varphi' \land \varphi'' \).

Within the framework of CU, resolution of constraints means applying the algorithm of CU to constraints (Pinkal 1996a; Niehren et al. 1997). The algorithm of CU is a restricted version of Huet’s higher order unification (HOU) algorithm (Huet 1975). Its task is to find substitutions for free variables to make the two sides of an equation identical. However, the substitutions must obey an additional constraint, the so-called \textit{linearity constraint}:

For any second-order variable \( C \), a well-formed substitution must assign to \( C \) a lambda term of the form \( \lambda X.t \) in which the lambda bound variable occurs exactly once in the second-order term \( t \). \( t \) is termed as \textit{body} of \( \lambda X.t \).
The intuitive motivation of the linearity constraint is that the semantic information of an expression must contribute exactly once to the semantic information of a compound expression in which it occurs. With the substitution constraint, two cases allowed in HOU are excluded: one case is that the lambda bound variable \( X \) does not occur in the body of \( \lambda X.t \), and another case is that it occurs more than once in the body \( \lambda X.t \). The first case would lead to overwriting of semantic information. For example, let \( C_1 \) in (13) be substituted by \( \lambda X.\text{walk@j} \). Applying \( C_1 \) to \( m \) and \( j \) result in \( \text{walk@j} \). The semantic information of \( m \) and \( j \) is overwritten. The second case causes multiple copies of one piece of semantic information. For instance, if \( C_1 \) is substituted by \( \lambda X.X@X \), applying \( \lambda X.X@X \) to \( m \) leads to \( m@m \).

In the following, we apply CU to (14) (= (1)), an example of scope ambiguity. As mentioned earlier, (14) is two ways ambiguous with respect to the scope relation of the quantifying noun phrases ‘every man’ and ‘a woman’.

(14) Every man likes a woman.

We assume that (5) is the syntactic analysis for (14) on which we can build the underspecified representation. (15) shows the semantic interpretation rules for obtaining constraints from binary syntax trees. They are defined in Nihren et al. (1997). In (15), each node \( P \) in the syntactic structure is associated with three semantic meta-variables \( X_P, X'_P \) and \( C_P \), and \( I(P) \) is the scope boundary for each node \( P \).

(15) (i) For every S-node \( P \) add \( X_P = C_P(X'_P) \), for any other node add \( X_P = X'_P \).

(ii) If \( [P \ Q \ R] \), \( Q \) and \( R \) are not NP nodes, add \( X_P = X_Q@X_R \) or \( X'_P = X_R@X_Q \), according to their types in object language\(^2\).

(iii) If \( [P \ Q \ R] \) or \( [P \ R \ Q] \), and \( R \) is an NP node, then add \( X'_P = X_Q@\text{var}_x \) and \( C_{I(P)} = \lambda X.C_0(X_R@\text{lam}_x(C_1(X))) \).

The application of rule (iii) to the two quantified NPs, where the values for the quantifiers \( \text{every@man} \) and \( \text{a@woman} \) are substituted in for the variables \( X_R \) in both cases, results in the first two constraints in (16). The quantifiers themselves are put together by rule (ii). The third constraint results from rule (i) when the semantics of \( X'_s \) is filled in. \( X'_s \), i.e., \( (\text{likes@var}_y)@\text{var}_x \), is a byproduct of the applications of rule (iii) to the two NPs.

(16) 1. \( C_s = \lambda X.C_1((\text{every@man}@\text{lam}_x(C_3(X)))) \land \)

2. \( C_s = \lambda X.C_2((\text{a@woman}@\text{lam}_y(C_4(X)))) \land \)

3. \( X_s = C_s((\text{likes@var}_y)@\text{var}_x) \)

In (16-1) and (16-2), the right hand side of each equation is a lambda term, and its body contains the description of a quantifier. The meta-variables occurring in the body indicate that the nuclear scope of each quantifier and the scope in which each quantifier occur is still

\(^2\) Types in the object language determine the functor and the argument of functional application in the meta language. As defined in CU, the right hand side argument of \( \odot \) is the functor, while the left hand side argument is the argument.
underspecified. However, the application of $C_s$ to $(\text{likes} @ \text{var}_y) @ \text{var}_x$ in (16-3) ensures that $(\text{likes} @ \text{var}_y) @ \text{var}_x$ occurs inside of the nuclear scope of each quantifier.

In the approach of CU, a closure operation is needed, i.e., the remaining free context variables are projected to $\lambda X. X$. The closure operation gives the minimal solutions of the constraints. The application of the algorithm of CU to (16) after closure yields the two possible scope readings of (14) given in (17). (17-1) specifies the reading in which ‘every man’ has wide scope whereas (17-2) corresponds to the reading in which ‘a woman’ has wide scope.

(17) 1. $X_s \rightarrow$
      \[
      (\text{every} @ \text{man}) @ \text{lam}_x ((a @ \text{woman}) @ \text{lam}_y ((\text{likes} @ \text{var}_y) @ \text{var}_x))
      \]

2. $X_s \rightarrow$
      \[
      (a @ \text{woman}) @ \text{lam}_y ((\text{every} @ \text{man}) @ \text{lam}_x ((\text{likes} @ \text{var}_y) @ \text{var}_x))
      \]

The advantage of the approach of CU is that it can represent parallelism between semantic representations by using parallelism constraints. The underspecified representation of an elliptical construction is a set of constraints, which consists of the constraints obtained by interpreting the source and the target clause independently of each other, and the parallelism constraint pair given in (18). $X_s$ and $X_t$ represent the semantics of the source clause and the target clause whereas $X_{ps}$ and $X_{pt}$ refer to the semantics of the parallel elements in the source and the target. It states that the trees representing the semantics of the source and the target clause must be equal up-to the positions corresponding to the semantics of the parallel elements.

(18) $X_s = C(X_{ps}) \land X_t = C(X_{pt})$

As a simple example, consider the elliptical sentence (19). For simplicity, we assume that proper names are interpreted by constants and not quantifiers.

(19) Dan likes Mary, and Bill does too.

(19) is analyzed as the constraint in (20).

(20) $X_s = \text{likes} @ \text{mary} @ \text{dan} \land$

$X_{ps} = \text{dan} \land X_{pt} = \text{bill} \land$

$X_s = C(X_{ps}) \land X_t = C(X_{pt})$

By applying the algorithm of CU to (20), $C$ is substituted by $\lambda X. \text{likes} @ \text{mary} @ X$. This yields the interpretation of the target clause, which is given by $X_t \rightarrow \text{likes} @ \text{mary} @ \text{bill}$.

We now look into the interpretation of the Hirschbühler sentence (21) (= (3)), in which the source clause is scope ambiguous and interpretation of the target clause must obey the parallelism requirement in the discourse. ‘a Canadian flag’ and ‘an American one’ are the parallel elements.
(21) A Canadian flag was hanging in front of most windows, and an American one was too.

(21) is analyzed as a constraint, which is a conjunction of the constraints (22) and (23).

\[
C_s = \lambda X.\text{C}_1((a\@\text{Can\_flag})@\text{lam}_x(\text{C}_3(X))) \land \\
C_s = \lambda X.\text{C}_2((\text{most\_windows})@\text{lam}_y(\text{C}_4(X))) \land \\
X_s = C_s((\text{hang\_in\_front\_of\_var}_y)@\text{var}_x)
\]

(23) \[
X_s = C(a\@\text{Can\_flag}) \land \\
X_t = C(a\@\text{Amer\_one})
\]

The conjunction of the constraints in (22) and (23) has the two correct solutions (24) and (25) with corresponding scopings in $X_s$ and $X_t$ and excluded the unwanted ones.

\[
X_s \mapsto (a\@\text{Can\_flag})@\text{lam}_x((\text{most\_windows})@\text{lam}_y((\text{hang\_in\_front\_of\_var}_y)@\text{var}_x))
\]

\[
X_t \mapsto (a\@\text{Amer\_one})@\text{lam}_x((\text{most\_windows})@\text{lam}_y((\text{hang\_in\_front\_of\_var}_y)@\text{var}_x))
\]

\[
C \mapsto \lambda X.\text{X}@\text{lam}_x((\text{most\_windows})@\text{lam}_y((\text{hang\_in\_front\_of\_var}_y)@\text{var}_x))
\]

(25) \[
X_s \mapsto (\text{most\_windows})@\text{lam}_y((a\@\text{Can\_flag})@\text{lam}_x((\text{hang\_in\_front\_of\_var}_y)@\text{var}_x))
\]

\[
X_t \mapsto (\text{most\_windows})@\text{lam}_y((a\@\text{Amer\_one})@\text{lam}_x((\text{hang\_in\_front\_of\_var}_y)@\text{var}_x))
\]

\[
C \mapsto \lambda X.(\text{most\_windows})@\text{lam}_y(X@\text{lam}_x((\text{hang\_in\_front\_of\_var}_y)@\text{var}_x))
\]

In example (21), we have illustrated the underspecified and uniform approach of CU to VPE and its interaction with quantifier scope.

However, the constraint language of CU cannot distinguish the occurrences of trees. As a consequence, the semantic representation of the primary occurrence (the semantics associated with the source parallel element) cannot be distinguished from other identical semantic representations. That leads to overgeneration in the cases of multiple proper names like (26).

(26) Felix's mother thinks Felix is a genius, and Siegfried's mother does too.

(26) has only one reading (both Felix's mother and Siegfried's mother think that Felix is a genius). We assume that the parallel elements are the two embedded proper names in subjects: 'Felix' and 'Siegfried'. In the source clause, there is a second occurrence of the proper name 'Felix'.

(26) is analyzed by (27).

\[
X_s = \text{think}@((a\_\text{genius}@\text{felix})@\text{mother}@\text{of}@\text{felix}) \land 
\]
\[ X_s = C(\text{felix}) \land X_t = C(\text{siegfried}) \]

After applying the algorithm, we get two solutions (28) and (29). (28) is the wanted solution and corresponds to the accepted reading of the target clause. But (29) is an overgenerated one, which represents a reading such that Felix's mother thinks Siegfried is a genius.

(28) \[ C \mapsto \lambda X.\text{think}_@(@\text{genius}@\text{felix})@(@\text{mother}_@@X) \]
\[ X_t \mapsto \text{think}_@(@\text{genius}@\text{felix})@(@\text{mother}_@@\text{siegfried}) \]

(29) \[ C \mapsto \lambda X.\text{think}_@(@\text{genius}@X)@(@\text{mother}_@@\text{felix}) \]
\[ X_t \mapsto \text{think}_@(@\text{genius}@\text{siegfried})@(@\text{mother}_@@\text{felix}) \]

The overgeneration problem arising from sentences like (26) indicates that the framework of CU must be extended and modified to cover the interaction of VPE and anaphora. As discussed in Chapter 2, the semantic representation on which the resolution process operates must be able to model the linking relation between pronouns and their antecedents to deal with anaphora in VPE. It states that the constraint language of CU should be able to distinguish between the occurrences of semantic material. However, the constraint language of CU does not meet this requirement.

To sum up, the framework of CU employs a constraint language which can express constraints over trees such as equality, subtree and parallelism constraints. The constraints are solved by context unification. The constraint language is suitable to treat semantic underspecification and parallelism phenomena in discourse such as quantifier scope, VPE and their interaction in a uniform way. Actually, the idea of using the equality and subtree constraints can also be found in Hole Semantics. For example, labels as well as holes behave very much like first order variables. A labeled formula expresses also an equation. For example, \( l_1 : \text{woman}(x) \) can be rewritten as an equation like \( X_1 = \text{woman}@x \). In addition, the relations \( \leq \) between labels and holes correspond to the subtree constraint. For instance, \( l_1 \leq h_1 \) could be expressed as a constraint, \( X_2 = C(X_1) \), in which \( X_2 \) stands for \( h_1 \) and \( X_1 \) for \( l_1 \).

### 3.5 Conclusion

According to the approaches of Hole Semantics and the framework of CU, the underspecified semantic representation of a natural language expression is a description of a set of logical formulas, a compact representation of a set of readings. Hole semantics and the framework of CU can be applied to scope ambiguity straightforwardly. Their basic ideas are similar. Both of them use meta-variables to specify the unresolved scopes of quantifiers.

In Hole Semantics, an underspecified representation consists of a set of labeled formulas, a set of meta variables that represent scope (holes), and a set of constraints on these. However, in this approach there is no clear separation between the object language and the meta language symbols in the representation.

The constraint language employed by CU can express equality and subtree constraints with the help of meta-variables. In CU, the equality and subtree constraints are applied to represent
3.5 Conclusion

scope underspecification. In contrast to Hole Semantics, the object language symbols do not appear in the representation of the constraint language of CU.

CU is more expressive than Hole Semantics, since its constraint language can express parallelism between trees. Therefore, it can handle the interaction of quantifier scopes and VPE.

In general, an adequate ellipsis analysis should be able to treat the interaction of ellipsis and anaphora as well. The framework of CU does not deal with anaphora. A semantic representation language would be required to be able to distinguish the occurrences of the representations of pronouns from each other as well as their antecedents to represent the linking relations. But the constraint language of CU does not meet this requirement. The aim of this thesis is to extend the framework of context unification so that phenomena of VPE and anaphora can be dealt with. In the next chapter, we will provide an analysis of the interaction of VPE and anaphora. In addition, we integrate the analysis into an extended framework of CU, called Constraint Language for Lambda-link Structures (CLLS), in which the occurrences of semantic information can be distinguished from each other and relations between the occurrences can be expressed as well.
Chapter 4

Underspecified Treatment of Verb Phrase Ellipsis and its Interaction with Anaphora and Quantifiers

This chapter deals with the underspecified semantic representation and resolution of ellipsis. We focus on the interaction of Verb Phrase Ellipsis (VPE) and anaphora. We present a linguistic analysis of the distribution of the strict and sloppy readings in VPE. We predict that the strict/sloppy ambiguity results from the interaction of anaphoric links and parallelism requirement in ellipsis. In section 4.1, we present a Constraint Language for Lambda-link Structures (CLLS), which we adopt for our analysis. In section 4.2, we will provide the analysis, illustrate its integration into CLLS with the help of some examples and show that the analysis solves a number of benchmark puzzles discussed in the literature. In section 4.3, we will apply CLLS to some examples of the interaction of quantifier scope, anaphora and VPE. In section 4.4, we show that CLLS can be generalized straightforwardly to deal with extended parallelism (Hobbs and Kehler 1997) and some examples of stripping and gapping.

4.1 A Constraint Language for Lambda-link Structures

CLLS is an extension of the constraint language of Context Unification (CU) (Niehren et al. 1997), which was discussed in Chapter 3. CLLS cannot only describe the interaction between quantifier scope and parallelism, but also the interaction between anaphora and parallelism. CLLS expresses constraints over lambda-link structures. Lambda-link structures are tree structures with additional binding relations (for lambda bindings) and linking relations (for anaphoric links).

We use CLLS as our description language for semantic underspecification which takes the typed Higher Order Language (HOL) or some other logical language as its object language. Every lambda-link structure characterizes a unique λ-term or a logical formula modulo consistent renaming of bound variables. We represent lambda-link structures as graphs in order to facilitate understanding. For example, the lambda-link structure given in (2) characterizes the HOL-formula (1).
4.1 A Constraint Language for Lambda-link Structures

(1) \((\text{many}(\text{language}))(\lambda x.\text{spokenby}(\text{john})(x))\)

(2)

\[
\begin{array}{c}
\text{many} \\
\text{language} \\
\text{spokenby} \\
\text{john}
\end{array}
\]

binding of the variable \(x\) by the \(\lambda\)-operator \(\lambda x\) is represented by an explicit binding relation between two nodes, which are labeled as \(\text{var}\) (for \(x\)) and \(\text{lam}\) (for \(\lambda x\)), respectively. In (2), the binding relation is represented as a dotted arrow. As the binding relation is explicit, the variable and the binder need not be given a name or index.

CLLS expresses constraints over lambda-link structures. CLLS consists

(i) a set of labels \(\Sigma\) which consists of

1. non-logical individual constants denoting the object language constants (e.g., \text{mary} for \text{mary}, \text{read} for \text{read}), and the object variables (e.g., \text{var} for each object language variable)

2. n-place function constants denoting object language operations (e.g. \text{@} for functional application, \text{lam} for a \(\lambda\)-binder such as \(\lambda x\) which binds a object language variable \(x\)).

(ii) a set of first order variables, written as \(X_1, \ldots, X_n\) or \(Y_1, \ldots, Y_n\). They denote nodes in a tree.

(iii) a set of constraints. There are various types of constraints which specify the different kinds of relations between nodes in a tree. They are

1. Labeling: \(X:f(X_1, \ldots, X_n)\) \(\quad (f^n \in \Sigma)\)
2. Dominance: \(X_1 <^* X_2\)
3. Divergence: \(X_1 \perp X_2\)
4. Binding: \(\lambda(X_1, X_2)\)
5. Linking: \(\text{link}(X_1, X_2)\)
6. Parallelism: \(X_s/X_{sp} \sim X_t/X_{tp}\).

The conjunction of two constraints \(\varphi\) and \(\varphi'\) like \(\varphi \land \varphi'\) is a constraint too.

In the following, we will give a precise description of various types of constraints with the help of some examples. A constraint is an underspecified description of a lambda-link structure. We represent the constraints for lambda-link structures as graphs as well. It is important to distinguish the graphs for lambda-link structures from the graphs for constraints which are seen as descriptions of the lambda-link structures. For instance, we can fully describe
the above lambda-link structure (2) by means of a constraint $\varphi$ in (3). $\varphi$ is a conjunction of labeling constraints and a binding constraint. $\varphi$ can be represented as a graph, given in (4). Note that (4) is the description of (2), not (2) itself.

\[
\varphi \equiv X_1: @(X_2, X_5) \land X_2: @(X_3, X_4) \land X_3: \text{many} \land X_4: \text{language} \land \\
\lambda(X_5, X_{10}) \land X_5: \text{lam}(X_6) \land X_{10}: \text{var} \land \\
X_6: @(X_7, X_{10}) \land X_7: @(X_8, X_{10}) \land X_8: \text{spokenby} \land X_9: \text{john}
\]

(4)

Labeling Constraint Each node has a single label. The labels are constants. Nodes labeled as nonlogical individual constants (e.g., var, ana, john etc.) are terminal nodes. They do not have daughters. Nodes labeled with $n$-place function constants ($1 \leq n$) have $n$ daughters. For example, the two-place function constant @ has two daughters. Let's consider an example constraint $\varphi$ that is represented in (5).

\[
\varphi \equiv X_1: @(X_2, X_3) \land X_2: \text{walk} \land X_3: \text{john}
\]

(5)

\varphi is a conjunction of three labeling constraints. $X_2: \text{walk}$ describes that $X_2$ is a terminal node labeled as walk. $X_3$ is also a terminal node labeled as john. $X_1: @(X_2, X_3)$ describes that $X_1$ labeled as @ has $X_2$ and $X_3$ as its daughters. In this case, $X_1$ dominates $X_2$ and $X_3$ immediately. We represent immediate dominance as undotted lines.

Dominance Constraint Dominance is the transitive and reflexive closure of immediate dominance (Vijay-Shanker 1992). We represent dominance constraints graphically as dotted lines. Constraint (6) describes a lambda-link structure in which the two nodes $X_1$ and $X_2$ lie between an upper bound $X_0$ and a lower bound $X_3$. The two solution alternatives for (6) are represented in (7), in which either $X_1$ dominates $X_2$ or $X_2$ dominates $X_1$.

(6)
4.1 A Constraint Language for Lambda-link Structures

Consider the constraint in (8). If there is a dominance relation between the two nodes $X_1$ and $X_2$ and no other nodes and relations (e.g., linking or binding) intervene, a closure operation can be applied to let $X_1$ and $X_2$ be the identical node. The solution is represented in (9).

\begin{equation}
\begin{array}{c}
\mathbf{X_0} \\
\mathbf{X_1} \\
\mathbf{X_2} \\
\mathbf{X_3}
\end{array}
\end{equation}

\begin{equation}
\begin{array}{c}
\mathbf{X_0} \\
\mathbf{X_1} \\
\mathbf{X_2} \\
\mathbf{X_3}
\end{array}
\end{equation}

\begin{equation}
\begin{array}{c}
\mathbf{X_0} \\
\mathbf{X_1, X_2} \\
\mathbf{X_3}
\end{array}
\end{equation}

**Divergence Constraint** Divergence is defined as the absence of dominance in both directions. In (5), the relation between $X_2$ and $X_3$ can be described as a divergence constraint $X_2 \perp X_3$, since neither $X_2$ dominates $X_3$ nor $X_3$ dominates $X_2$.

**Binding Constraint** A binding constraint is represented as $\lambda(X,Y)$. It states that node $Y$ labeled as var may be bound by node $X$ labeled as lam. A binding constraint between $X$ and $Y$ holds if $X$ dominates $Y$, and $X$ is the single binder of $Y$.

The interaction of binding constraints and dominance constraints offers a simple means of representing scope underspecification. In (10), we have the typical case of the undetermined scopes of quantifying noun phrases. (10) is analyzed by constraint (11), where nodes $X_3$ and $X_{11}$ (labeled as lam), bind nodes $X_8$ and $X_7$ (labeled as var), respectively, represented as dotted arrows from binder nodes to variable nodes. The underspecified dominance relation between $X_1$ and $X_9$ describes the scope underspecification between the two quantifying noun phrases 'every man' and 'a woman'.

\begin{equation}
\text{Every man loves a woman.}
\end{equation}
The graph can be linearized by adding either a constraint \( X_1 \equiv X_9 \) or \( X_9 \equiv X_1 \). After applying closure operation, the linearized solutions are (12) and (13), in which bindings are not destroyed and the variables are inside of the scopes of their binders. They correspond to the two possible scoping readings for the sentence (10). (12) corresponds to the HOL formula (14-1), while (13) describes (14-2).
4.2 Underspecified Representation and Resolution of Verb Phrase Ellipsis and Anaphora

(14) 1. every_man(λx.(a_woman(λy.(loves(x,y)))))
     2. a_woman(λy.(every_man(λx.(loves(x,y)))))

Parallelism Constraint  We analyze elliptical constructions by means of a parallelism constraint of the form

(15) \(X_s/X_{sp} \sim X_t/X_{tp}\).

A parallelism constraint holds if the two contexts, i.e., the trees immediately below \(X_s\) and \(X_t\), have the same structure and the same labels up to the nodes \(X_{sp}\) and \(X_{tp}\), see (16). We call \(X_{sp}\) and \(X_{tp}\) exception nodes.

Linking constraints and their interaction with parallelism constraints will be discussed in the next section. The interaction of binding constraints and parallelism constraints will be discussed in section 4.3.

The semantics of CLLS is formally defined in Egg et al. (1998). Disregarding linking constraints and their interaction with parallelism constraints, CLLS extends the expressiveness of CU in the following points: First, the first order variables in CLLS are interpreted as nodes rather than whole trees. Since each node occurs exactly once in a tree, we can express the distinction of different occurrences of semantic material by the distinction of nodes. The single occurrence feature of nodes is crucial for the analysis of the interaction between VPE and anaphora. Second, in CLLS binding relations without explicit representation of the variable names avoid the variable capturing problem which CU suffers from. This point is discussed in Egg et al. (1998). Since it is not the central concern of this thesis, we will not go deeper into it. CLLS extends the coverage of the approach of CU to a framework that integrates an analysis of VPE with anaphora. In the next section, we will discuss our analysis and its integration into CLLS.

4.2 Underspecified Representation and Resolution of Verb Phrase Ellipsis and Anaphora

In this section, we pay special attention to the strict/sloppy ambiguity in the interaction of VPE and anaphora, as in sentence (17).
(17) Max$_1$ saw his$_1$ mother, and Bill did, too.

As discussed earlier, in the reading of the source clause in which ‘his’ refers to ‘Max’, the target clause in (17) may receive two readings, namely one in which Bill saw Max’s mother (the *strict* reading), and one in which Bill saw his own mother (the *sloppy* reading). In (17), ‘Max’ and ‘Bill’ are *parallel elements*. Parallel elements are constituents in the source clause and their overt parallel counterparts in the target clause.

We claim that the strict/sloppy ambiguity results from the interaction of the linking relation between ‘his’ and ‘John’ and parallelism requirement between the source and the target clause. The linking relation describes the referential dependence between an anaphoric element and its antecedent. The definition of linking relations in our analysis is based on the *role linking rule* in Kehler (1995). Both Kehler (1995) and our analysis claim that there is a correspondence between binding relations in syntactic representations (Chomsky 1981) and linking relations in semantic representations. Normally, syntactic representations are trees too. Note that trees for the syntactic representations are not the same as trees in CLLS. Constituents are nodes in a syntactic tree. According to Chomsky (1981), node $A$ is defined to bind node $B$ in the syntactic representation iff $A$ and $B$ are co-indexed and $A$ c-commands $B$. A c-commands $B$ iff $A$ does not dominate $B$ and $B$ does not dominate $A$ and the first branching node dominating $A$ also dominates $B$. In the syntactic literature, there are various versions of definitions for binding and c-command. For our purpose, we only need those presented above. The definition for *linking* is given as follows.

**Definition for linking**

The semantic representation of an anaphoric element is linked to the semantic representation of another element which binds it most immediately in the syntax. If there is no binder available, then it is linked to the most salient discourse referent.

In our linguistic analysis, the strict/sloppy ambiguity arises from the interaction of linking relations in the source clause and parallelism requirement which is defined below.

**Parallelism requirement for the interaction of VPE and anaphora**

If an anaphoric constituent in the source clause refers to the source parallel element and it has an elided parallel anaphoric constituent in the target clause, then the semantic representation of the elided anaphoric constituent must either

(i) be coreferential with the semantic representation of its parallel constituent in the source clause, i.e., linked to the semantic representation of its parallel constituent in the source clause (referential parallelism)

or

(ii) be linked to the semantic representation of a constituent in the target clause. The newly created link must be structurally parallel to the corresponding link in the source clause. (structural parallelism)

Our linguistic analysis is integrated into CLLS in the following way: Linking relations between the semantic representations of pronouns and their antecedents are expressed by linking constraints; the options for anaphoric linking of the elided pronouns as expressed in parallelism requirement are spelt out in the resolution algorithm.
**Linking Constraint** A linking constraint \( \text{link}(X,Y) \) means that \( Y \) labeled as \( \text{ana} \) (i.e. an anaphora) is linked to another terminal node \( X \). In a semantic interpretation of lambda-link structures, linked nodes get identical denotations. The construction of a linking constraint is based on the linking concept in our linguistic analysis.

For example, in (17), the pronoun ‘his’ is linked to the subject ‘Max’ in the semantic representation, since ‘his’ is bound by ‘Max’ most immediately in the syntax (There is no other binder available.). (18) is a representation for the source clause of (17). \( X_5 \) stands for the semantic representation of ‘his’ whereas \( X_{sp} \) is the semantic representation of ‘Max’. The linking relation between ‘his’ and ‘Max’ is expressed by \( \text{link}(X_5, X_{sp}) \). We represent linking relations as dashed arrows.

\[
(18)
\]

Links can form chains. The constraint (20) analyzes (19), where the second pronoun is linked to the first, rather than linked to the proper name. We call \( X_{10} \) the *first element of the linking chain*.

\[
(19) \quad \text{John}_1 \text{ said he}_1 \text{ liked his}_1 \text{ mother.}
\]

\[
(20)
\]

**Interaction of linking constraints and parallelism constraints** If linking constraints interact with a parallelism constraint \( X_s/X_{sp} \sim X_t/X_{tp} \), two isomorphic anaphora (nodes labeled as \( \text{ana} \)) within the parallel contexts (i.e., trees immediately below \( X_s \) and \( X_t \)) must have either isomorphic links (structural parallelism) or have identical reference, i.e., one is linked to the other (referential parallelism), and the two anaphoric nodes must be elements in linking chains which have exception nodes as their first elements.

Assuming that \( X_s \) and \( X_t \) are semantic representations for the source and the target clause, and \( X_{sp} \) and \( X_{tp} \) are semantic representations for parallel elements (the exception nodes). (21) is an underspecified representation of the elliptical sentence (17).
(22) and (23) are the two solutions for (21). The parallelism constraint is satisfied in both solutions, since the two contexts below $X_s$ and $X_t$ have the same structure and the same labels up to the exception nodes $X_{sp}$ and $X_{tp}$. $Y_5$ is the semantic representation for the elided pronoun. In (22), $Y_5$ is linked to $X_5$ (referential parallelism). In (23), the link between $X_5$ and $X_{sp}$ is isomorphic to the link between $Y_5$ and $Y_{tp}$ (structural parallelism). In (22) and (23), $X_5$ and $Y_5$ are elements of linking chains which contain the exception nodes as their first elements. A linked node receives the same interpretation of the node it is linked to. In (22), $Y_5$ has the same interpretation as $X_5$ (strict reading). In (23), $Y_5$ has the same interpretation as $X_{tp}$ (sloppy reading).

(22)

(23)

We have shown that our approach is suitable to treat the basic case of the strict/sloppy ambiguity in example (17).

In contrast to the pronoun in example (17), the pronoun ‘he’ in sentence (24) is not bound by the coreferential NP ‘Felix’, because ‘Felix’ does not c-command ‘he’. Nevertheless, Reinhart
(1983) and Dalrymple et al. (1991) note that this sentence has two readings, one in which Siegfried's mother thinks Felix is a genius (the strict reading), and one in which Siegfried's mother thinks her own son is a genius (the sloppy reading). The analyses of Sag (1976) and Williams (1977) fail to generate the sloppy reading.

We will show that bindings in the syntax are not a necessary condition for the sloppy interpretations of pronouns for our analysis.

(24) Felix$_1$'s mother thinks he$_1$'s a genius and so does Siegfried$_2$'s mother.

In (24), 'he' is not bound by any referential elements in the syntax. 'Felix' is the single coreferential NP in this case. According to the definition of linking, 'Felix' is therefore the most salient discourse referent to be linked to.

We assume that 'Felix' and 'Siegfried' are parallel elements in this elliptical construction. The description for (24) is depicted in (25). The pronoun, represented as node $X_6$ (labeled as ana), is linked to the representation of the source parallel element $X_{sp}$ (labeled as felix). For simplicity, we omit the intension operator for the interpretation of 'think'.

(25)

Solving the parallelism constraint $X_s/X_{sp} \sim X_t/X_{tp}$ yields two $\lambda$-link structures, represented as (26) and (27). In both solutions, the parallelism constraint is satisfied: The tree below $X_t$ is a copy of $X_s$ up to $X_{tp}$. In (26), node $Y_6$ labeled as ana is linked to $X_6$ to gain referential parallelism, resulting in the strict reading. In (27), the link from $X_6$ to $X_{sp}$ is isomorphic to the link from $Y_6$ to $X_{tp}$ (structural parallelism), yielding the sloppy reading.
(28) differs from (24), since its subject in the embedded clause of the source clause is a proper name, not a pronoun.

(28) Felix is a genius and so does Siegfried's mother.

In contrast to (24), (28) has only one reading such that Siegfried's mother thinks Felix's a genius. There is no sloppy reading available. As discussed in Chapter 3, the approach of CU overgenerates in the case of (28), because the constraint language of CU cannot express the distinction of the occurrences of the semantic representation for 'Felix'. In our approach, we can distinguish the different occurrences of the semantic representation of 'Felix' by using different first order variables. For example, we represent the semantics 'Felix' occurring in 'Felix's mother' as $X_{sp}$. The semantics of the other occurrence of 'Felix' is represented as $X_6$. (28) is analyzed as (29). Since there is no linking constraint in the source representation, the resolution of (29) is (30), which corresponds to the admissible reading.
(30)

In the following, we demonstrate that our approach accounts for the case of the cascaded ellipsis, repeated as sentence (31).

(31) John\textsubscript{1} realizes that he\textsubscript{1} is a fool, but Bill\textsubscript{2} does not, even though his\textsubscript{2} wife does.

Assuming that the pronoun ‘he’ in the first clause refers to ‘John’ and the pronoun ‘his’ in the third clause refers to ‘Bill’, this sentence has not only the two readings in which the two elided pronouns in the second and the third clause have a strict or a sloppy reading simultaneously, but also a third reading (32), in which the elided pronoun in the third clause has a strict reading whereas the elided pronoun in the second clause has a sloppy reading. We attempt to derive (32). (32) is problematic for the identity-of-relations analyses, e.g., Williams’ analysis, discussed in Chapter 2.

(32) John realizes that John is a fool, but
Bill does not realize that Bill is a fool, even though
Bill’s wife realizes that Bill is a fool.
(sloppy, strict)

In the source clause, the pronoun ‘he’ is linked to the source parallel element ‘John’, represented graphically in (33). (33) is a simplified representation for the description of the source clause (31). For simplicity, we use simplified representations for the further examples in this section.
John$_1$ realizes that he$_1$ is a fool.

To derive (32), the second clause must have a sloppy reading. The sloppy reading, represented as (34), is obtained by linking the elided pronoun to the target parallel element ‘Bill’. The link in (34) is structurally parallel to the link in (33).

Bill$_2$ does not realize he$_2$ is a fool.

Taking (34) as the source representation for the third clause, its link licenses two options for the reconstructed link in the third clause. Referential parallelism yields the strict reading of the third clause, represented in (35).

Bill$_2$ does not realize he$_2$ is a fool,

his$_2$ wife realizes he$_2$ is a fool.

(33) and (35) together form the representation for (32).

We now turn to the case in which there are several coreferential pronouns in the source clause, the many-pronouns-puzzle (36). As mentioned earlier, if the second elided pronoun has a sloppy interpretation in (36), the first elided pronoun must be interpreted sloppily as well, hence the unavailability of the interpretation in (37-4).

John$_1$ said that he$_1$ liked his$_1$ mother, and Bill did too.

Bill did too.

1. Bill said that John liked John's mother. (strict, strict)
2. Bill said that Bill liked Bill's mother. (sloppy, sloppy)
3. Bill said that Bill liked John's mother. (sloppy, strict)
4. *Bill said that John liked Bill's mother. (strict, sloppy)

Let us consider the linking relations in the source clause of (36). The first pronoun ‘he’ is bound most immediately by the subject ‘John’. Therefore, ‘he’ is linked to ‘John’. The interesting point about this example is that the second pronoun ‘his’ can be either bound by ‘he’ or ‘John’ in the syntactic analysis. But ‘his’ must be linked to ‘he’ because it is bound by ‘he’ most immediately. Hence, the admissible links for the source clause of (36) are depicted in (38), in which ‘John’ links ‘he’ and ‘he’ links ‘his’. The two links form a linking chain. One of the features of a linking chain is that the referential information of its first element is copied down the chain.

John$_1$ said he$_1$ liked his$_1$ mother

(38)
There are four possibilities to reconstruct the ellipsis, while obeying parallelism requirement. Let us step through the four possibilities to see that our analysis solves this puzzle. In the first case, as given in (39), the two elided pronouns are linked to their parallel pronouns respectively. Both of them have the same referential value as their source pronouns. Therefore, they refer to ‘John’, yielding the strict-strict reading.

\begin{align*}
\text{John}_1 \text{ said } &\text{he}_1 \text{ liked his}_1 \text{ mother.} \\
\text{(39)} \\
\text{Bill}_2 \text{ said } &\text{he}_1 \text{ liked his}_1 \text{ mother.}
\end{align*}

In the second possibility (40), the link for the first elided pronoun is gained according to the referential parallelism while the link for the second elided pronoun is received by obeying the structural parallelism. In this case, the first elided pronoun refers to the first source pronoun which refers to ‘John’, yielding the strict interpretation. The second elided pronoun is linked to the first elided pronoun and refers to ‘John’ as well. (40) corresponds to the strict-strict reading. (39) and (40) together show that the second elided pronoun must have a strict reading, if the first elided pronoun has a strict reading.

\begin{align*}
\text{John}_1 \text{ said } &\text{he}_1 \text{ liked his}_1 \text{ mother.} \\
\text{(40)} \\
\text{Bill}_2 \text{ said } &\text{he}_1 \text{ liked his}_1 \text{ mother.}
\end{align*}

In the third case (41) the first elided pronoun is linked to the target parallel element to receive a sloppy interpretation whereas the second elided pronoun is linked to the second source pronoun to get a strict interpretation, resulting in the sloppy-strict reading.

\begin{align*}
\text{John}_1 \text{ said } &\text{he}_1 \text{ liked his}_1 \text{ mother.} \\
\text{(41)} \\
\text{Bill}_2 \text{ said } &\text{he}_2 \text{ liked his}_1 \text{ mother.}
\end{align*}

In the fourth case (42) both elided pronouns are linked to elements in the target clause to establish the structural parallelism, yielding the sloppy-sloppy reading in the target clause. (42) shows that the sloppy interpretation of the second elided pronoun is dependent on the sloppy interpretation of the first elided pronoun.

\begin{align*}
\text{John}_1 \text{ said } &\text{he}_1 \text{ liked his}_1 \text{ mother.} \\
\text{(42)} \\
\text{Bill}_2 \text{ said } &\text{he}_2 \text{ liked his}_2 \text{ mother.}
\end{align*}

The linking chain in the source clause does not allow the combination of links in the target clause in (43) since the link between the second elided pronoun and the target parallel element
does not have a parallel link in the source clause. Thus, the link does not obey the parallelism theorem. Hence it follows that (43), which corresponds to the reading (37-4), is ruled out in our analysis.

(43) * John_1 said he_1 liked his_2 mother.

Assuming that ‘John’ links both ‘he’ and ‘his’ as given in (44), the combination of the links allows the first target pronoun to be linked to the first source pronoun to receive the referential parallelism and the second target pronoun to be linked to the target parallel element ‘Bill’ to gain the structural parallelism. The links in (44) lead to the unwanted reading for the target clause in which Bill said John liked Bill’s mother. Since the definition of linking rules out the link from ‘his’ to ‘John’ in the source clause, the representation in (44) is undervariable in our analysis.

(44) * John_1 said he_1 liked his_2 mother.

Briefly, the four possibilities above produce the three readings in which the sloppy interpretation of the second elided pronoun is dependent on the sloppy interpretation of the first elided pronoun. With respect to the linking chain, our analysis predicts the following explanation for the missing readings in the cases of the many-pronouns-puzzle:

If there is a linking chain in the source clause formed by n links and its first element is the source parallel element, then n + 1 strict and sloppy readings can be expected for the target clause, instead of 2^n, since the sloppy interpretation of the ith (2 ≤ i ≤ n) elided pronoun, which is parallel to the ith pronoun in the chain, is dependent on the sloppy interpretation of the (i-1)th elided pronoun, which is parallel to the (i-1)th pronoun in the chain.

In the many-pronouns-puzzle, the length of the linking chain is two, yielding three readings. All of them obey the constraint that the second elided pronoun can be interpreted as sloppy only if the first elided pronoun has a sloppy interpretation. Let us consider sentence (45) which is similar to (36), but has all four readings with respect to the strict/sloppy ambiguity.

(45) John_1 said that his_1 mother liked him_1, and Bill did too.

In this example, the second pronoun is only bound by the source parallel element. There is no binding relation between it and the first pronoun, because the first pronoun does not c-command it. Therefore, it is linked by the source parallel element. The two links in the source clause do not form a linking chain (see the source representation in (46)). This means
that the second pronoun is not referentially dependent on the first pronoun. The consequence is that the sloppy interpretation of the second elided pronoun is not dependent on the sloppy interpretation of the first elided pronoun. Thus, the combination of the links in the source clause allows the links for the target clause, given in (46), where the second elided pronoun is linked to the target parallel element. (46) corresponds to the reading in which the first elided pronoun has a strict interpretation while the second elided pronoun receives a sloppy interpretation. (45) contains two linking chains with a length of one link, and each of them consists of the source parallel element as its first element. The combination of two such chains causes four readings for the target clause.

\[
\begin{array}{c}
\text{John}_1 \text{ said his}_1 \text{ mother liked him}_1. \\
\text{Bill}_2 \text{ said his}_1 \text{ mother liked his}_2.
\end{array}
\]

(46)

We claim that our analysis explains the five-reading sentence (47) as well. Our analysis predicts exactly the five wanted readings. The five readings are given in (48).

(47) John\textsubscript{1} revised his\textsubscript{1} paper before the teacher\textsubscript{2} did, and Bill\textsubscript{3} did too.

(48) 1. John revised John's paper before the teacher revised John's paper, and Bill revised John's paper before the teacher revised John's paper. (JJJJ)

2. John revised John's paper before the teacher revised John's paper, and Bill revised Bill's paper before the teacher revised John's paper. (JJBJ)

3. John revised John's paper before the teacher revised John's paper, and Bill revised Bill's paper before the teacher revised Bill's paper. (JJBB)

4. John revised John's paper before the teacher revised the Teacher's paper, and Bill revised John's paper before the teacher revised the Teacher's paper. (JTJT)

5. John revised John's paper before the teacher revised the Teacher's paper, and Bill revised Bill's paper before the teacher revised the Teacher's paper. (JTBT)

In (47), the source clause is ‘John\textsubscript{1} revised his\textsubscript{1} paper before the teacher did’. ‘John’ and ‘Bill’ are parallel elements. (47) is difficult for many analyses, because there is an interaction of VPE and anaphora in the source clause. The parallel elements in the elliptical construction of the source clause are ‘John’ and ‘the teacher’. The source clause has two readings, the strict reading (49) and the sloppy reading (50).

\[
\begin{array}{c}
\text{John}_1 \text{ revised his}_1 \text{ paper before the teacher}_2 \text{ revised his}_2 \text{ paper}.
\end{array}
\]

(49)

\[
\begin{array}{c}
\text{John}_1 \text{ revised his}_1 \text{ paper before the teacher}_2 \text{ revised his}_2 \text{ paper}.
\end{array}
\]

(50)
In the case of (49), the two links including the reconstructed one form a linking chain with a length of two. Assuming that (49) be the source representation of the whole clause, we predict that there will be three readings for the target clause in which the second elided pronoun has a sloppy interpretation only if the first elided pronoun is interpreted sloppily. In (50), there exists only one link in which the first element is the source parallel element. We predict that (50) yields two readings for the target clause based on the strict/sloppy ambiguity of the first elided pronoun. The three readings in the strict case plus the other two readings in the sloppy case add up to five readings for the whole clause.

Let us see whether our prediction is correct. First, let us check the strict reading case. Comparing (49) with the source representation of the many-pronouns-puzzle (38), we discover their similarity. Both of them contain a linking chain with a length of two. According to parallelism requirement, there are four possibilities to reconstruct the links for the target clause. The four cases are graphically represented in (51), (52), (53) and (54).

(51) John1 revised his1 paper before the teacher2 revised his1 paper
    Bill3 revised his1 paper before the teacher2 revised his1 paper
    (JJJJ)

(52) John1 revised his1 paper before the teacher2 revised his1 paper
    Bill3 revised his1 paper before the teacher2 revised his1 paper
    (JJJJ)

(53) John1 revised his1 paper before the teacher2 revised his1 paper
    Bill3 revised his3 paper before the teacher2 revised his3 paper
    (JJBB)

(54) John1 revised his1 paper before the teacher2 revised his1 paper
    Bill3 revised his3 paper before the teacher2 revised his1 paper
    (JJBJ)

As we predicted, the above four representations correspond to three readings **JJJJ, JJBB** and **JJBJ**. The other two readings are obtained on the basis of the sloppy reading of the source clause, i.e., (50). In (50), the first link contains the source parallel element ‘John’ as its first
element. Therefore, we find a strict and a sloppy reading for the elided pronoun. The second link consisting of ‘the teacher’ as its first element does not play a role in the reconstruction of the strict and sloppy readings of the target clause. Our parallelism requirement is not applicable for the second link. Therefore, the sloppy reading of the source clause results in the two readings JTJT and JTBT.

\[
\begin{align*}
(55) & \quad \text{John}_1 \text{ revised his}_1 \text{ paper before the teacher}_2 \text{ revised his}_2 \text{ paper} \\
& \quad \text{Bill}_3 \text{ revised his}_1 \text{ paper before the teacher}_2 \text{ revised his}_2 \text{ paper} \\
& \quad \text{JTJT}
\end{align*}
\]

\[
\begin{align*}
(56) & \quad \text{John}_1 \text{ revised his}_1 \text{ paper before the teacher}_2 \text{ revised his}_2 \text{ paper} \\
& \quad \text{Bill}_3 \text{ revised his}_3 \text{ paper before the teacher}_2 \text{ revised his}_2 \text{ paper} \\
& \quad \text{JTBT}
\end{align*}
\]

Since our parallelism requirement does not consider linking chains which do not contain parallel elements as its first element, our approach avoids spurious ambiguity as generated in (Kehler 1995; Hobbs and Kehler 1997). For example, if we apply the parallelism requirement to the second link too, there will be four representations for the target clause: (57), (58), (59) and (60). They correspond to the two readings JTJT and JTBT. As discussed in Chapter 2, both Kehler (1995) and Hobbs and Kehler (1997) derive four representations for the two readings based on the sloppy reading of the source clause, since they apply referential parallelism and structural parallelism to the second link too.

\[
\begin{align*}
(57) & \quad \text{John}_1 \text{ revised his}_1 \text{ paper before the teacher}_2 \text{ revised his}_2 \text{ paper} \\
& \quad \text{Bill}_3 \text{ revised his}_1 \text{ paper before the teacher}_2 \text{ revised his}_2 \text{ paper} \\
& \quad \text{JTJT}
\end{align*}
\]

\[
\begin{align*}
(58) & \quad \text{John}_1 \text{ revised his}_1 \text{ paper before the teacher}_2 \text{ revised his}_2 \text{ paper} \\
& \quad \text{Bill}_3 \text{ revised his}_1 \text{ paper before the teacher}_2 \text{ revised his}_2 \text{ paper} \\
& \quad \text{JTJT}
\end{align*}
\]
To sum up, the interaction of linking constraints and parallelism constraints causes the missing reading in the many-pronouns-puzzle. As shown above, our approach accounts for the other benchmark phenomena in the literature, such as the cascaded ellipsis and the five-reading sentence. Actually, the approach presented in this section agrees with the proposals made in (Kehler 1995; Hobbs and Kehler 1997). However, the frameworks used there cannot treat the ellipsis problem in an underspecified way and they cannot be extended straightforwardly to deal with other semantically underspecified phenomena, such as the scope ambiguities. It is an advantage of our approach that it integrates the analysis of VPE and anaphora into a general underspecification mechanism. In the next section, we will apply CLLS to two examples for the interaction of scope, anaphora and VPE and show that it provides a simple, integrated, and underspecified treatment.

4.3 Applications

In our approach, the description of the semantics of a VPE is a union of the constraint sets obtained by interpreting source clause and target clause plus the parallelism constraint. Its resolution involves solving the constraints. We use dominance constraints and binding constraints for scope underspecification. Linking relations between pronouns and their antecedents are expressed by linking constraints. Parallelism constraints represent parallelism between semantic representations. The interaction of quantifier scope, anaphora, and ellipsis is modeled as the interaction of binding, linking, and parallelism constraints. In the following, we will illustrate our approach with two examples.

Interaction of Quantifiers and Verb Phrase Ellipsis

Sentence (61) has two readings, corresponding to whether the quantifying noun phrase 'every person' takes wide scope over both the source and the target clause or not.

(61) John greeted every person when Bill did.

The two readings are shown as HOL formulas in (62).
(62) 1. \( \textit{every	extunderscore person}(\lambda x. \textit{when}(\textit{greet}(\textit{john}, x), \textit{greet}(\textit{bill}, x))) \)

2. \( \text{when}(\textit{every	extunderscore person}(\lambda x. (\textit{greet}(\textit{john}, x))), \textit{every	extunderscore person}(\lambda y. (\textit{greet}(\textit{bill}, y)))) \)

The constraint (63) is an underspecified representation of (61). It contains a binding constraint \( \lambda(X_3, X_7) \). The scope of \textit{every	extunderscore person}, is still underspecified. Therefore, the parallel lambda-link structures below \( X_s \) and \( X_t \) are underspecified as well.

(63)

In CLLS, the interaction of binding constraints and parallelism constraints is defined as follows:

If binding constraints interact with a parallelism constraint \( X_s/X_{sp} \sim X_t/X_{tp} \), it is required that the parallelism constraint must be satisfied and the bindings in the two contexts (i.e., the trees immediately below \( X_s \) and \( X_t \)), must be structurally isomorphic; otherwise, if two variables (nodes labeled as \textit{var}) in identical positions are bound outside their contexts, they must be bound by the same binder.

According to the definition above, there are two solutions for (63) after closure. The first solution, with the NP taking wide scope, results when the relative scope between \( X_1 \) and \( X_4 \) is resolved in a way such that \( X_1 \) dominates \( X_4 \). The corresponding solution of the constraint is visualized in (64). The parallelism constraint is satisfied in the solution since the tree dominated by \( X_t \) is a copy of the tree dominated by \( X_s \) up to \( X_{tp} \), and the two variables in the identical positions \( X_7 \) and \( Y_7 \), are bound by the same binder \( X_3 \) outside of their contexts. (64) corresponds to the reading (62-1).
The other possible scoping is that $X_4$ dominates $X_1$. The solution is depicted in (65). The tree below $X_s$ contains the interpretation of the quantifying noun phrase. The binding between $X_3$ and $X_7$ enters into the parallelism. With respect to the bindings, the parallelism constraint holds in (65), since the binding between $X_3$ and $X_7$ is isomorphic to the binding between $Y_3$ and $Y_7$. (65) corresponding to the reading (62-2).

Interaction of Quantifiers, Anaphora, and Verb Phrase Ellipsis

In (66), both anaphora and quantification interact with verb phrase ellipsis.

(66) Mary read a book she liked before Sue did.

In (66), the source clause is ‘Mary read a book she liked’. The parallel elements are ‘Mary’ and ‘Sue’. (66) has three readings according to Gawron and Peters (1990). The first reading is that the indefinite NP ‘a book she liked’ takes wide scope over both clauses (a particular book liked by Mary is read by both Mary and Sue). The two others are characterized by having the operator ‘before’ outscope the indefinite NP. The two options result from the two possibilities of reconstructing the pronoun ‘she’ in the ellipsis interpretation, viz., strict (both read some book that Mary liked) and sloppy (each of them read some book liked by herself).

The constraint for (66), displayed in (67), is an underspecified representation of the above three readings. $X_s$ and $X_t$ represent the semantics of the source and the target clause, while
$X_{sp}$ and $X_{tp}$ stand for the semantics of the parallel elements (Mary and Sue), respectively. For readability, we represent the semantics of the complex NP 'a book she liked' by a triangle dominated by $X_2$, which only makes the anaphoric content $X_4$ of the pronoun 'she' within the NP explicit. Since 'Mary' binds 'she' most immediately in the syntax, there is a linking relation between them, represented by the linking constraint from $X_4$ to $X_{sp}$. The dominance relation between $X_1$ and $X_5$ is still underspecified.

\[(67)\]

\[\begin{array}{c}
\bullet X_0 \\
\circ X_1 \\
\circ X_3 \\
\circ X_4 \\
\circ X_5 \\
\circ X_t \\
\circ X_{sp} \\
\circ X_{tp} \\
\bullet X_7 \\
\bullet X_8 \\
\end{array}\]

The first reading, with the NP taking wide scope, results when the relative scope between $X_1$ and $X_5$ is resolved such that $X_1$ dominates $X_5$. The corresponding solution of the constraint is visualized in (68).

\[(68)\]

\[\begin{array}{c}
\bullet X_0, X_1 \\
\circ X_2 \\
\circ X_3 \\
\circ X_4 \\
\circ X_5 \\
\circ X_t \\
\circ X_{sp} \\
\circ Y_6 \\
\circ Y_8 \\
\bullet X_7 \\
\bullet X_8 \\
\end{array}\]

The parallelism constraint $X_s/X_{sp}$ $X_t/X_{tp}$ is satisfied in the solution because node $X_t$ dominates a tree that is a copy of the tree dominated by $X_s$ up to $X_{tp}$. In particular, it contains a node $Y_8$ labeled as var, which is similar to $X_8$, as it is $\lambda$-bound by $X_3$ which is outside of the two parallel contexts.

The other possible scoping is that $X_5$ dominates $X_1$. The two solutions this gives rise to are drawn in (69) and (70). Here $X_1$ and the interpretation of the indefinite NP immediately below $X_1$ enter into the parallelism as a whole, as these nodes lie below the source node $X_s$. In (69) and (70), the parallelism constraint holds since the trees below $X_s$ and $X_t$ have the same structure and the same labels up to the parallel element nodes $X_{sp}$ and $X_{tp}$, and the binding between $X_3$ and $X_7$ is isomorphic to that between $Y_3$ and $Y_7$. In addition, for the
interaction of linking and parallelism, there are two anaphoric nodes: X₄ in the source and its ‘copy’ Y₄ in the target semantics. In (69), Y₄ is linked to X₄ to have the same referential value (strict reading), and in (70), Y₄ receives a link to Xₛₚ that is structurally parallel to the link from X₄ to Xₛₚ, and hence leads to the node of the parallel element Sue (sloppy reading).

From the above examples, we have illustrated that CLLS is able to treat the interaction of quantifier scopes, anaphora and VPE in an underspecified, uniform way, without using additional means. Its coverage is comparable to Crouch (1995) and Shieber et al. (1996). However, Shieber et al. (1996) must combine two formalisms (one for scoping and one for ellipsis resolution). In addition, Shieber et al. (1996) must use additional constraints to handle the many-pronouns-puzzle and the five-reading sentence. Crouch (1995) integrates the analysis from Kehler (1993), a previous version of Kehler (1995), into the QLF formalism Alshawi and Crouch (1992). But within the framework of QLF, occurrences of semantic material can only be modeled in an indirect and less efficient way. Therefore, Crouch (1995) has to use different indices like in Kehler (1995) to express linking relations.

4.4 Extensions of the Approach

In this section, we want to show that our approach can be extended straightforwardly to examples of extended parallelism, and some other elliptical phenomena, such as stripping and gapping.
We argue that parallelism requirement for the interaction of VPE and anaphora can be generalized for all ellipsis cases.

In the present version of CLLS, the parallelism constraint \( X_s/X_{sp} \sim X_t/X_{tp} \) is restricted to one pair of exceptions \( X_{sp} \) and \( X_{tp} \). However, for extended parallelism and gapping constructions, contexts with multiple pairs of exceptions are needed. Disregarding the computational complexity problem, parallelism constraints can be generalized to have multiple pairs of exception nodes, like (71).

\[
(71) \quad X_s/(X_{sp1}, \ldots, X_{spn}) \sim X_t/(X_{tp1}, \ldots, X_{tpm})
\]

This means that the trees immediately below \( X_s \) and \( X_t \) have the same structure and labels up-to the exception nodes \( X_{sp1}, \ldots, X_{spn} \) and \( X_{tp1}, \ldots, X_{tpm} \).

**Extended Parallelism**

Kehler (1995) notes that elements involved in a sloppy reading may not be contained in the minimal clause of the ellipsis. In (72), ‘Mary likes him₁’ and ‘Susan does’ are the minimal clause pair of the ellipsis. Accounts that operate only within the minimal clauses (taking ‘Mary likes him₁’ as the source clause and ‘Susan does’ as the target clause) fail to generate the sloppy reading of the second clause, i.e., Bill said Susan likes Bill.

\[
(72) \quad \text{John₁ said that Mary likes him₁, and Bill said that Susan does too.}
\]

In our approach, we can treat (72) straightforwardly by taking the first clause as our source clause and the second as our target clause. The overt constituents in the target clause and their counterparts are treated as the parallel elements. Then the source parallel elements are ‘John’ and ‘Mary’, while the target parallel elements are ‘Bill’ and ‘Susan’. In the source clause, the pronoun ‘him’ is linked to ‘John’.

The constraint for (72) is graphically represented in (73).

\[
(73)
\]

Solving the interaction of the linking constraint link(\( X_{sp1}, X_b \)) and the parallelism constraint \( X_s/(X_{sp1}, X_{sp2}) \sim X_t/(X_{tp1}, X_{tp2}) \) results in the two solutions: (74) and (75). In the two solutions, the two trees below \( X_s \) and \( X_t \) have the same structures and labels up-to the nodes standing for the parallel elements. In (74), the anaphoric node \( Y_6 \) is linked to \( X_6 \) to obtain referential parallelism. In (75) the two links are structurally parallel, yielding the sloppy reading.
(74)

\[
\begin{array}{c}
\text{said} \bullet X_3 \quad \text{like} \bullet X_5 \quad \text{ana} \bullet X_6 \\
\text{said} \bullet X_3 \quad \text{like} \bullet Y_5 \quad \text{ana} \bullet Y_6 \\
\end{array}
\]

\[
\begin{array}{c}
\text{John} \quad X_{sp1} \quad \text{Mary} \quad X_{sp2} \\
\text{Bill} \quad X_{tp1} \\
\end{array}
\]

\[
\begin{array}{c}
\text{said} \bullet X_3 \quad \text{like} \bullet X_5 \quad \text{ana} \bullet X_6 \\
\text{said} \bullet Y_3 \quad \text{like} \bullet Y_5 \quad \text{ana} \bullet Y_6 \\
\end{array}
\]

(75)

\[
\begin{array}{c}
\text{Stripping}
\end{array}
\]

In stripping, there is only one overt constituent in the target clause without a stranded auxiliary, exemplified as (76) (Kehler 1995).

(76) John liked his mother, and Mary’s too.

Assuming that ‘his’ and ‘Mary’s’ are parallel elements in this case, although the pronoun ‘his’ is linked to ‘John’ in the source, this sentence has only one reading, i.e. (77).

(77) John liked John’s mother, and John liked Mary’s mother too.

The reason is that the pronoun ‘his’ has the overt constituent ‘Mary’s’ as its counterpart — there is no parallel elided pronoun in the target clause. Therefore, (77) does not have a sloppy reading. In DSP’s analysis, a sloppy reading is derivable. Let us show that our analysis derives only one reading.
(78) is an underspecified representation of (76). Solving the parallelism constraint, we obtain only the solution (79). The context below $X_i$ is a copy of the context below $X_s$ except the node $X_{tp}$. Since $X_{tp}$ is an exception node, it should not be similar to $X_{sp}$. Moreover, $X_{tp}$ is not an anaphoric node. Therefore, it cannot be linked to other nodes.

(79)

Gapping

In a gapping construction, the target clause contains two or more overt constituents. Sentence (80) is an example for gapping.

(80) John gave chocolate to Mary, and Fred to Peter.

The two overt constituents in the target clause are ‘Fred’ and ‘Peter’. There are different ways to determine the parallel elements in a gapping construction which leads to ambiguities. Although we do not deal with the determination of parallel elements in this work, it belongs to one of the interesting questions in ellipsis theories. For our purpose, we assume that ‘John’ and ‘Fred’ form a pair of parallel elements while ‘Mary’ and ‘Peter’ are the other pair. Thus, (80) has one reading in which Fred gave chocolate to Peter.

The constraint, illustrated in (81), is the underspecified representation for (80).
(81)

\[
\begin{array}{c}
X_0 \\
\downarrow \quad \downarrow \quad \downarrow \\
X_2 \\
\downarrow \quad \downarrow \\
X_3 \\
\downarrow \quad \downarrow \quad \downarrow \\
X_4 \quad \text{chocolate} \quad X_5 \\
\end{array}
\]

The solution, given in (82), corresponds to the wanted reading. In this case, the structure below \( X_t \) is resolved and it is a copy of the structure below \( X_s \) up to the exception nodes \( X_{tp1} \) and \( X_{tp2} \).

(82)

\[
\begin{array}{c}
X_0 \\
\downarrow \\
X_2 \\
\downarrow \\
X_3 \\
\downarrow \\
X_4 \quad \text{chocolate} \quad X_5 \\
\downarrow \\
X_6 \\
\end{array}
\]

4.5 Conclusion

In this chapter, we have presented an analysis to explain the strict/sloppy ambiguity in the cases of VPE and anaphora. The basic idea is that an elided pronoun can either be linked to its parallel pronoun in the source clause (referential parallelism) or be linked in a structurally parallel way (structural parallelism). In addition, we claim that linking chains and the parallelism requirement yield the best explanation for the distribution of the strict and sloppy readings involving many pronouns. It covers a series of problematic cases in the literature such as the many-pronouns-puzzle and the five-reading sentence. The framework of CLLS, into which our analysis is integrated, allows a simple, integrated, and underspecified treatment of scope, parallelism and anaphora. The underspecified representation and resolution of scope, anaphora and parallelism is modeled as the interaction of binding, linking and parallelism constraints and the resolution of them. Although CLLS reaches a comparable coverage to the approaches of (Shieber, Pereira, and Dalrymple 1996) and (Crouch 1995), it avoids e.g. the problems Shieber, Pereira, and Dalrymple (1996) get by combining two formalisms (for scoping and for ellipsis resolution) and allows a more simple and transparent treatment. The application of CLLS and its extensions have shown that it is well suited for the task of representing underspecified semantic information.
Chapter 5

Conclusion and Outlook

In this thesis, we have discussed the underspecified semantic representation and resolution of ellipsis. We have presented a Constraint Language for Lambda-link Structures (CLLS). The application of CLLS extends the coverage of the approach of (Niehren et al. 1997) to include the interaction of VPE and anaphora. Lambda-link structures are tree structures with additional relations to express lambda-binding and anaphoric links. CLLS expresses constraints over lambda-link structures. We use dominance constraints and binding constraints to express semantically underspecified information such as scope ambiguities. The parallelism constraints are used to express the parallelism between semantic representations. In CLLS, first order variables are interpreted as nodes rather than trees. This gives a simple means of handling anaphora in VPE. Therefore, we can integrate our linguistic analysis for the interaction of VPE and anaphora into CLLS.

We have argued that an adequate treatment of VPE and anaphora cannot be achieved by simply identifying a target VP with its source VP. The resolution process must establish a parallelism between larger units (clauses) that a VP occurs in. In particular, the semantic representation on which the resolution process operates should be able to model the linking relations between pronouns and their antecedents. In our linguistic analysis, we have claimed that the strict/sloppy ambiguity in VPE arises from the interaction of linking relations and the parallelism requirement in ellipsis. The analysis is integrated into CLLS in the following way: Linking relations between the semantic representations of pronouns and their antecedents are expressed by linking constraints; the options for anaphoric linking of the elided pronouns as expressed in the parallelism requirement are spelled out in the resolution algorithm. We have shown that our analysis explains the complex behavior of the elliptical constructions which many ellipsis analyses fail to deal with. Our analysis agrees with the proposals made in (Kehler 1995; Hobbs and Kehler 1997). But, those do not treat the ambiguities in an underspecified way and they do not include a treatment of scope ambiguities so far. It is an advantage of our approach to integrate the analysis of VPE and anaphora into a general underspecification mechanism. Compared to other accounts, CLLS allows a transparent and uniform treatment of underspecification phenomena. The interaction of scope, anaphora and ellipsis is simply modeled as the interaction of binding, dominance, linking and parallelism constraints.

Finally, we identify some points of future research suggested by the approach presented
here. The first point is that we will investigate a special type of the interaction of VPE and anaphora, namely, the interaction of VPE and reflexives (Hestvik 1993). We intend to find out the different behaviors of anaphoric pronouns and reflexives in VPE and hope to extend our analysis to deal with them. Let’s consider two simple examples of VPE with reflexives.

(1) John hit himself before Peter did.
(2) John hit himself, and Peter did too.

In (1), the reflexive gives rise both a strict and a sloppy reading for the target clause. It behaves similarly as a pronoun in VPE. However, in (2) the target clause has only a sloppy reading. It is easy to see that the distribution of the strict and sloppy readings of reflexives is different from that of the pronouns in some way.

The second point is that we attempt to work out a deduction concept for CLLS to recognize parallelism in ellipsis. As mentioned in Hobbs and Kehler (1997), the antecedent of sentences like (3), provided by Webber (1978), must be derived by using inference.

(3) Mary wants to go to Spain and Fred wants to go to Peru, but because of limited resources, only one of them can.

The third point is the question how to apply CLLS to deal with incomplete utterances occurring in spoken language (Pinkal 1996a). One possibility is that we can use dominance constraints to describe absent information.
Bibliography


Smolka, G. (1994). The definition of kernel Oz. Dfki oz documentation series, German Research Center for Artificial Intelligence (DFKI), Saarbrücken.


