Projection in discourse

Venhuizen, Noortje Joost

Publication date:
2015

Projection in Discourse

A data-driven formal semantic analysis

Noortje Joost Venhuizen

The work presented in this thesis was carried out under the auspices of the Center for Language and Cognition Groningen (CLCG) of the Faculty of Arts of the University of Groningen.

Groningen Dissertations in Linguistics 141
ISSN: 0928-0030
ISBN: 978-90-367-8071-1 (printed version)
ISBN: 978-90-367-8070-4 (electronic version)

Document prepared with LaTeX 2$_\varepsilon$ and typeset by pdfTeX (TeX Gyre Schola font)
Cover design by Noortje Joost Venhuizen and Harm Brouwer
Printed by Ridderprint BV, Ridderkerk. Printed on FSC certified paper.

rijksuniversiteit
groningen

Projection in Discourse

A data-driven formal semantic analysis

Proefschrift

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. E. Sterken
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

donderdag 26 november 2015 om 12.45 uur

door

Noortje Joost Venhuizen

geboren op 19 oktober 1986
te Groningen

Promotores
Prof. dr. ing. J. Bos
Prof. dr. P. Hendriks

Beoordelingscommissie
Prof. dr. N. M. Asher
Prof. dr. J. Hoeksema
Prof. dr. R. A. van der Sandt

Voor Harm

En voor Kees, aan wie ik dit werk
zo graag had willen laten lezen

Acknowledgements

Bovenaan het lijstje met mensen die ik wil bedanken voor hun steun en inzet tijdens de ontwikkeling van dit proefschrift staan natuurlijk mijn twee *promotores* en directe begeleiders: Johan Bos en Petra Hendriks. Johan, jij hebt mij geleerd *data-driven* te denken, en de waarde daarvan in te zien. Vandaar ook een *data-driven* aanpak tot mijn dankwoord aan jou. Eerste observatie: jouw deur staat altijd open. Tweede observatie: je bent enorm enthousiast over wat je doet, en dat werkt aanstekelijk. Derde observatie: je bent open en toegankelijk, ook buiten werktijd om (bijvoorbeeld voor een borrel, of een lift van Groningen naar Saarbrücken). Hieruit kan ik maar één conclusie trekken: heel erg bedankt voor een prettige en inspirerende samenwerking! Petra, jij bent wat later bij mijn promotietraject betrokken geraakt, maar dat doet niet af aan de invloed die je hebt gehad op de ontwikkeling van dit proefschrift en van mijzelf. Ik heb veel geleerd van jouw gedetailleerd en constructief commentaar, en heb onze uitgebreide gesprekken over 'de wetenschap en meer' altijd erg gewaardeerd. Ontzettend bedankt voor je waardevolle adviezen, ik hoop van harte dat we onze regelmatige Skype-gesprekken ook in de toekomst nog zullen voortzetten!

Next, I would like to express my gratitude to Nick Asher, Jack Hoeksema, and Rob van der Sandt, who participated in my reading committee. I highly value each of your opinions, and therefore I feel honoured that you have taken the time to read my dissertation, and have deemed it scientifically sound (@Rob: I hope that the name of the implementation "PDRT-SANDBOX" did not go unnoticed). There is a number of people from the scientific community that I have talked to about my work and that have contributed to it in different ways. In this context, I would especially like to thank Bart Geurts, Larry Horn, Hans Kamp, Andrew Kehler, and Emar Maier.

I have been very lucky to work within a close group of colleagues, who have warmed many office hours, lunches, and the not-too-occasional 'vrijdagmiddagborrel' at Hugo's jazz café. For this, I would like to thank the entire Informatiekunde department: Dieke, George, Gertjan, Gosse, Gregory, Hessel, Johan, Johannes, John, Kilian, Leonie, Malvina, Martijn, Rob, Simon, and Valerio. Special thanks go to Kilian, Valerio and Johannes, who were my direct collaborators in Johan's group, my very 'gezellige' roommates, as well as great company outside the office, of course together with their respective (to be) spouses Sara and Vivian. Furthermore, there are various colleagues outside Informatiekunde who I have really enjoyed working and/or chatting with. Thanks, Albert, Aletta, Alice, Dicky, Jennifer, John, Karin, Marijke, Mark, Meriel, and Ryanne. This last year, I have been warmly welcomed by the colleagues of the Universität des Saarlandes in Saarbrücken. Many thanks to Matt Crocker, for so generously providing me with a place in his department, and to the other colleagues there: Claudia, Elizabeth, Elli, Francesca, Heiner, Jesús, Les, Lisa, Mindaugas, Sahri, Torsten and Yoana.

Ook buiten het werk om is er een aantal mensen dat ik graag zou willen bedanken voor de nodige ontspanning en relativering tijdens de afgelopen vier jaar. Lieve Lotte en Ben, jullie zijn een baken voor mij, en het is altijd een feest om met jullie te eten, te praten, en te lachen. Lieve Merel, Lieke en Lodi, hoewel ons contact soms wat (te) schaars is, is jullie vriendschap enorm belangrijk voor mij. Lieve Bas, Emmy, Frank, Manja, Mark, Peter, Richard, en Sytze, bedankt voor het mij zo makkelijk opnemen in jullie warme vriendenclub. En ook de volgende mensen zijn op verschillende punten erg belangrijk geweest: Bedankt, John, Catharina, Anna en David, Lotte en Jeroen, Nienke, Wouter, Sanne, en Theo. Op deze plek zou ik graag mijn paranimfen speciaal willen bedanken. Kilian, ik heb het altijd heel erg prettig gevonden om met jou samen te werken: of het nou gaat om grote wetenschappelijke vraagstukken, kleine problemen oplossen, of organisatorische zaken, jouw

toewijding, zorgvuldigheid, en geduld heb ik altijd erg gewaardeerd. Niet alleen als collega, maar ook als goede vriend. Emmy, wij kennen elkaar nog niet zo heel erg lang, maar het voelt al vanaf het begin ontzettend vertrouwd om met elkaar om te gaan. Ik vind het heerlijk om zo nu en dan een kijkje te krijgen in jullie gezinsleven, en om te praten over zowel de grote als de kleine dingen in het leven. Bedankt allebei, ik ben er trots op jullie als mijn paranimfen te hebben!

Tenslotte wil ik hier mijn familie noemen, hoewel een 'bedankt' niet helemaal recht doet aan mijn gevoel ten opzichte van hen. Elly, Tonnis en Netty, betere ouders had ik mij niet kunnen wensen. Lieve broer Guus, en Karin, voor mij is ons contact ontzettend waardevol. Ik ben heel blij dat ik ook Wiebout en Richelle, en Sjoerdieke, Charles, Christopher en Simon tot mijn familie mag rekenen. Mijn skoanfamylje Oebele, Siska en Sytze, bij wie ik mij al vanaf dag één thuis heb gevoeld. En natuurlijk Harm. Uit het feit dat jij hebt meegewerkt aan een groot deel van de hoofdstukken in dit proefschrift blijkt al dat jij voor mij meer bent dan mijn man. Je bent mijn ijkpunt, mijn inspiratie, en mijn drijfveer. Je bent mijn alles.

Saarbrücken,
september 2015

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Reporter: Mr. Gandhi, what do you think of Western civilization?
Gandhi: I think that it would be a very good idea.[1]

At the risk of ruining a joke by explaining it, let us take a look at what happens here exactly. The reporter asks Gandhi for his opinion about something he takes to exist, namely 'Western civilization'. Instead of giving his opinion about Western civilization, however, Gandhi responds by putting the assumption of its existence into question. Formally, we say that the reporter's utterance *presupposes* (i.e., takes for granted) the existence of Western civilization. Gandhi's response, in turn, is a case of *presupposition cancellation*: the *presupposition* that Western civilization exists is directly contradicted.

Interestingly, the expression 'Western civilization', which gives rise to—or *triggers*—the existence presupposition, appears in a sentence that is in question form. Questions are generally considered *entailment-cancelling* constructions, which means that the regular entailments of an assertive sentence disappear when the sentence is put into question form. For instance, the question "*Is Hillary a suitable presidential candidate?*" does not entail that Hilary is in fact a presidential candidate, whereas the statement "*Hillary is a suitable presidential candidate*" does entail this. In the example above, however, the implication that Western civilization exists is not cancelled by the question form; rather, it functions as a kind of precondition for the interpretation of the question, which results in the humorous effect of Gandhi's answer. This

[1]From Beatrice Santorini: `http://www.ling.upenn.edu/~beatrice/humor/`

property of presuppositions to survive in entailment-cancelling environments is called *projection*: presuppositions behave as if they *project* out of the scope of entailment-cancelling operators, such as questions, negation, and modal operators (Langendoen and Savin, 1971).

1.1 Projection in Discourse

Presuppositions have been at the center of linguistic debate already since the works of Frege (1892) and Russell (1905), due to their non-straightforward interaction with semantic interpretation. Semantic theories have traditionally been concerned with the propositional content of language. However, language is used to communicate much more than merely propositions; it may be used to express intentions, desires, demands, assumptions, and emotions, among many other things. The formal investigation of these functions of language has often been assigned to the field of *Pragmatics*, which focuses on "the inter-relation of language structure and principles of language usage" (Levinson, 1983, p.9). However, this strict division of labor between semantics and pragmatics is challenged by the observation that 'pragmatic' features of language may interact with semantic interpretation. This is for instance the case for presuppositions: their contribution is 'pragmatic' and remains unaffected by semantic operators, but at the same time it critically affects the truth-conditional interpretation of the sentence in which it is introduced (Russell, 1905; Strawson, 1950). This means that a linguistic theory that treats presuppositions as pragmatic contributions should explain how and why this content affects semantic interpretation; and conversely, a semantic account of presuppositions should be able to account for the 'pragmatic' property of projection.

In this thesis, I will provide a formal *semantic* analysis of presuppositions as part of the broader class of *projection phenomena*, which also include anaphoric expressions (van der Sandt, 1992) and conventional implicatures (Potts, 2003, 2005). While previous approaches have mainly treated projection as a deviation from the standard way of meaning construction, I will propose a unified analysis of projection phenomena and other semantic content based on the notion of *information status*, i.e., how their contribution is related to the unfolding discourse context. This analysis will form the basis for the development of a robust formal semantic system, and will be employed in a data-driven analysis of the behavior of different types of projection phenomena as part of a larger discourse. Taken together, these results will pave the way for a more integrated empirical analysis of different semantic and pragmatic aspects of linguistic meaning.

1.2 Data-driven formal semantics

Few empirical studies exist that investigate the behavior of projection phenomena in discourse. The main reason for this is that such empirical investigations require large resources of semantically annotated texts. The work presented in this thesis aims to improve upon this situation by providing a data-driven analysis of the information status of different projection phenomena in discourse. The requirement of a large, semantically annotated resource will be met through the employment of the Groningen Meaning Bank: a corpus of public domain texts, annotated with semantic representations, which was developed as part of the "Deep Meaning Annotation" project in which I participated.

1.3 About this thesis

The main question that this thesis seeks to answer is: *How to formalize, implement, and study the behavior of projection phenomena in discourse?*—or in more specific subquestions:

(i) A wide range of linguistic phenomena has been attributed the property of projection, including presuppositions, anaphora and conventional implicatures. Can we describe these projection phenomena in a unified manner, and formulate a categorization that demonstrates their differences as well as their similarities? (Chapter 2)

(ii) Can we provide a formal semantic representation of the wide range of projection phenomena, while respecting this unified analysis? (Chapter 3)

(iii) What does the formal semantic representation say about the *how* and *why* of the projection behavior of different types of expressions? (Chapter 4)

(iv) What are the practical implications of the semantic formalism? How can it be worked out as to provide a robust, widely applicable semantic framework? (Chapter 5 & Appendix A)

(v) Robust, large-scale automatic semantic analysis requires large amounts of annotated data. How to collect human judgements about formal linguistic properties, such as projection behavior, from expert as well as non-expert linguists? (Chapter 6 & Chapter 7)

(vi) What lexical, contextual and semantic features affect the semantic behavior of projection phenomena? Can we use the annotated data to predict this behavior, based on the proposed formal analysis (and thereby improve automatic classifiers)? (Chapter 7)

This thesis is organized in three parts. The first part provides a general overview of the literature on presuppositions, anaphora and conventional implicatures, and aims to answer subquestion (i) by presenting a categorization of these different phenomena (Chapter 2). In the second part, I introduce ‘Projective Discourse Representation Theory’ (PDRT), a semantic formalism that extends Discourse Representation Theory (Kamp, 1981) by providing a uniform treatment of projection phenomena. In this part, I seek to answer subquestions (ii), (iii) and (iv): I first introduce PDRT and compare it to other analyses of presupposition projection (Chapter 3), accordingly I describe its implications in particular for the theory of conventional implicatures (Chapter 4), and finally I work out all the formal details of the framework based on an implementation of the formalism (Chapter 5, Appendix A). Part three presents a data-driven analysis of projection phenomena. I introduce the Groningen Meaning Bank corpus (Chapter 6), and aim to answer subquestions (v) and (vi) by describing different ways of data collection, and an analysis of the obtained data in terms of the formal behavior of (a subset of) the different projection phenomena (Chapter 7). The final chapter of this thesis (Chapter 8) provides the general conclusions and summarizes the conclusions of the individual parts, thereby answering the subquestions defined above. Moreover, this chapter provides directions for future work within the realms of data-driven formal semantics.

1.4 Publications

Several chapters in this thesis are adapted versions of peer-reviewed publications:

- **Chapter 3**: Venhuizen, Noortje J., Johan Bos, and Harm Brouwer (2013b). Parsimonious semantic representations with projection pointers. In Erk, Katrin and Alexander Koller, editors, *Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013) – Long Papers*, pages 252–263, Potsdam, Germany. Association for Computational Linguistics.

- **Chapter 4**: Venhuizen, Noortje J., Johan Bos, Petra Hendriks, and Harm Brouwer (2014b). How and why conventional implicatures project. In Snider, Todd, Sarah D'Antonio, and Mia Wiegand, editors, *Semantics and Linguistic Theory (SALT)*, volume 24, pages 63–83, New York, USA. LSA and CLC Publications.

Moreover, Chapter 5 and Chapter 7 are in the process of submission to peer-reviewed journals.

The work in this thesis has inspired the development of the following software package, and its associated publication (see Appendix A):

- **PDRT-SANDBOX**: A Haskell NLP library implementing DRT and PDRT. Brouwer, Harm and Noortje Venhuizen. Available at: `http://hbrouwer.github.io/pdrt-sandbox/`

- Venhuizen, Noortje and Harm Brouwer (2014). PDRT-SANDBOX: An implementation of Projective Discourse Representation Theory. In Rieser, Verena and Philippe Muller, editors, *Proceedings of the 18th Workshop on the Semantics and Pragmatics of Dialogue (DialWatt - SemDial 2014)*, pages 249–251, Edinburgh

Part I

Projection Phenomena in Discourse

Chapter 2

Projection phenomena and information status

Abstract. The property of projection, i.e., the indifference of semantic content to the syntactic scope of entailment-cancelling operators, has traditionally been associated with presuppositions. However, over the years the class of projection phenomena has been extended to incorporate other types of semantic content as well, of which anaphora and conventional implicatures are the most salient examples. This unification of different types of expressions into a single class of projection phenomena aids in the identification of the unifying constraints underlying the property of projection. At the same time, however, it obscures any differentiating properties that caused the classes to be described as separate semantic phenomena in the first place. We describe a classification of projected and asserted content that captures their commonalities as well as their differences. The hypothesized classification is based on the notion of *information status*, which describes the relation between semantic content and the discourse context in which it is introduced.

2.1 Introduction

The property of projection, i.e., the indifference of semantic content to the syntactic scope of entailment-cancelling operators, can be illustrated using the so-called "family of sentences" test (Chierchia and McConnell-Ginet, 1990):

(1)
a. The president of Bolivia is rich.
b. It is not the case that the president of Bolivia is rich.
c. If the president of Bolivia is rich, he is happy.
d. It is possible that the president of Bolivia is rich.
e. Is the president of Bolivia rich?

The sentence in (1a) has at least two implications, namely (i) that there exists an individual that is the president of Bolivia, and (ii) that this individual is rich. These implications are affected differently when embedded under sentential operators, such as negation, as in (1b), the antecedent of a conditional, as in (1c), an epistemic modal, as in (1d), or when the sentence is put in question form, as in (1e). While the implication that the president of Bolivia exists is also implied by each of (1b-e), the implication that he is rich is not; in other words, the first implication projects, while the second one does not.

Traditionally, the property of projection was considered a characteristic property of presuppositions (such as the implication triggered by "the president of Bolivia" in the example above). However, this clear dichotomy between presuppositional and asserted semantic content was challenged in two ways; firstly, the homogeneity of the class of presuppositions was disputed by various studies investigating the behavioral differences of different types of presupposition triggers (e.g., Zeevat, 2002; Spenader, 2002). Secondly, the property of projection was shown not to be exclusive to presuppositions, but could also be observed for expressions with different semantic properties than regular presuppositions. In particular, anaphoric expressions project from the same embedded contexts as presuppositions, but they pose different felicity constraints on their discourse context, that is, they need an antecedent. Similarly, Potts (2003, 2005) defined a class of expressions, which he dubbed 'conventional implicatures' (CIs; a term adapted from Grice, 1975), with contained expressions with the same projection properties as presuppositions, but that make a different contribution to the discourse; while presuppositions signal established (*pre-supposed*) information, CIs contribute novel information. The identification of this broader class of projection phenomena has inspired efforts to provide a unified analysis of projection, e.g. in terms of *at-issueness* (Simons et al., 2010). Critically, however, a unified analysis of projection phenomena demands for an explanation of the different contributions made by, for instance, presuppositions and conventional implicatures, that respects their underlying common nature.

In this chapter, we will make an inventory of the different phenomena that have been considered projection phenomena, focusing on the three major classes in the literature: presuppositions, anaphora, and conventional implicatures. We briefly discuss the different analyses that have been proposed to describe their behavior, and how these relate to each other. Accordingly, we propose a categorization of projected and asserted content in terms of their *information status*, which describes the relation between the semantic content and the discourse context. We show that the categorization makes explicit and testable predictions about the differences and similarities between the main classes of projection phenomena in terms of information status, as well as their relation to asserted content.

2.2 Projection phenomena

Presuppositions have traditionally been considered the most paradigmatic cases of projection phenomena. Recently, however, other expressions have been argued to possess this property as well (see, e.g. Simons et al., 2010; Tonhauser et al., 2013). Below, we describe the three major classes of projection phenomena found in the literature: presuppositions, anaphora, and conventional implicatures.

2.2.1 Presuppositions

The earliest descriptions of presuppositions date back to Frege (1892), who focused on definite descriptions, like 'the president of Bolivia' in the example above. He considered these presuppositions to be special conditions that must be met in order for a linguistic expression to have a denotation. On Frege's account, all cases of *presupposition failure*, i.e. falsehood of presupposition, come from a failure of reference. The implications of such failure have been famously disputed in the Russell-Strawson debate (Russell, 1905; Strawson, 1950; Russell, 1957; Strawson, 1964), illustrated by the following traditional example:

(2) The present king of France is bald.

The debate centered around the question whether this sentence could be assigned a truth value, given the observation that "the present king of France" does not exist. Russell (1905, 1957) refused to treat the existence implication of definite descriptions such as "the present king of France" as presuppositional, maintaining that a sentence containing a description that fails to refer results in a false sentence. Strawson (1950, 1964), on the other hand, argued, following Frege (1892), that sentences containing such a description lack truth value. Although it is difficult to reach a final conclusion regarding judgements about whether a sentence is meaningless or false, the observation that the implications that were identified as presupposition share the property of projection is in line with the Fregian/Strawsonian stance, where presuppositions and assertions impose different constraints on the context.

2.2.1.1 Presupposition triggers

In the vast amount of literature on presuppositions that followed Frege's formal introduction of the phenomenon, a wide variety of linguistic structures has been categorized as *presupposition triggers*. These include aspectual verbs ('to start', 'to stop') and factive verbs ('to regret', 'to know'), but also cleft sentences ('It was John who..'), iteratives ('again', 'too'), and special intonation patterns ('JOHN called Mary'). Table 2.1 shows a non-comprehensive list of lexical items and constructions that have been dubbed presupposition triggers (see also Langendoen and Savin, 1971; Soames, 1982; Beaver, 1997; Beaver and Geurts, 2011; Potts, 2013, for overviews).

This table suggests that presupposition triggers can be roughly categorized into three types on the basis of the type of presupposition they trigger: 'existence-triggers', 'prejacent-triggers', and 'precondition-triggers'. Existence-triggers are generally referential expressions that

Table 2.1: Examples of presupposition triggers

Class	Example trigger	Presupposition
Definite descriptions	*the barkeeper*	existence, uniqueness
Proper names	*John*	existence
Possessives	*his children*	existence
Quantifiers	*every*	existence
Factive verbs	*to regret, to know*	truth of prejacent
Manner adverbs	*quickly, slowly*	truth of prejacent
Cleft sentences	*It was John who..*	truth of prejacent
Temporal modifiers	*before, after*	truth of prejacent
Wh-questions	*Who called?*	truth of prejacent
Intonation patterns	*$John_F$ called Mary*	truth of prejacent
Additives	*also, only*	precondition holds
Implicature verbs	*to manage, to fail*	precondition holds
Iterative expressions	*again, too*	precondition holds
Aspectual verbs	*to stop, to start*	previous state holds
Sortally restricted predicates	*bachelor*	restriction holds

presuppose the existence of some entity, such as the definite description 'the president of Bolivia' from example (1). In addition to the existence presupposition, definite descriptions also trigger a uniqueness presupposition, indicating that the intended referent is uniquely identifiable (in the current context); i.e., the use of the definite description 'the president of Bolivia' suggests that there is only one such person. In the case of prejacent-triggers, the triggered presupposition depends on the proposition that is embedded by the trigger. For instance, the sentence 'John knows that Mary is in Paris', which contains a factive verb, presupposes that Mary is in fact in Paris. Similarly, the question 'Who called?' presupposes that someone called. The final set of triggers, which we dubbed precondition-triggers, is somewhat more heterogenous, including expressions that in some way impose a precondition, or requirement on their context. Utterances containing iterative expressions, such as 'again' and 'too', for instance, presuppose that the described event already occurred before, or that someone else fulfills the same constraint, as in 'John went to the party too', which presupposes that someone else went to the party. Sortally restricted predicates, in turn, involve presuppositions about certain sortal properties; in the case of the adjective 'bachelor', it is presupposed that the referent is male.

As mentioned above, the set of presupposition triggers described in Table 2.1 is non-comprehensive, nor is it generally agreed upon; different classifications may include more triggers, or exclude them. We will here mainly focus on what Potts (2013) calls *semantic presuppositions*, which are triggered by the conventional meaning of specific words and constructions. This means that from the list of examples in Table 2.1, the presuppositions triggered by intonational patterns are disregarded. The wide variety of presupposition triggers has inspired various diversifying approaches, which aim to classify presupposition triggers, for example according to their presuppositional strength (i.e., *weak / soft* versus *strong / hard* triggers; cf. Karttunen, 1971; Abusch, 2002, 2010), or according to differences in their projection behavior (see, e.g. Zeevat, 2002; Spenader, 2002).

2.2.1.2 The projection problem

While the property of projection successfully identifies presuppositional phenomena, it also poses new issues for theories of presupposition, as it demands an explanation about why some implications project and under which conditions this happens. Or as Beaver (2001) puts it: projection is both "the curse and the blessing of modern presupposition theory" (Beaver, 2001, p.13).

As shown in example (1), the presupposition triggered by the definite noun phrase "the president of Bolivia" projects out of the scope of entailment-cancelling operators, such as negation, modals and questions. Now consider (3), where this same implication is not projected, that is, it is not entailed by the entire sentence.

(3) If Bolivia is a republic, then the president of Bolivia is rich.

Here, we use our world-knowledge to conclude that if Bolivia is not a republic, then it doesn't have a president. So, in this case the presupposition that the president of Bolivia exists does not project from outside its scope, but it is somehow connected to the antecedent of the implication. The fact that presuppositions in some cases do, and in other cases do not project has become known as the *projection problem* of presuppositions and has inspired a wide range of analyses (see Stalnaker, 1973; Karttunen, 1974; Gazdar, 1979a,b; Soames, 1979, 1982; Heim, 1983, among many others). The particular case of presupposition cancellation in conditional statements was dubbed the *proviso problem* by Geurts (1999) and has elicited a large discussion. Some say that implications like (3) yield a conditional presupposition, like 'If Bolivia is a republic, then Bolivia has a president' (Heim, 1983; Beaver, 2001; Schlenker, 2008, 2009, 2011; Singh, 2008). Others explain the fact that the presupposition does not project in terms of pragmatic inferences about the relation between the antecedent and the presupposition, also called *bridging inferences* (van der Sandt, 1992; Geurts, 1999; Bos et al., 1995; Irmer, 2009).

The projection problem has divided the presupposition literature into two main approaches: the *satisfaction-based approach*, which treats presuppositions as setting a precondition on updating the context (see, e.g., Karttunen, 1974; Stalnaker, 1974; Heim, 1982, 1983, 1992; Beaver, 1992, 2001; Chierchia, 1995; Zeevat, 1992; Schlenker, 2007, 2008), and *binding and accommodation theory*, according to which presuppositions are anaphorically linked to previously established information (see, e.g., Soames, 1979; Kripke, 2009; van der Sandt, 1989, 1992; Geurts, 1999). Critically, the latter approach builds upon the idea that two of the three major classes of projection phenomena (namely, presuppositions and anaphora) can be described according to the same principles. We will argue that this analysis can be extended to incorporate conventional implicatures, by reformulating the theory in terms of information status. The close relation between presuppositions and anaphora, as well as the comparison to conventional implicatures, will be described in more detail below.

2.2.2 Anaphora

Anaphoric expressions, such as anaphoric pronouns, obtain their interpretation by virtue of the context in which they occur. Just like in the case of presuppositions, anaphoric dependencies are not affected by the scope of entailment-cancelling operators, i.e., the entity referred to by the anaphoric expression projects. Anaphora behave in different ways, depending on their direct linguistic context; they may serve as referring expressions that are used to co-refer with some other referential expression, called the *antecedent*, as in (4a), or they may behave like bound variables, when occurring within the scope of some logical quantifier, as illustrated in (4b) (Quine, 1960; see also King, 2013).

(4) a. John is happy because <u>he</u> passed the test.
b. Every student believes <u>he</u> will get a good grade.

In (4a), the anaphoric pronoun "he" is used to identify an individual that was introduced in the text before, namely "John". In (4b), on the other hand, the pronoun does not refer to some individual that was explicitly introduced before, but rather functions as a variable that refers to any individual that is selected by the universal quantifier.

2.2.2.1 Discourse anaphora

Despite their different functional properties, the interpretation of both anaphoric pronouns shown in example (4) can be described in terms of some kind of syntactic binding; both the proper name "John" and the quantifying noun phrase "Every student" provide an appropriate antecedent for anaphoric reference. Interestingly, however, such an analysis cannot account for the interpretations of the anaphora in (5):

(5) a. Few students passed the test. <u>They</u> were very happy.
b. If a farmer owns a donkey, he feeds <u>it</u>.

The pronoun "they" in (5a) is an example of *discourse anaphora*. In contrast to the examples in (4), the interpretation of "they" in (5a) cannot be described in terms of syntactic binding due to the fact that the pronoun and the antecedent occur in different sentences. Moreover, any analysis that can account for cross-sentential binding would result in an interpretation in which 'few students are such that they passed the test and were very happy', which is clearly not the intended interpretation of this sentence (Evans, 1977). In a similar manner, the 'donkey-sentence' in (5b), first proposed by Geach (1962), cannot be explained in terms of quantifier binding, since the existential quantification of "a donkey" cannot take scope over the pronoun "it". Linguistic examples reflecting discourse-level dependencies, as illustrated in (5a) and (5b), formed one of the major motivations for the development of *dynamic semantics*. In dynamic semantics, the meaning of an utterance is considered to be more than its truth-conditions only; rather, utterances are viewed as "instructions to update an existing context with new information, with an updated context as result" (van Eijck and Visser, 2012): in other words, meaning is 'context change potential'. Various formalisms have been developed that adhere to this view of meaning, including File-change semantics (Heim, 1982), Discourse Representation Theory (DRT; Kamp, 1981; Kamp and Reyle, 1993), Dynamic Predicate Logic (DPL; Groenendijk and Stokhof, 1991), and Update Semantics (Veltman, 1991).

Together with Heim's (1982) File-change Semantics, Kamp's (1981) DRT was one of the first fully developed dynamic semantic framework with an analysis of discourse-level anaphoric dependencies. On this account, all referential expressions introduce variables that are unselectively bound on the discourse level; this means that the variables introduced by the indefinite "a donkey" and the pronoun "it" in (5a) are both bound by the same referent, which is introduced in the restriction of the if-clause.

2.2.2.2 Presuppositions as anaphora

The DRT analysis of anaphora has been shown to be successful in accounting for different sentence- and discourse-level anaphoric dependencies (Kamp and Reyle, 1993; Kamp et al., 2011). Moreover, it has formed the basis for one of the most influential theories of presupposition: van der Sandt's (1992) analysis of 'presupposition projection as anaphora resolution' (also referred to as *binding and accommodation theory*) (van der Sandt, 1989, 1992). According to this view, presuppositions behave like anaphora in that they can anaphorically bind to an earlier introduced antecedent (see also Soames, 1979; Kripke, 2009).

This analysis is motivated by the observation that presuppositions and anaphora have very similar felicity constraints, as illustrated in (6) (adapted from Beaver and Geurts, 2011).

(6) a. If a farmer owns a donkey, he feeds it.
b. If France has a king, the king of France is bald.
c. #If a farmer doesn't own a donkey, he feeds it.
d. #If France doesn't have a king, the king of France is bald.

Both the pronoun 'it' and the presupposition that there exists a king of France can be bound dynamically by the antecedent of the conditional, as in (6a) and (6b), but are infelicitous when the existence of their antecedent is negated, as in (6c) and (6d). In order to account for these parallels, van der Sandt (1992) proposes a unified account of presuppositions and anaphora in the framework of Discourse Representation Theory (DRT; Kamp, 1981). The main idea is that, like anaphora, presuppositions can bind to (dynamically) accessible antecedents. However, unlike anaphora, they can occur felicitously in contexts where no suitable antecedent can be found. For presuppositions, failure to bind is resolved by creating their own antecedent at an accessible discourse level; this is called *accommodation*. In Chapter 3, we will describe van der Sandt's analysis in more detail, and propose an extension that will be aimed at accounting for the larger class of projection phenomena, including conventional implicatures, as described by Potts (2005).

2.2.3 Conventional implicatures

The term 'conventional implicature' was first introduced by Grice (1975), who used it to refer to implicatures that get their meaning by virtue of the conventional meaning of words. Grice (1975) took the expressions 'therefore' and 'but' to be prime examples of triggers for conventional implicatures because their usage commits the speaker to the conventional meaning of the expression (i.e. implicating consequence and contrast, respectively). This classification was countered by Bach (1999), who argued that expressions like 'but' and 'therefore' cannot trigger conventional implicatures, because their contribution cannot be separated from the main contribution of the utterance. While for Bach (1999) this reasoning meant doing away with the class of conventional implicatures as a whole, Potts (2003, 2005) agreed with Grice that some expressions trigger conventional implicatures that commit the speaker to the conventional meaning of the expression while being logically independent of what is said, but defined an entirely new class of expressions triggering such conventional implicatures (CIs, for short).

Potts (2003, 2005) describes conventional implicatures based on a set of specific criteria: *non-cancellability*, *not at-issueness*, *scopelessness*, and *speaker-orientedness*. The first two properties are shared between presuppositions and CIs, while the latter two can be considered typical to CI content. Non-cancellability means that the implication is triggered by the conventional meaning of words and therefore can not directly be denied, in contrast to, for example, conversational implicatures. At-issueness is not explicitly defined by Potts, but used "as a coverterm for regular asserted content ('what is said', in Grice's terms)" (Potts, 2005, p.24). Potts takes the property of scopelessness (as well as speaker-orientedness, see below) to distinguish CIs from presuppositions, as CIs are not sensitive to 'presupposition plugs', which prevent presuppositions from being projected (cf. Karttunen, 1973). This is illustrated in (7) (example 4.15 from Potts, 2003).

(7) a. Clinton: The damn Republicans should be less partisan.
 b. Bush: Clinton says the damn Republicans should be less partisan.

Here, the expressive attributive adjective *damn* triggers a CI that expresses disapproval about the Republicans; in (7a) this disapproval is ascribed to Clinton, who is also the speaker of the sentence. In (7b), the CI content occurs within the scope of the propositional attitude verb *says* (a presupposition plug), of which Clinton is the subject. Nevertheless, the CI content is attributed to Bush, i.e., the speaker of the sentence, illustrating that the CI content necessarily projects to the highest possible context. This is closely related to the last defining property of CIs that Potts describes; speaker-orientedness, which means that the speaker of a sentence containing an (embedded) CI is always committed to the CI-content. In fact, Amaral et al. (2007) argue that scopelessness and speaker-orientedness stem from a single characteristic that can be described as “a species of indexicality: CIs are generally anchored to the point of view of the speaker” (Amaral et al., 2007, p.733). Similarly, Simons et al. (2010) have argued that the former two properties (non-cancellability and not at-issueness) are also highly related; according to their analysis, the property of not at-issueness is what causes presuppositions and CIs to project. Their distinction between at-issue and not at-issue content (similar to the distinction made by Geurts, 2010, between foregrounded and backgrounded material) is based on the concept of the Question Under Discussion (QUD; Roberts, 1996). The QUD is a set of alternative propositions that represent the topic of the discourse. The goal of the discourse is to resolve this question, and felicitous conversational moves are taken to be those that address the QUD. According to Simons et al. (2010), the projection behavior of presuppositions, anaphora, and CIs can be explained by the observation that these expressions do not address the at-issue content of the utterance in which they occur (i.e., they are ‘not at-issue’). Since operators such as modals and negation typically target at-issue content, not at-issue content remains unaffected by entailment-cancelling operators, which explains their projection behavior.

Potts identifies two classes of CI triggers: supplemental expressions and expressives. Supplemental expressions are again separated into two subclasses: supplemental clauses, and supplemental adverbs. The class of expressives includes expressive attributive adjectives, epithets and honorifics. The next examples show three different CI triggers: an appositive construction in (8a) and a supplementary relative in (8b) (both supplemental clauses), a supplemental adverb in (8c) and an expressive attributive adjective in (8d) (examples taken from Potts, 2003).

(8) a. Ames, <u>a successful spy,</u> is now behind bars.
   b. Ames, <u>who was a successful spy,</u> is now behind bars.
   c. <u>Thoughtfully,</u> Ed destroyed the evidence for us.
   d. Sue’s dog is really <u>fucking</u> mean.

In order to account for the contribution made by CI content, Potts proposes a multi-dimensional semantic analysis in which the CI contribution and the at-issue contribution are analyzed in separate dimensions of meaning. As we will see in Chapter 4, however, van der Sandt's (1992) analysis of presuppositions as anaphora (formalized in DRT) can be extended to account for conventional implicatures as well, without introducing multiple meaning dimensions. This analysis centers around the idea that CIs, presuppositions, anaphora, and asserted content do not differ in terms of the *type* of content that is contributed, but rather in terms of *how* their content is related to the unfolding discourse context. That is, the different contributions of these expressions can be explained within a single meaning dimension, based on the observation that their content differs in *information status*. In the next section, we will describe the differences in information status in more detail, and propose a classification of semantic content based on this distinction.

2.3 The information status of projected content

In the linguistic literature, information status has mostly been used to describe differences between referential expressions (Prince, 1981; Ariel, 1988; Gundel et al., 1993). In its simplest form, the information status of a referential expression describes whether the entity that it refers to has already been introduced in the discourse (in which case it is *given*), or if the entity is *new*. Here, we will argue that information status can be used to describe the differences between the contributions made by presuppositions, anaphora and conventional implicatures, as well as their similarities, and their relation to asserted content. More specifically, we will provide a classification in terms of information status that (i) accounts for the differences as well as the similarities between different classes of projection phenomena, (ii) does not introduce an inherent distinction between the class of projection phenomena and the class of asserted content, and (iii) results in testable predictions with respect to the semantic behavior of the different types of expressions. Before turning to this classification, we first define our notion of information status in more detail, by describing it as a relation between semantic content and the communicative context.

2.3.1 The structure of communication

In order to investigate the information status of projected content, we need a formal way of defining how information is structured in a communicative setting. Following Stalnaker (1970, 1973, 1974), we can describe a conversation as a situation in which two participants P_1 and P_2 both have their own beliefs, as well as a set of shared beliefs that functions as the *common ground* for the conversation. Critically, the two participants may have different conceptions about the propositions that they take to be part of the common ground. The *discourse context* is the subset of the common ground that is part of the current exchange – either explicitly, by previous mention, or indirectly, for example as part of the extra-linguistic context. Figure 2.1 shows an illustration of the interdependence between the beliefs of the participants of a conversation (B_1 and B_2), the actual common ground (CG), what both participants take to be the common ground (CG_1 and CG_2), and the current discourse context (DC).

When producing an utterance, the speaker's goal is for the hearer to accept the content of the utterance, so that it can be added to the current discourse context (and therefore also to the common ground). In order to provide an informative and coherent contribution, an utterance by P_1 generally contains different types of information, including information from P_1's beliefs set (B_1), as well as information from–what P_1 takes to be–the common ground (C_1). These different types of information are signalled by the use of different words or lexical constructions; for instance, definite descriptions signal that their content is available in the common ground, whereas indefinite noun phrases signal that their contribution is novel. More generally, we can say that by using a presupposition trigger, the speaker signals the information to be part of the common ground, while asserted content signals information that is new to the hearer (P_2). Critically, in case the presupposed content is not part of the common ground according to the hearer, he/she may either decide to add—i.e., *accommodate*—this information to the common ground, or he/she may not accept it, in which case the communication fails. Anaphoric expressions, in turn, signal information that the speaker assumes to be part of the current discourse context. These

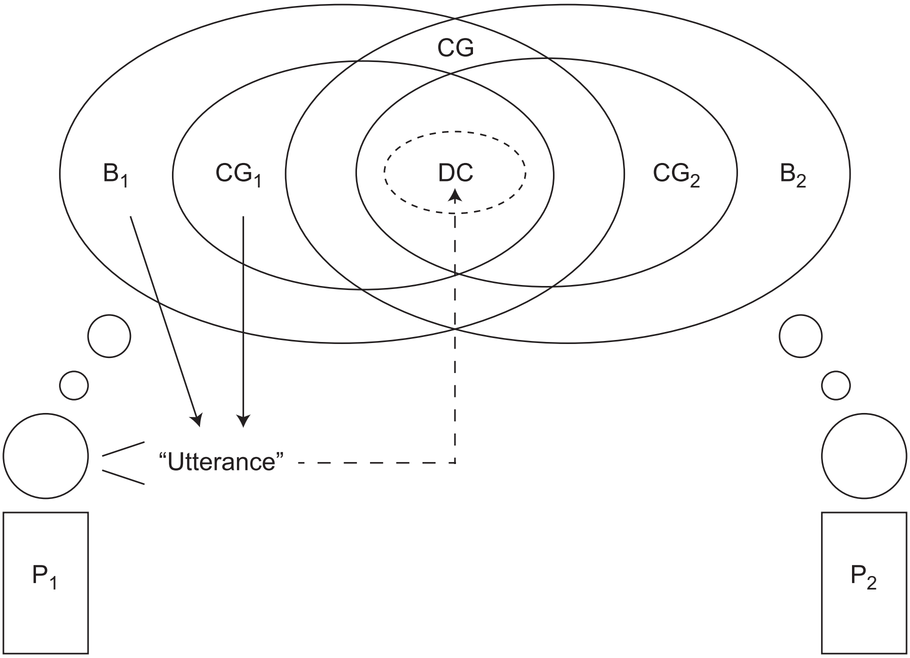


Figure 2.1: This figure shows a graphical representation of the setting of a conversation. The set of beliefs of the two participants are represented as B_1 and B_2. CG_1 and CG_2 represent the sets of beliefs that P_1 and P_2, respectively, take to be the common ground, and CG represents the actual common ground. The discourse context is represented as DC.

expressions are much harder to accommodate in case the givenness assumption fails, due to the fact that they contain little descriptive information. Finally, conventional implicatures (CIs) behave like assertions in the sense that they signal their content to be new to the hearer. In addition, however, CIs signal that their contribution is *backgrounded* (as opposed to *foregrounded*), which means that it needs to be added to the common ground, rather than to the current discourse context. In this respect, CIs behave like presuppositions, which in the case of presupposition failure are also added to the common ground.

To summarize, anaphoric expressions signal information that is *given* in the current discourse context, and are infelicitous in case this requirement is not met. Presuppositions, in turn, signal common ground-information, and can therefore be used both when the information is *given* in the current discourse context, and when it is *new*. Importantly, in case the common ground-assumption fails (presupposition failure), the presupposition is added to the common ground as *backgrounded* information (cf. Geurts, 2010). Conventional implicatures signal *new* and *backgrounded* information, i.e., information that is not part of the discourse context, and needs to be added to the common ground. In contrast, asserted content signals *new* and *foregrounded* information, which needs to be added to the current discourse context.

2.3.2 Categorizing information status

We can define a classification of semantic content based on the information structural notions 'givenness' (reflecting whether the content is *given* or *new* relative to the discourse context), and 'backgroundedness' (reflecting whether the content is *backgrounded*, i.e., needs to be added to the common ground, or *foregrounded*, i.e., needs to be added to the discourse context). That is, the classification describes the contexts in which different types of expressions can be used, based on three types of information status: (i) *given* with respect to the discourse context, (ii) *backgrounded* (i.e., new with respect to the discourse context and added to the common ground), (iii) *foregrounded* (i.e., new with respect to the discourse context and added to the discourse context).

The three types of information status effectively define three different contexts that each affect the acceptability of an expression in a distinct manner. Hence, we can determine the acceptability of different types of expressions across these different contexts. Given a binary acceptability rating and three types of information status, we predict a total of $2^3 = 8$ different categories of semantic content, as illustrated in Table 2.2 (here, the plus-sign '$+$' indicates that the expression represented in the row can be felicitously used in the context represented by the information status in the column). Based on the definition of the constraints, however, we can eliminate two of these classes in advance. Firstly, the category describing expressions for which all three types of information status are unacceptable (class E) is by definition empty,

Table 2.2: Classes of information status

Class	**Given**	**New** *backgrounded*	*foregrounded*
A	+	−	−
B	+	+	−
C	+	−	+
D	+	+	+
E	−	−	−
F	−	+	−
G	−	−	+
H	−	+	+

since the three types of information status together exhaust all the possible ways in which information can be related to the larger communicative setting (i.e., any semantic content must by definition be either given or new with respect to the discourse context, and in case it is new, it must be either foregrounded or backgrounded). The exclusion of the second category is due to the hierarchical nature of the described types of information status (cf. Gundel et al., 1993). Since *backgrounded* combines properties from both other types of information status, the three types of information status can be represented as a hierarchical ordering: *given* > *backgrounded* > *foregrounded*, where the former two both involve reference to the common ground, and the latter two both involve the contribution of novel information. As a result, the class describing expressions that are acceptable only in case their content is either *given* or *foregrounded* (class C) can be excluded.

In what follows, we describe the properties of each of the six other classes, and the type of expressions associated with these classes.

Class A: Anaphoric expressions. As described above, anaphoric expressions are only felicitous in case their content is *given* with respect to the current discourse context. Because of their low descriptive content, they are infelicitous in case this requirement is not met.

Classes B and D: Presuppositions. Presuppositions have often been differentiated based on differences in 'presuppositional strength'. In the literature on presuppositions, this notion has been described in various ways, for instance by adhering to the cancellability or deniability of presuppositions (see, e.g., Abusch, 2002, 2010; Kadmon, 2001; Von Fintel and Matthewson, 2008), or by describing different preferences for the accommodation of presuppositions (van der Sandt, 1992; Geurts, 1999; Geurts and van der Sandt, 2004). In terms of the latter account, presuppositions may differ in *where* they allow their content to be accommodated; on the *global* discourse level (i.e., in the common ground), or *locally*, as part of the discourse context. Local accommodation means that the presupposition does not project to the sentence-level, which happens in case it is cancelled or 'trapped' (van der Sandt, 1992). Critically, however, not all presuppositions allow local accommodation; proper names, for instance, have a strong urge to be projected to the common ground, and therefore are not felicitous in contexts where they are foregrounded (but see Geurts, 1997). This is illustrated in (9).

(9) a. If France is a monarchy, then <u>the king of France</u> is a rich man.
b. If France is a monarchy, then <u>Louis XIX</u> is a rich man.

In (9a), the existence presupposition triggered by the definite description "the king of France" is locally accommodated (i.e., added to the discourse context), because the existence of a king of France is conditioned on France being a monarchy, which is the restriction introduced by the implication. In contrast, (9b) does not allow for a reading in which the existence presupposition of the proper name "Louis XIX" is accommodated locally; the proper name can only be used to refer to some specific entity in the common ground, despite the fact that "Louis XIX" and "the king of France" may be used to refer to the same entity. This distinction between strong presuppositions, which cannot be *foregrounded* (i.e., locally accommodated), and weak presuppositions, which can be both *foregrounded* and *backgrounded*, is reflected in the categorization by the classes B and D.

Class F & G: Conventional implicatures and Assertions. As described above, CIs and asserted content can be distinguished in terms of the notion of 'backgroundedness': while CIs signal novel information that needs to be contributed to the common ground, i.e., *backgrounded* information, assertions signal novel information that needs to be added to the current discourse context, i.e., *foregrounded* information.

Class H: Indefinites. The categorization predicts a separate class of expressions that can introduce both *foregrounded* and *backgrounded* information, but not *given* information (class F). Interestingly, this fits the acceptability constraints for indefinite noun phrases, which can be used either *specifically*, or *non-specifically*. On their specific use, indefinites introduce a specific entity into the common ground (i.e., *backgrounded*). On their non-specific reading, on other hand, indefinites introduce a novel entity into the discourse context. Example (10) illustrates this seemingly ambiguous nature of indefinites (Kasher and Gabbay, 1976):

(10) Anna wants to marry a Swede.

In this example, the indefinite description "a Swede" can be interpreted both specifically, and non-specifically. On the non-specific reading, the sentence states that Anna wants to marry someone, and she wants this person to be a Swede. On the specific reading, it states that Anna wants to marry a certain person and that this person is a Swede. In other words, on the first reading, a novel entity is introduced in the local discourse context, i.e., inside the scope of the modal verb "to want". On the specific reading, by contrast, the entity described by the indefinite is added as a referent to the common ground, which means that it is not affected by the scope of the modal verb; i.e., it is projected. This treatment of the specific reading of indefinites as signalling projection is in line with the account proposed by van Geenhoven (1998), who describes specific indefinites as having a strong preference for projection (see also Geurts, 2010).

Final classification. The final categorization, consisting of six categories of semantic content, is shown in Table 2.3.

Table 2.3: Categorization of projected and asserted content

Class	**Type**	**Given**	**New** *backgrounded*	*foregrounded*
A	Anaphora	+	−	−
B	Strong presuppositions	+	+	−
D	Weak presuppositions	+	+	+
F	CIs	−	+	−
G	Assertions	−	−	+
H	Indefinites	−	+	+

2.3.3 Related categorizations

As described in the introduction of Section 2.3, the aim of the classification presented in Table 2.3 was to provide a unifying categorization of projected and asserted content, that makes testable predictions about the contexts in which the different types of expression can occur. This relates to a number of other categorizations that have been proposed in the literature, a large part of which focusses on discriminating between a subset of these expressions, e.g. providing a fine-grained classification of different types of presuppositions (e.g. Zeevat, 1992; Spenader, 2002; Abusch, 2010, among many others). Here, we take a more general approach, on which we aim to describe the unifying and the diversifying properties of a larger class of semantic phenomena, including both asserted and projected content. In this context, it is interesting to consider two recent classifications in more detail; the first one is proposed by Potts (2013), and the second one by Tonhauser et al. (2013).

Potts (2013): Backgrounding and Projection. The rough distinction between presuppositions, CIs and asserted content that follows from our categorisation is in line with the typology of meaning classes described by Potts (2013, p.30). Potts characterizes classes of meaning based on three descriptive properties: 'conventionality', 'backgroundedness and 'projectivity'. Potts's (2013) notion of *backgrounding* differs from ours, as it describes the property of presuppositions to be presumed 'mutual public knowledge' by the speaker. This property is very similar to our notion of *givenness*, which states that presuppositions (like anaphora) can be used to refer to established information, which may either be part of the discourse context, or part of the common ground. Potts distinguishes conventional implicatures from assertions on the basis of the property of *projection*. In our categorization, the property of projection follows from the acceptability in contexts in which information is *backgrounded* or *given*, since these both involve a direct reference to the common ground. Therefore, projection does not need to be incorporated as a separate property. In other words, the distinctions between the meaning classes proposed by Potts on the basis of *backgrounding* and *projection*, can be described using 'low-level' properties that describe how linguistic content is related to its context. Moreover, these 'low-level' properties result in a more comprehensive classification that incorporates differences between presuppositions, as well as between indefinite noun phrases and other asserted content.

Tonhauser et al. (2013): Contextual Felicity and Local Effect. On the basis of the distinction between at-issue and not at-issue content by Simons et al. (2010), (Tonhauser et al., 2013) propose a taxonomy of different projection phenomena. They categorize projection phenomena according to two properties: 'Strong Contextual Felicity' and 'Obligatory Local Effect'. If a projective expression m has the property of *Strong Contextual Felicity*, this means that it can only occur in contexts in which m is entailed. In our categorization described above, this means that the expression can only occur in contexts in which its content is *given*. This property allows for distinguishing between anaphoric expressions on the one hand, and presuppositions and conventional implicatures on the other. The property of *Obligatory Local Effect*, in turn, states that when embedded under a semantic operator, such as epistemic operators (e.g. "believe"), modals and conditionals, the projected content is both interpreted locally, and at its projection site. According to Tonhauser et al. (2013), this property distinguishes between presuppositions and conventional implicatures, since presuppositions always contribute their content to the local context of the operator. This is illustrated in (11):

(11) a. Jane believes that Bill has stopped smoking.
b. Jane believes that Bill, who is Sue's cousin, is Sue's brother.

In (11a), the proposition that Bill used to smoke is presupposed by the use of the presupposition trigger "stop", and in (11b) the non-restrictive relative clause triggers the conventional implicature that Bill is Sue's cousin. The presupposition in (11a) is necessarily attributed to Jane's belief state (in order for Jane to believe that Bill stopped smoking, she must believe that Bill used to smoke). Moreover, the presupposition is not projected, since the sentence can be felicitously continued with a statement that directly contradicts the presupposition (e.g. "...but he's never been a smoker"). In (11b), on the other hand, the CI content is not attributed to Jane's belief state – in fact it cannot be because this would result in attributing a contradictory belief to Jane (assuming a world in which one cannot be Sue's brother and Sue's cousin at the same time). Tonhauser et al. (2013) explain the difference illustrated in (11a) and (11b) based on an inherent distinction between presuppositions and CIs; presuppositions are considered to have *Obligatory Local Effect*, while CIs do not. However, this distinction is not persistent, and seems to heavily depend on the different contexts in which the expressions are used. This is illustrated in (12), which is a variant of (11b) in which the CI is replaced for a presupposition with the same informational content as the CI.

(12) Jane believes that Sue's cousin Bill is Sue's brother.

This example has two readings, one on which the statement that Bill is Sue's cousin is considered to be part of Jane's belief state and one on which it is not; this corresponds to the '*de dicto*/*de re*' distinction (cf. Quine, 1956). On the *de dicto* reading, the sentence in (12) attributes to Jane the somewhat contradictory belief that Bill is both Sue's cousin and Sue's brother. On the *de re* reading, on the other hand, the referential expression "Sue's cousin Bill" rigidly (in the Kripkian sense) refers to the person about whom Jane believes that he is Sue's brother. So, on the latter *de re* reading the presupposition has no local effect, while on the *de dicto* reading it has. Moreover, we can state that on the *de dicto* reading the presupposition is not projected, because the utterance in (12) can be continued with a something like "... but Bill isn't family of Sue at all". So, the *de dicto* reading of (12) (as well the sentence in (11a)) merely show examples of local presupposition accommodation, illustrating that what (Tonhauser et al., 2013) call *Obligatory Local Effect* is captured in our categorization by the notion of *foregroundedness*, which indicates that presuppositions can be accommodated locally (i.e., foregrounded), while CIs cannot. Critically, since our categorization describes the differences between semantic categories in terms of acceptability ratings for contexts with different types of information status, the classification accounts for both the non-projecting reading of presuppositions illustrated in example (11a), and the projecting reading illustrated in (12).

2.4 Conclusions

Projection phenomena, which include presuppositions, anaphora, and conventional implicatures, have primarily been studied as constituting separate dimensions of linguistic meaning. However, such accounts fail to describe the commonalities underlying (the projection property of) these phenomena. In this chapter, we therefore proposed a categorization that describes the differences as well as the similarities between anaphora, (strong and weak) presuppositions, CIs, indefinites, and assertions. The categorization is based on the notion of *information status*, which describes whether linguistic content is *given*, *backgrounded*, or *foregrounded* relative to the current discourse context.

The categorization makes explicit predictions about the different contexts in which different types of expressions can occur. As such, it provides an important starting point for an empirical analysis of the behavior of these different types of meaning, which is the subject of recent work in the field of empirical semantics and pragmatics (see, e.g. Schwarz, 2015, for a collection of experimental studies of presupposition). In Chapter 7 of this thesis, we will use a corpus-based approach to validate (part of) the predictions made by the proposed categorization. At the theoretical level, a classification that unifies projective and non-projective meaning, poses a challenge for formal semantic theories, in which these phenomena are often analyzed as introducing separate dimensions of meaning. In the Chapters 3-5, we will describe an extension of Kamp's (1981) Discourse Representation Theory, called Projective Discourse Representation Theory, that incorporates this unified analysis of projective and non-projective content by means of an explicit representation of information status as part of the semantic representations.

Part II

Projective Discourse Representation Theory

Chapter 3

Toward a formal treatment of projection*

Abstract. The influential idea by van der Sandt (1992) to treat presuppositions as anaphora in the framework of Discourse Representation Theory (DRT, Kamp and Reyle, 1993) has inspired a lot of debate as well as elaborations of his account. In this chapter, we propose an extension of DRT, called Projective DRT, which adds pointers to all DRT referents and conditions, indicating their projection site. This means that projected content need not be moved from the context in which it is introduced, while it remains clearly discernible from asserted content. This approach inherits the attractive properties from van der Sandt's approach to presupposition, but precludes a two-step resolution algorithm by treating projection as variable binding, which increases compositionality and computational efficiency. The result is a flexible representational framework for a descriptive theory of projection phenomena.

3.1 Introduction

When it comes to presupposition projection, or more general 'projection phenomena', there seems to be some unpleasant friction between

*Chapter adapted from Venhuizen, Noortje J., Johan Bos, and Harm Brouwer (2013b). Parsimonious semantic representations with projection pointers. In Erk, Katrin and Alexander Koller, editors, *Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013) – Long Papers*, pages 252–263, Potsdam, Germany. Association for Computational Linguistics neat compositional approaches to discourse representation, and empirically driven theories. A case in point is Discourse Representation Theory (DRT), in which proper names are treated with a special procedure in order to account for their availability as antecedent for subsequent anaphora (Kamp and Reyle, 1993). This behavior is due to the *projective* nature of proper names, that is, their existential indifference to logical operators such as negation and conditionals. In van der Sandt's (1992) empirically-driven theory of presupposition projection, formalized in the DRT framework, this discrepancy between compositionality and empirical soundness becomes very clear: presuppositions are only resolved in a second stage of processing by moving them from an embedded context to their context of interpretation. In purely compositional accounts of DRT, on the other hand, treatment of projection phenomena is usually simply left out (Muskens, 1996).

The goal of this chapter is to investigate whether van der Sandt's idea to treat presuppositions in the same way as anaphora can be generalized to account for other projection phenomena, such as Potts's (2005) conventional implicatures, in a more compositional manner. To this purpose, we propose a representational extension of DRT, called Projective DRT (PDRT), that deals with presuppositions and other projection phenomena without moving semantic material within the representation. The approach is a simplification of Layered DRT, as proposed by Geurts and Maier (2003), since presuppositions and asserted content are treated on the same level. In PDRT, projection is represented by assigning variables ranging over DRSs, just as anaphora in dynamic frameworks are dealt with by assigning variables ranging over entities. This results in semantic representations that are close to the linguistic surface structure, while clearly distinguishing between asserted and projected content.

This chapter is organized as follows. First, a theoretical background on projection phenomena in DRT is provided, focusing on van der Sandt's (1992) approach to presuppositions. In Section 3.3 we introduce Projective DRT, describing its preliminaries and how it deals with different types of (projective) content. The interpretation of PDRT is given via a translation to standard DRT, described in Section 3.4. Finally, Section 3.5 presents the conclusion and indicates directions for future work, describing an ongoing effort to implement PDRT into a large corpus of semantically annotated texts: the Groningen Meaning Bank (Basile et al., 2012a).

3.2 Projection in Discourse

Presuppositions have a long history in the formal semantics and pragmatics literature (see, e.g., Beaver and Geurts, 2011, for an overview). In this chapter, we focus on a specific representational theory of presuppositions based on Discourse Representation Theory (DRT; Kamp and Reyle, 1993).

3.2.1 Discourse Representation Theory

The original motivation for Kamp (1981) to construct a dynamic semantic framework was twofold, as he tried to fill a gap in accounting for both anaphora resolution and temporal reference. Especially the former posed an important issue for traditional truth-conditional theories, as was illustrated by the so-called 'donkey sentences' (Geach, 1962).

(1) a. Every farmer who owns a donkey beats it.
b. If a farmer owns a donkey then he beats it.

In these sentences, the pronoun 'it' needs to be bound by the (unavailable) antecedent 'a donkey', but traditional theories predict a free variable occurrence instead (e.g., the formal representation of (1a) becomes something like: $\forall x((Farmer(x) \land \exists y(Donkey(y) \land own(x, y)) \rightarrow beats(x, y))$, where the last occurrence of the variable y is free). In the representational framework of Discourse Representation Theory (DRT, Kamp, 1981; Kamp and Reyle, 1993) this problem is resolved by means of the introduction of *discourse referents* into the representation of the discourse. Each (non-anaphoric) noun phrase introduces a new discourse referent, which becomes in turn available for the binding of anaphoric expressions. The set of discourse referents together with a set of conditions that contain information about these referents, forms a complete *Discourse Representation Structure* (DRS). These DRSs can be represented in a box representation, as illustrated in (2).

(2) John loves Mary.

x, y
John(x) Mary(y) loves(x,y)

The set of discourse referents (universe) of this DRS consists of the discourse referents x and y, and the set of conditions on these referents is *John(x), Mary(y)* and *loves(x,y)*. The truthconditions of the DRS are satisfied if there exists an individual for each discourse referent in the universe of the DRS, such that the DRS conditions with respect to these referents are satisfied. I.e. (2) is satisfied if there exist two individuals such that one is called John and the other is called Mary and the former loves the latter.

DRSs are recursive structures in that they can embed other DRSs in their conditions, which are sometimes called subDRSs. Anaphoric reference in DRT involves creating a relation between the referent introduced by the anaphoric expression and an available antecedent. Informally, we can say that available antecedents are those that are introduced in the current DRS, or in a DRS of which the current DRS is a subDRS. In the case of a logical operator such as implication, the referents in the antecedent are also available for reference in the consequent. This is illustrated in (3), where (3a) shows the DRS for a multi-sentence discourse, and (3b) shows DRT's solution for the donkey sentences in (1).

(3) a. A man walks in the park. He wistles.

x, y, z
man(x) park(y) walks_in(x,y) z=x wistles(z)

b. Every farmer who owns a donkey beats it.

<table><tr><td colspan="3"></td></tr><tr><td><table><tr><td>x, y</td></tr><tr><td>farmer(x)
donkey(y)
owns(x,y)</td></tr></table></td><td>⇒</td><td><table><tr><td>v,w</td></tr><tr><td>v=x
w=y
beats(v,w)</td></tr></table></td></tr></table>

In the case of a multi-sentence discourse the interpretation succeeds incrementally; first the DRS for the first sentence is constructed, which is then combined with the information in the second sentence. The referent that is introduced by the pronoun of the second sentence is linked to the referent already introduced by ‘a man’ in the first sentence and the content of the second sentence is added to the existing DRS, resulting in the representation shown in (3a). For the donkey sentence shown in (3b), we see that the scope problems that logical theories ran into are now resolved, because the introduction of discourse referents precludes the need for quantifiers.

So, a DRS consists of a set of discourse referents and a set of conditions and can be written as a tuple $\langle U_k, C_k \rangle$, where U_k represents the universe of the DRS and C_k the set of conditions. Definition 1 formalizes the syntax of Discourse Representation Structures (Bos, 2003).

Definition 1 (DRS syntax)**.**

1. If $\{x_1 \dots x_n\}$ is a finite set of variables, and $\{\gamma_1 \dots \gamma_m\}$ is a finite set of DRS-conditions, then the ordered pair $\langle \{x_1 \dots x_n\}, \{\gamma_1 \dots \gamma_m\} \rangle$ is a basic DRS.

2. If R is a relation symbol for an n-place predicate and $x_1 \dots x_n$ are variables, then $R(x_1, \dots, x_n)$ is a basic DRS-condition.

3. If x_1 and x_2 are variables, then $x_1 = x_2$ is a basic DRS-condition.

4. Every basic DRS-condition is a DRS condition.

5. If B is a DRS, then $\neg B$, $\Box B$, and $\Diamond B$ are DRS-conditions.

6. If B_1 and B_2 are DRSs, then $B_1 \vee B_2$, and $B_1 \Rightarrow B_2$ are DRS-conditions.

7. If x is a variable and B is a DRS, then $x : B$ is a DRS-condition.

The semantics of DRSs can be given in terms of a model-theoretic interpretation. This interpretation can be given directly (Kamp and Reyle, 1993) or via a translation into first-order logic (Muskens, 1996). This property is interesting from both a theoretical point of view and from a practical perspective, because it permits the use of efficient existing inference engines (e.g. theorem provers and model builders) developed by the automated deduction community. The translation to first-order logic (as given in Bos, 2003) is provided below.

Definition 2 (Translation $(.\,,\ .)^{fo}$ from DRSs to first-order logic)**.**

1. $(w, \langle\{x_1 \ldots x_n\}, \{\gamma_1 \ldots \gamma_m\}\rangle)^{fo} \ := \ \exists x_1 \ldots \exists x_n((w, \gamma_1)^{fo} \wedge \ldots \wedge (w, \gamma_m)^{fo})$
2. $(w, R(x_1, \ldots, x_n))^{fo} \ := \ R(w, x_1, \ldots, x_n)$
3. $(w, x_1 = x_2)^{fo} \ := \ x_1 = x_2$
4. $(w, \neg B)^{fo} \ := \ \neg(w, B)^{fo}$
5. $(w, B_1 \vee B_2)^{fo} \ := \ (w, B_1)^{fo} \vee (w, B_2)^{fo}$
6. $(w, \langle\{x_1 \ldots x_n\}, \{\gamma_1 \ldots \gamma_m\}\rangle \Rightarrow B)^{fo} \ := \ \forall x_1 \ldots \forall x_n(((w, \gamma_1)^{fo} \wedge \ldots \wedge (w, \gamma_m)^{fo}) \rightarrow (w, B)^{fo})$
7. $(w, \Diamond B)^{fo} \ := \ \exists v(R(w, v) \wedge (v, B)^{fo})$
8. $(w, \Box B)^{fo} \ := \ \forall v(R(w, v) \rightarrow (v, B)^{fo})$
9. $(w, v : B)^{fo} \ := \ (R(w, v) \wedge (v, B)^{fo})$

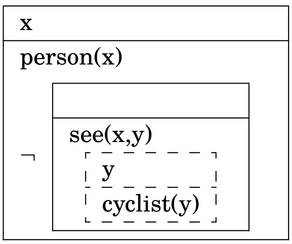
3.2.2 Presuppositions as anaphora

In the theory of van der Sandt (1992), presupposition projection is treated on a par with anaphora resolution. This approach is motivated by the observation that presuppositions and anaphora display similar behavior, since they both project their content from the scope of entailment-cancelling operators and show a preference for binding to an accessible antecedent. Unlike anaphora, however, presuppositions can occur felicitously in contexts where no suitable antecedent can be found. In these cases a new antecedent is created at an accessible discourse level, a process that has been called '*accommodation*'. The framework used by van der Sandt to implement his theory is DRT. In this account, each DRS is associated with a so-called A-structure, in which all presuppositions of that DRS are collected. In a second stage of processing, these presuppositions are resolved by either *binding* them to earlier introduced discourse referents or *accommodating* them at a suitable level of discourse. Presupposition resolution is secured by applying several constraints that determine relative preferences between alternative interpretations. These constraints include, for example, that binding is preferred over accommodation, and that global accommodation is preferred over local accommodation (see also Geurts, 1999).

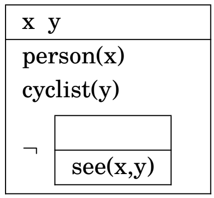
An example of the working of van der Sandt's algorithm is shown in (4) (the A-structure introduced by the presuppositional content is indicated by a dashed box). In the unresolved representation in (4a) the presupposition triggered by the definite description "the cyclist" occurs in the A-structure at the introduction site. In the second stage of processing, this A-structure is resolved by accommodating the presupposition in the global DRS, resulting in the representation shown in (4b).

(4) Someone did not see the cyclist.

a. *Unresolved DRS:*



b. *Resolved DRS:*



One of the main issues with van der Sandt's analysis of presuppositions in DRT is that, after presupposition projection, accommodated presuppositions and asserted content are indistinguishable. For example, in (4b) the accommodated presupposition "the cyclist" is added to the global context and therefore obtains the same status as the asserted content introduced by "someone". Krahmer (1998) argues, following Kracht (1994), that accommodated presuppositions should maintain their *presupposition-hood* because they are interpreted different from asserted content. For example, falsehood of a presupposition, also called *presupposition failure*, makes the sentence in which it occurs undefined (as in "The king of France is bald", where the existence of a king of France is presupposed), while in the case of falsely asserted content, the sentence is simply false (as in "France is a monarchy"). Moreover, given a compositional approach to semantics, we have to take into account that accommodated presuppositions may become bound later on, when more information of the surrounding context becomes available. This is not the case for asserted content, which implies that at each stage of processing these types of content should be distinguishable.

In order to resolve this issue, Krahmer (1998) introduces a marker for presuppositional content, such that presuppositions are accommodated at a higher discourse level *as presuppositions*, allowing for an interpretation distinct from asserted content. While this increases compositionality, the presupposition is still moved away from its introduction site in case of accommodation, which makes it difficult to retrieve the linguistic surface structure. This is problematic for applications such as surface realisation – text generation from semantic representations – and for the treatment of phenomena that depend on this surface structure, such as factive constructions and VP-ellipsis. Introducing yet another marker to identify the introduction site of a presupposition would clutter the representation and severely reduce readability and computational efficiency. Another issue with this approach is that recently the property of projection has been associated with a wider range of linguistic expressions outside of presuppositions (see Simons et al., 2010, for an overview). An important example are conventional implicatures (CIs), as described by Potts (2005). An example of a CI is shown in (5) (adapted from Potts, 2005).

(5) It is not true that Lance Armstrong, an Arkansan, won the 2002 Tour de France.

The conventional implicature triggered in the appositive (that Lance Armstrong is an Arkansan), is projected from out of the scope of the negation, just like the presupposition triggered by the proper name. However, CIs show a different projective behavior than presuppositions, since they have a strong resistance against binding to an antecedent. This is explained by the observation that they intuitively convey 'new' information, like asserted content. This preference for accommodation contrasts with the theoretical assumptions of van der Sandt (1992) and

Krahmer (1998), who implement accommodation as a repair strategy.

In sum, we need a single representational framework that allows for a separate treatment of asserted and projected content. An important step in this direction is Layered DRT (Geurts and Maier, 2003) where different types of information are treated on different layers. We will show that although this representation accounts for the differences between asserted, presupposed and conventionally implied content, it fails to capture their similarities and interactions.

3.2.3 Layered DRT

In Layered DRT (LDRT), the distinction between different types of information is implemented by introducing different layers (Geurts and Maier, 2003). Each discourse referent and condition is associated with a set of labels that indicate the layers on which the information is interpreted. These layers allow for a distinction between asserted and presupposed content, but also for a separate interpretation of implicated, indexical and formal content. An example is shown in (6), where the label p indicates presupposed content, the label a implicates asserted content and ci indicates a conventional implicature.

(6) Bill, a linguist, does not like Mary.

$$\begin{array}{|l|}\hline \text{x}_p \\ \hline \text{Bill}_p(\text{x}) \\ \text{linguist}_{ci}(\text{x}) \\ \neg_a \begin{array}{|l|}\hline \text{y}_p \\ \hline \text{Mary}_p(\text{y}) \\ \text{like}_a(\text{x,y}) \\ \hline \end{array} \\ \hline \end{array}$$

This example shows that the different types of content are represented within a single framework, while being clearly distinguishable through the labels. The different layers are connected by sharing discourse referents, indicating the interaction between different types of content. Since all conditions are indexed with a label, projected material can remain at its introduction site, because it is interpreted at a separate layer and therefore it is not targeted by logical operators. The interpretation of LDRT is defined on the basis of the truth-conditional content of sets of layers. For example, the presupposed meaning of (6) is true in the set of worlds in which the individuals called Bill and Mary exist. The asserted content can only be defined in combination with the presupposed content, representing the set of worlds in which Bill does not like Mary.

Although LDRT nicely captures the differences and dependencies between the various types of information, the separation into different layers comes at a cost. Firstly, it is unclear under which conditions a new layer is created. According to Geurts and Maier (2003, pp.15–16), all information that has a "special status" may be put on a separate layer. However, this may result in abundance of layers that all have their specific interpretation, which would fail to account for any similarities between phenomena interpreted on different layers. In particular, the similar felicity conditions for anaphora and presuppositions described by van der Sandt (1992) and the strong correspondence between asserted content and conventional implicatures (see, e.g., Amaral et al., 2007) cannot be captured in a multi-dimensional (multi-layered) framework.

Secondly, not all material seems to strictly belong to a specific layer. For example, Maier (2009b) adapts Layered DRT to account for the special behavior of proper names and indexicals, which are taken to constitute a special layer for 'reference-fixing' content (Maier calls this the *'kripke-kaplan'* or kk-layer, separating its content from the *'fregian'* fr-layer). However, some expressions, such as proper names and third person pronouns must be allowed to 'hop' between layers in order to account for their different usages (e.g., third person pronouns are regularly used in both deictic and anaphoric constructions). This solution is criticized by Hunter (2010, 2013), who argues that a relaxation of the separation between layers seems to defeat their purpose, since apparently they do not represent strictly distinct parts of meaning. Hunter provides an alternative analysis in which she shows that no extra layer is needed for indexical content; the behavior of reference-fixing expressions can be accounted for by adding an extra-linguistic context level to standard DRT, the content of which is determined by the actual state of the world. This context allows indexicals to pick out a unique object in the actual world, without the need for a separate layer of meaning.

The goal of the current chapter is to apply a similar kind of dimension reduction for projection phenomena, and to show that their behavior can be accounted for within a unidimensional framework. To this purpose, we develop Projective DRT, which extends standard DRT with a set of pointers to indicate the accommodation site of linguistic material. The framework can be seen as a refinement of Layered DRT, which integrates a subset of its layers into one and thereby accounts for the distinction, as well as the similarities between the different phenomena.

3.3 Projective Discourse Representation Theory

Projective DRT (PDRT) is an extension of standard DRT in which each referent and condition is associated with a pointer to indicate projection behavior. The basic idea of PDRT is that all projected content is represented locally, i.e., at the introduction site, and that projection is signalled by means of pointers that indicate where the content is to be interpreted. This means that projection is not realised by physically moving semantic material in the resolution stage, but by setting a variable equation on pointers and PDRS labels. This representation stays closer to the linguistic surface structure, and reduces computational complexity born out of a two-stage resolution mechanism. Moreover, presupposed and asserted (i.e. non-projected) content are clearly discernible in the representation at each step of composition, while remaining subject to the same interpretation mechanism.

3.3.1 Projection as variable binding

In PDRT, asserted and projected material is treated in the same way, by associating the content with a pointer to its context of interpretation. The differences between asserted and projected material arise from the different contexts they point to. Asserted material gets as pointer the label of the PDRS in which it is introduced, and is thus interpreted locally. In the case of projected material, the pointer may refer to the label of an accessible PDRS (in van der Sandt's terminology: a PDRS on the projection path), or it may be a free variable. As a result, projected content is interpreted in the appointed PDRS or in the global PDRS in case the pointer is a free variable. An example is shown in (7), where we use integers to denote labels and bound pointers, and f for free pointers.

(7) a. A man smiles.

$$\begin{array}{c} 1 \\ \begin{array}{|l|} \hline 1 \leftarrow \text{x} \\ \hline 1 \leftarrow \text{man(x)} \\ 1 \leftarrow \text{smile(x)} \\ \hline \end{array} \end{array}$$

b. The man smiles.

$$\begin{array}{c} 1 \\ \begin{array}{|l|} \hline f \leftarrow \text{x} \\ \hline f \leftarrow \text{man(x)} \\ 1 \leftarrow \text{smile(x)} \\ \hline \end{array} \end{array}$$

c. It is not the case that the man smiles.

$$\begin{array}{c} 1 \\ \begin{array}{|c|} \hline \\ \hline 1 \leftarrow \neg \begin{array}{c} 2 \\ \begin{array}{|l|} \hline f \leftarrow \text{x} \\ \hline f \leftarrow \text{man(x)} \\ 2 \leftarrow \text{smile(x)} \\ \hline \end{array} \end{array} \\ \hline \end{array} \end{array}$$

Each PDRS introduces a label, represented on top of the PDRS, and all referents and conditions associate with a label via a pointer, represented with an inverted arrow. If no material is projected, as in (7a), all material points to the PDRS in which it is introduced (the PDRS labeled '1'). In (7b) and (7c), on the other hand, the definite description 'the man' triggers a presupposition about the existence of its referent. In PDRT this is indicated by using a free variable as pointer for the presupposed material (here, 'f'). Free pointers are interpreted as pointing to the outermost PDRS (representing the discourse context), which both in (7b) and (7c) is the PDRS labeled '1'. As a result, the interpretations of (7a) and (7b) are equivalent, as desired, but on the representational level they are clearly distinguishable in order to account for their different compositional properties.

3.3.2 Preliminaries

The vocabulary of PDRT extends the standard DRT language with labels for DRSs and pointers for referents and conditions. A structure in PDRT (a PDRS) consists of a label ϕ, a set of projected referents D

and a set of projected conditions C, resulting in a triple: $\langle \phi, D, C \rangle$. The projected referents and conditions are defined as follows:

Definition 3 (Projected referents)**.**
If p is a pointer and d is a discourse referent, then $\langle p, d \rangle$ is a projected discourse referent.

Definition 4 (Projected conditions)**.**

- If p is a pointer and P is an n-place predicate and $u_1, \ldots, u_n$ are discourse referents, then
$\langle p, P(u_1, \ldots, u_n) \rangle$ is a projected condition.
- If p is a pointer and ϕ and ψ are PDRSs, then $\langle p, \neg\phi \rangle$, $\langle p, \phi \vee \psi \rangle$, $\langle p, \phi \rightarrow \psi \rangle$ are
projected conditions.

Furthermore, accessibility between PDRSs and free variables are defined just as in standard DRT (Kamp and Reyle, 1993). Below, when possible, we will simply refer to the referents and conditions of PDRSs, instead of projected referents and projected conditions.

In the current implementation, the semantics of a PDRS is provided via a translation to standard DRT (see Section 3.4). This is computationally advantageous because of the model-theoretic properties of standard DRT, which are interpretable via first order logic (Muskens, 1996). This means that although in PDRT the movement of projected material is precluded at the representational level, in the interpretation it will be moved in order to obtain equivalence to DRT. This way, the theory inherits some attractive properties from the DRT account to presupposition, such as its inference mechanisms and predictions with respect to, for example, the proviso problem (cf. Geurts, 1999). However, the approach can easily be adapted to incorporate other interpretative models, for example a three-valued logic to account for presupposition failure in terms of undefinedness (see, e.g., Krahmer, 1998).

3.3.3 PDRS composition

Most presuppositional theories are lexically driven, i.e., based on the assumption that specific lexical items give rise to presuppositions (so-called 'presupposition triggers'). Therefore, projected material will be manifested in the lexical semantics of projection triggers. Various authors have proposed a compositional treatment of DRT using basic tools from Montague Grammar and lambda calculus (Muskens, 1996; Bos, 2003; de Groote, 2006). Compositionality in PDRT is realised by providing every lexical item with an (unresolved) semantics in the form of a typed lambda term. In order to combine these unresolved semantics, a merge operation can be applied that combines two PDRSs into one by means of *merge-reduction* (see, e.g., Bos, 2003). In the current framework, we use different types of merge for asserted, presupposed and conventionally implied material in order to account for their different compositional properties.

In PDRT, projected material is not interpreted on a different level than asserted material, it only contributes to the context in a different way. This is realised by implementing distinct types of merge for asserted and presupposed material: assertive merge ($+$) and projective merge ($*$). Assertive merge between two (unresolved) PDRSs can be defined in the usual way by the union of the referents and conditions. Additionally, however, the pointers that refer to the merged PDRSs (i.e., the bound pointers) must be unified with the label of the resulting PDRS, in order to secure that asserted material is interpreted locally. The definition of assertive merge operations is shown below. For the renaming of pointers we use the notation '$A[x/y]$', which is taken to represent the set resulting from replacing every instance of y in the set A by x.

Definition 5 (Assertive merge)**.**

$$\overset{i}{\boxed{\begin{array}{c} D_i \\ \hline C_i \end{array}}} + \overset{j}{\boxed{\begin{array}{c} D_j \\ \hline C_j \end{array}}} := \overset{j}{\boxed{\begin{array}{c} D_i[j/i] \cup D_j \\ \hline C_i[j/i] \cup C_j \end{array}}}$$

In words, the definition for assertive merge defines the merge of two asserted PDRSs as the union of the domains and conditions of the PDRSs, with the local pointers of the PDRS in the first argument of the merge (labeled i) replaced by the label of the second argument of the merge (labeled j).

Projected material, on the other hand, is not affected by the local context, but keeps its own pointer, which either refers to its accommodation site or is a free variable. Therefore, projective merge only involves adding the projected referents and conditions to the resulting DRS, without affecting their interpretation. This results in the following definition:

Definition 6 (Projective merge)**.**

$$\overset{i}{\begin{array}{|c|}\hline D_i \\ \hline C_i \\ \hline\end{array}} \quad * \quad \overset{j}{\begin{array}{|c|}\hline D_j \\ \hline C_j \\ \hline\end{array}} \quad := \quad \overset{j}{\begin{array}{|c|}\hline D_i \cup D_j \\ \hline C_i \cup C_j \\ \hline\end{array}}$$

Conventional implicatures, in turn, exhibit yet a different type of compositional behavior (Potts, 2005). Like presuppositions, CIs project out of their local context. Unlike presuppositions, however, they cannot bind to an antecedent, nor accommodate locally (i.e., non-globally). In PDRT, this is realised by always projecting conventionally implied content to the outermost context (the “global” PDRS). This way, conventional implicatures receive an interpretation that is in some way between that of presuppositions and assertions: CIs accommodate at the highest possible context, while assertions accommodate locally and presuppositions remain free to indicate binding possibilities. In the definition for implicative merge, this means that all (bound) pointers of the conventionally implied content are replaced by a constant, say ’0’, which always refers to the outermost discourse context. This results in the following definition:

Definition 7 (Implicative merge)**.**

$$\overset{i}{\begin{array}{|c|}\hline D_i \\ \hline C_i \\ \hline\end{array}} \quad \bullet \quad \overset{j}{\begin{array}{|c|}\hline D_j \\ \hline C_j \\ \hline\end{array}} \quad := \quad \overset{j}{\begin{array}{|c|}\hline D_i[0/i] \cup D_j \\ \hline C_i[0/i] \cup C_j \\ \hline\end{array}}$$

3.3.4 Projection in PDRT

Next we will show how the different merge definitions are implemented in the lexical semantics of the linguistic material, resulting in a unified compositional framework for the representation of asserted content, presuppositions and conventional implicatures.

3.3.4.1 Asserted versus projected content

The distinction between asserted content and projected content is achieved by making use of different merge operations, reflecting the different ways in how the content is added to the discourse context. As an example, we look at the lexical semantics of definite descriptions and indefinites. In order to obtain the representations shown in (7), the indefinite should be added to the local context and the definite description should project using a free variable as pointer. This can be achieved by using different types of merge in the lexical semantics of "a" and "the". An indefinite description combines with the local context using an assertive merge, which means that the referent inherits the label from the merged PDRS and thus becomes asserted content. Definite descriptions, on the other hand, project out of their local context, which can be achieved using projective merge. The resulting lexical semantics for the determiners "a" and "the" are shown in (8).

(8) a. "a": $\lambda p.\lambda q.((\overset{i}{\begin{array}{|c|}\hline i \leftarrow x \\ \hline \\ \hline \end{array}} + p(x)) + q(x))$

b. "the": $\lambda p.\lambda q.((\overset{i}{\begin{array}{|c|}\hline i \leftarrow x \\ \hline \\ \hline \end{array}} + p(x)) * q(x))$

The lexical semantics of the indefinite article "a" introduces a discourse referent in a local PDRS. This PDRS is first combined with a predicate (e.g. a noun like "man") using assertive merge. The result of this merge operation is then combined with another predicate (e.g. a verb like "smiles"), again using *assertive* merge. This results in a representation where the indefinite description ("a man") is interpreted locally in the PDRS introduced by the rest of the context ("smiles"). For the definite article "the", on the other hand, the *projective* merge is used to combine the result of the first, assertive merge with the rest of the context. This means that the definite description keeps its own pointer, which will either be bound by an accessible PDRS, or become a free variable in the final representation, indicating accommodation.

Other presupposition triggers, such as pronouns and proper names, receive a lexical semantics similar to definite descriptions, using projective merge. In case a presupposition gets bound, the standard DRT analysis can be used, introducing an equality relation between the referent and the antecedent (Kamp and Reyle, 1993). Alternatively, we can unify the referent with the antecedent, as in van der Sandt (1992).

3.3.4.2 Conventional Implicatures

Potts (2005) defined the class of conventional implicatures on the basis of a set of specific criteria, including non- cancellability, not at-issueness, scopelessness and speaker orientation. He roughly categorizes CIs into two groups: supplemental expressions (including appositives, non-restrictive relative clauses –NRRCs– and parenthetical adverbs) and expressives (including expressive attributive adjectives, epithets and honorifics). Potts (2005) presents a multi-dimensional framework in order capture the distinction between CIs and asserted content. However, there is strong evidence against such a multi-dimensional approach, as Amaral et al. (2007) argue that there is a strong interaction between CIs and asserted content and Simons et al. (2010) unify presuppositions and CIs as projection phenomena. Therefore, in Projective DRT conventional implicatures are treated in the same way as presuppositions and asserted content, with the peculiarity that CIs always accommodate to the global discourse level. This is realised by projecting CIs using the implicative merge defined in Section 3.3.3.

Conventional implicatures are often triggered by constructions rather than lexical items, for example the subordinating constructions of appositives and NRRCs. In PDRT this is reflected by creating a special semantics for the subordinating comma, which projects its second argument. Because of the directionality of the merge operator, this means that the subordinating comma must reorder its arguments, such that the subordinated content is projected. The resulting semantics is shown in (9).

(9) subordinating comma “,”: $\lambda p.\lambda q.(q \bullet p)$

An example of the PDRT representation of an appositive is shown in (10). Note that the pointer of the appositive is ‘0’, which is a constant referring to the label of the current global context, here '1'. Thus, both the presupposition introduced by the proper name and the CI introduced by the appositive accommodate to the global discourse context. The difference is that the pointer of the presupposition (indicated with '*f*') remains available for binding, while the pointer of the appositive will always refer to the most global context.

(10) It is not true that Lance Armstrong, a former cyclist, is a Tour-winner.

1

$1 \leftarrow \neg$

2

$f \leftarrow$x

$f \leftarrow$Lance_Armstrong(x)

$0 \leftarrow$former_cyclist(x)

$2 \leftarrow$Tour-winner(x)

3.3.5 Comparison with related approaches

As described above, Layered DRT, as proposed in Geurts and Maier (2003), is a multi-dimensional framework that can account for different linguistic phenomena within a single representation. Projective DRT provides a unidimensional treatment for a subset of the phenomena covered in LDRT, including asserted content, presuppositions and conventional implicatures. The advantage of treating these different phenomena on a single 'layer' is that they are not treated as different kinds of meaning; they merely contribute their content to the context in a different way. A similar endeavour was taken by Hunter (2010, 2013), who argues for a unidimensional account of indexicals and asserted content. She proposes a DRT-style analysis in which an extra context is created for reference-fixing content, which is interpreted relative to the actual state of the world. This fits neatly within the idea of Projective DRT, where linguistic expressions are differentiated on the basis of the context the project ('point') to, and thus allows for a straightforward extension along these lines. We will leave an implementation of this and other extensions of PDRT for future work.

The account presented here is also related to the work of Schlenker (2011), who proposes a DRT account in the spirit of Heim (1983). In his representation, presupposed propositions are indexed with context variables that explicitly represent local contexts in the logical form. In this sense, his analysis is in line with approaches that use update semantics (e.g., Zeevat, 1992), because the context variable defines the context in which the presupposition is interpreted. The anaphoric aspect is therefore not in the presupposition itself, but in the context variables, which can anaphorically refer to accessible contexts. The consequence of this analysis is that accommodation does not imply adding the presuppositional content to a higher context, but rather interpreting it within this higher context. So, the interpretation of the presupposition itself, rather than that of the context in which it is accommodated is affected. In this respect Schlenker's approach crucially differs from Projective DRT, since in PDRT the traditional DRT strategy of adding presuppositions to their context of interpretation is applied. This allows for a straightforward analysis of cases of intermediate accommodation, which are difficult to capture in Schlenker's account. Moreover, PDRT allows for a more fine-grained analysis, since each referent and condition is associated with an interpretation site, while Schlenker only projects complete propositions.

3.4 Translation PDRT to DRT

The semantics of PDRSs can be described via a translation to standard DRT (Kamp and Reyle, 1993). As described above, PDRT is not strictly limited to this interpretation and may be extended to incorporate other interpretation models. We implemented PDRT as part of the wide-coverage semantic parser Boxer (Bos, 2008), including an automatic translation to standard DRT. Below we only provide a sketch of the algorithmic translation to DRT, as space limitations do not permit a description of the full translation.

3.4.1 Translation procedure

For the translation to DRT we make use of PDRT's separation of logical structure and linguistic content. Since each referent and condition is associated with a pointer to its accommodation site, it is possible to first separate this content from the embedded PDRS structure and accordingly project each condition to its appointed site. We assume that α-conversion is applied to the PDRS in order to make sure that all labels, pointers and referents use unique variables.

For convenience, we here describe the algorithm for translating PDRSs to DRSs in three steps. In the first step, all accommodation sites referents and conditions are gathered in separate sets. In the second step, the referents and conditions are added to their appointed accommodation site. In the third and final step, the PDRSs in the set of accommodation sites are combined to form a DRS.

Step 1. We start by creating three empty sets: one for accommodation sites (Π), one for discourse referents (Δ) and one for conditions (Γ). Starting from a PDRS $\Phi = \langle \varphi, D, C \rangle$, we define the pointer of Φ to be a constant: $p(\Phi) = g$, and we add this pointer, together with an empty PDRS with the label of Φ to Π: $\Pi \cup \langle p(\Phi), \langle \varphi, \{\}, \{\} \rangle \rangle$. All referents $d \in D$ are added to Δ. For the conditions $c \in C$, the base case is that c contains no embedded PDRSs, i.e., $c = \langle p, R(x_1, \ldots, x_n) \rangle$. In this case c is added to Γ. If c does contain an embedded PDRS, e.g., $c = \langle p, \neg \langle l, D_l, C_l \rangle \rangle$, then a fresh label is created, say l_0. This label is used as a sort of 'trace' to indicate where the embedded PDRS was introduced. We add $\langle l_0, \langle l, \{\}, \{\} \rangle \rangle$ to Π and $\langle p, \neg \langle l_0, \{\}, \{\} \rangle \rangle$ to Γ. This way, the context introduced by the embedded PDRS becomes available as an accommodation site, and the condition containing the embedded PDRS is added to the list of conditions. Accordingly, the referents (D_l) and conditions (C_l) of the embedded PDRS are recursively resolved in the same manner as described above, with respect to the current Δ, Γ and Π. This procedure can also be applied for other complex conditions, such as disjuctions, implications, modal expressions or propositional PDRSs (e.g., $c = \langle p, v : \langle l, D_l, C_l \rangle \rangle$).

Step 2. In this step, all referents in Δ and all conditions in Γ are projected to an appropriate PDRS in the list of accommodation sites, Π. For each referent $\langle l, u\rangle \in \Delta$, this means that if $\langle p, \langle l, D_l, C_l\rangle\rangle \in \Gamma$, then u is added to the domain: $D_l \cup u$ (so without the pointer). Otherwise, the label occurs free, so u is added to the domain of the outermost PDRS, which has g as pointer: $\langle g, \langle m, D_m \cup u, C_m\rangle\rangle$. The same strategy can be applied for conditions and the process continues until Δ and Γ are empty.

Step 3. The last step is to put the accommodation sites in Π (which now contain all the accommodated material) back together in order to form a translated DRS. We start with the DRS $\Phi = \langle D_1, C_1\rangle$, such that: $\langle g, \langle l_1, D_1, C_1\rangle\rangle \in \Pi$. This accommodation site is accordingly removed from Π. Then we check the conditions of Φ for embedded PDRSs. If such a complex condition is found, e.g. $c = \neg\langle l_c, D_c, C_c\rangle$, then the embedded PDRS is replaced by the DRS $\Psi = \langle D_m, C_m\rangle$, such that: $\langle l_c, \langle l_m, D_m, C_m\rangle\rangle \in \Pi$, which is accordingly removed from Π. Then, the set of conditions C_m of Ψ is again checked for embedded PDRSs. Once no embedded PDRSs remain, the remainder of the conditions of the dominating DRS (in this case, Φ) are checked. This recursive process goes on until Π is empty. At that point we will have a DRS with all the projected (and asserted) material at its accommodation site.

3.4.2 Example translation

We now provide an example of the translation procedure explained in the last subsection. The PDRS is shown in (11a), the desired DRS translation is shown in (11b) and its first-order logic equivalent in (11c).

(11) a.
$$\begin{array}{c} 1 \\ \begin{array}{|l|} \hline 1 \leftarrow \mathrm{x} \\ \hline f \leftarrow \mathrm{P(x)} \\ 1 \leftarrow \neg \begin{array}{c} 2 \\ \begin{array}{|l|} \hline 1 \leftarrow \mathrm{y} \\ \hline 2 \leftarrow \mathrm{Q(y)} \\ \hline \end{array} \end{array} \\ \hline \end{array} \end{array}$$

b.
$$\begin{array}{|l|} \hline \mathrm{x}\ \ \mathrm{y} \\ \hline \mathrm{P(x)} \\ \neg \begin{array}{|l|} \hline \\ \hline \mathrm{Q(y)} \\ \hline \end{array} \\ \hline \end{array}$$

c. $\exists x \exists y (P(x) \wedge \neg Q(y))$

Step 1. We start with three empty sets: Δ, Γ and Π. First, we add an empty PDRS with the label of the outermost PDRS Φ and a fixed pointer, say 0, to the set of accommodation sites: $\Pi = \{\langle 0, \langle 1, \emptyset, \emptyset\rangle\rangle\}$. We add the referents and simple conditions of Φ to the correct sets: $\Delta = \{\langle 1,\text{x}\rangle\}$; $\Gamma = \{\langle f,\text{P(x)}\rangle\}$. Then, we create a fresh label, say 3, and add an empty PDRS with the label of the embedded PDRS and the fresh label as pointer to Π: $\Pi = \{\langle 0, \langle 1, \emptyset, \emptyset\rangle\rangle, \langle 3, \langle 2, \emptyset, \emptyset\rangle\rangle\}$. The condition with the operator and an empty PDRS with the fresh label is then added to Γ: $\Gamma = \{\langle f,\text{P(x)}\rangle, \langle 1, \neg\langle 3, \emptyset, \emptyset\rangle\rangle\}$. Finally, we add the content of the embedded PDRS to the corresponding sets: $\Delta = \{\langle 1,\text{x}\rangle, \langle 1,\text{y}\rangle\}$; $\Gamma = \{\langle f,\text{P(x)}\rangle, \langle 1, \neg\langle 3, \emptyset, \emptyset\rangle\rangle, \langle 2,\text{Q(y)}\rangle\}$.

Step 2. Now, we simply project each of the elements of Δ and Γ to the appropriate PDRS in Π, i.e., to the PDRS that has the pointer of the referent/condition as label, or to the PDRS with the pointer 0, in case of a free variable: $\Pi = \{\langle 0, \langle 1, \{x, y\}, \{\text{P(x)}, \neg\langle 3, \emptyset, \emptyset\rangle\}\rangle\rangle, \langle 3, \langle 2, \emptyset, \{\text{Q(y)}\}\rangle\rangle\}$.

Step 3. Finally, we create a DRS Ψ from the accommodation site in Π that has 0 as pointer: $\Psi = \langle\{x, y\}, \{\text{P(x)}, \neg\langle 3, \emptyset, \emptyset\rangle\}\rangle$. We check for embedded PDRSs in the conditions of Ψ and replace them with the DRS from the corresponding element in Π (matching the pointer to the label). The result is the following DRS: $\langle\{x, y\}, \{\text{P(x)}, \neg\langle\emptyset, \{\text{Q(y)}\}\rangle\}\rangle$, which is exactly the desired DRS shown in (11b).

3.5 Conclusions and future work

In this chapter we presented Projective DRT, and extension of DRT in which all linguistic material is associated with a pointer to indicate its accommodation site. This way, semantic material does not need to be moved or copied at the representational level, as projection is secured by using free variables as pointers, or by binding the pointers of projected material to labels introduced by higher level PDRSs. This is in line with van der Sandt's (1992) idea to treat presuppositions as anaphora, since in DRT anaphora resolution is also based on variable binding. The theory results in a simple and parsimonious representation of different linguistic phenomena, with a unified treatment of asserted content, presuppositions and conventional implicatures. Moreover, it allows for compositional construction of discourse structures with projected content while precluding a two-step resolution algorithm. The resulting representation structures have a straightforward interpretation via translation to standard DRT.

Projective DRT can be extended to account for other phenomena, as well as other interpretation models. For example, we above mentioned a possible extension with a special context for indexical content, as described by Hunter (2010, 2013). Other directions for future work include the incorporation of phenomena such as factive constructions and VP-ellipsis with presupposed content in PDRT. A proper treatment of such phenomena may ask for an extension of the PDRT syntax (for example, allowing multiple pointers for one condition) or a more elaborate semantics that is not necessarily interpretable via a translation to standard DRT.

All in all, PDRT provides a transparent and flexible compositional framework for investigating projection phenomena. The robustness of the framework has already been put to test through an implementation into Bos's (2008) wide-coverage semantic parser: Boxer. Future work will aim at evaluating and refining the PDRSs produced by Boxer via an integration into the Groningen Meaning Bank, a large-scale corpus of semantically annotated texts (Basile et al., 2012a). PDRT allows for a coherent and easy-to-read representation of projection phenomena, since all content appears locally and the representation is therefore closer to the linguistic surface structure. This is important for a proper evaluation of semantic representations, as well as for studying the behavior of linguistic phenomena. Implementation of PDRT into a large resource of semantically annotated texts will make an important contribution to corpus-based investigations into the behavior of projection phenomena in discourse.

Chapter 4

How and why conventional implicatures project*

Abstract. Conventional Implicatures (CIs; in the sense of Potts, 2005) are part of a larger class of projection phenomena. These phenomena also include presuppositions and anaphora, and can be described as content that is not *at-issue* (see Simons et al., 2010). Despite the shared property of projection, CIs differ from other projection phenomena with respect to the information status of their contribution. Presuppositions, for instance, refer to established, or *old* information, whereas CIs contribute *novel* information to the discourse, like at-issue content. Here, we propose a unidimensional analysis of CIs and at-issue content, which highlights the similarity in projection behavior of CIs, presuppositions, and anaphora. This analysis treats CIs as 'piggybacking' on their anchor; they introduce an anaphoric dependency on the interpretation site of their anchor, while at the same time requiring their anchor to refer to a specific referent in the discourse context. CIs are thus elaborations on the description of the referent referred to by their anchor. This analysis of CIs is formalized in Projective Discourse Representation Theory (PDRT; Venhuizen et al., 2013b), a representational framework in which the property of projection is accounted for by

*Chapter adapted from: Venhuizen, Noortje J., Johan Bos, Petra Hendriks, and Harm Brouwer (2014b). How and why conventional implicatures project. In Snider, Todd, Sarah D'Antonio, and Mia Wiegand, editors, *Semantics and Linguistic Theory (SALT)*, volume 24, pages 63–83, New York, USA. LSA and CLC Publications explicitly distinguishing between the introduction and interpretation site of semantic content. Our formal analysis explains the interpretational differences between CIs, presuppositions, anaphora, and at-issue content, without stipulating a fundamental distinction between them.

4.1 Introduction

The property of *projection* has posed a challenge for semantic theories because of its apparent non-compositional nature. Projection refers to the indifference of semantic content to the syntactic scope of its embedding operators, such as negation, implication, modal operators, and interrogative constructions. Presuppositions have generally been considered the most paradigmatic class of projection phenomena. However, since Potts' redefinition of the class of Conventional Implicatures (CIs; Potts, 2003, 2005), interest has shifted toward a broader class of phenomena that project, including CIs, presuppositions, and anaphora (see, e.g., Simons et al., 2010). Critically, these projection phenomena can be differentiated from *at-issue* content, since they convey backgrounded information. On the other hand, despite their shared property of projection, presuppositions and conventional implicatures also exhibit clear differences in terms of their discourse contribution; while presuppositions signal established, or *old* information (from the speaker's perspective), CIs signal *novel* information, in the sense that by using a CI the speaker communicates to the hearer that he is introducing some additional, backgrounded information that may be new to the hearer. Consider an example of a presupposition in (1a) and a CI in (1b).

(1) a. It is not the case that John and his sister went to the party.
b. It is not the case that John, who has a sister, went to the party with her.

It follows from both (1a) and (1b) that John has a sister, despite the fact that the constructions that trigger this content syntactically occur within the scope of a negation operator. In (1a) the possessive construction *his sister* triggers a presupposition, i.e., it is taken for granted that John has a sister, while in (1b) the appositive construction signals this same piece of information to be backgrounded but *novel*. The challenge for a formal theory of projection phenomena, now, is to account for the projection behavior of presuppositions and CIs in a unified way, while appreciating their information structural differences. This is not only interesting from a theoretical perspective, but also aids practical approaches to semantic interpretation, such as Information Extraction (see Karttunen and Zaenen, 2005).

In this chapter, we describe a unified analysis of projection phenomena, formalized in the framework of Projective Discourse Representation Theory (PDRT; Venhuizen et al., 2013b). This framework was specifically developed to deal with the projection behavior of presuppositions, without introducing an extra dimension of meaning (as in Geurts and Maier, 2003, 2013), or a separate stage of processing (as in van der Sandt, 1992). Here, we will show that we can account for CIs within this framework as well, by explicitly incorporating constraints reflecting the information structure of the discourse as part of the semantic representations of PDRT. The analysis of CIs is based on the observations that (i) CIs always attach to an anchor, (ii) this anchor refers to a specific referent in the discourse context, and (iii) CIs project to the same interpretation site as their anchor.

This chapter is organized as follows: in section 4.2, we first discuss the projection behavior of CIs, and propose an analysis in which CI projection is based on discourse anchoring. In section 4.3, we then describe the semantic framework of PDRT, and show how it can account for the projection behavior of presuppositions. Next, in section 4.4, the PDRT analysis is extended to account for CIs, on the basis of the analysis presented in the previous sections. The coverage of our novel analysis is discussed in section 4.5, and section 4.6 concludes this chapter.

4.2 The projection behavior of conventional implicatures

CIs were first described by Grice (1975) as referring to implicatures that get their meaning by virtue of the conventional meaning of words. Grice took the expressions “therefore” and “but” to be prime examples of CI triggers, because their usage commits the speaker to the conventional meaning of the expression (i.e. implicating *consequence* and *contrast*, respectively). Potts (2003, 2005) agrees with Grice that some expressions trigger conventional implicatures that commit the speaker to the conventional meaning of the expression, while being logically independent of what is said. However, following Bach (1999), he rejects “but” and “therefore” as CI triggers. Instead, Potts identifies a new class of expressions that trigger CIs, which can be categorized into two main groups: *supplemental expressions*, including appositives, non-restrictive relative clauses (NRRCs) and parenthetical adverbs, and *expressives*, including expressive attributive adjectives, epithets and honorifics. Example (2) illustrates some of these CI triggers (from Potts, 2003):

(2) a. Ames, the former spy, is now behind bars. APPOSITIVE
b. Ames, who stole from the FBI, is now behind bars. NRRC
c. Ames was, as the press reported, a successful spy. AS-CLAUSE
d. Fortunately, Beck survived the descent. PARENTHETICAL ADVERB
e. Frankly (speaking), Ed fled. UTTERANCE MODIFIER
f. I hate your damn dog! EXPRESSIVE ADVERB
g. That bastard Conner got promoted. EPITHET
h. Yamada sensei - ga o - warai - ni nat - ta.
Yamada teacher - NOM HON - laugh - DAT be - PERF
‘Professor Yamada laughed.’ HONORIFIC

In each of the examples above, the speaker is committed to the contribution made by the underlined content; for example, the supplemental in (2b) conveys that Ames stole from the FBI, the expressive in (2f) conveys the speaker’s negative attitude toward the addressee’s dog, and the honorific in (2h) conveys the social status of Professor Yamada relative to the speaker. Crucially, the CI content contributes novel information that is backgrounded (i.e., not at-issue), and hence projects.

4.2.1 CIs, presuppositions, and at-issue content

Most formal analyses of CIs have focused on separating their contribution from at-issue content, e.g., by introducing separate meaning dimensions (see, e.g., Potts, 2005). Recently, however, converging evidence has emphasized a close interaction between CIs and other types of content, motivating a *unidimensional* analysis of CIs and at-issue content (see, e.g., Nouwen, 2007; Amaral et al., 2007; AnderBois et al., 2010; Schlenker, 2013; Koev, 2014). AnderBois et al. (2010), for instance, point out that the strong separation between at-issue content and CIs into different meaning dimensions is challenged by the observation that various semantic phenomena "cross the meaning boundary" between these different types of content. In particular, anaphoric dependencies, presuppositions, as well as ellipsis from the at-issue part of a sentence, can be resolved within the CI, as well as the other way around. This is illustrated in (3) (from AnderBois et al., 2010).

(3) John$_x$, who nearly killed a$_y$ woman with his$_x$ car, visited her$_y$ in the hospital.

In this example, the anaphoric expression *his* that occurs within the CI, is bound by the (presuppositional) antecedent *John*, which is introduced as part of the at-issue content. Moreover, the indefinite *a woman*, introduced in the CI, serves as the antecedent for the pronoun *her*, which occurs in the at-issue part of the sentence. This bidirectional dependency motivates a unidimensional and incremental analysis of CIs and at-issue content. Critically, such an analysis should capture the similarities between these types of content, as well as their differences.

CIs are similar to at-issue content in terms of the contribution they make to the discourse, as illustrated in the following example (from Nouwen, 2007).

(4) a. Jake, a famous Dutch boxer, lives in Utrecht.
  b. Jake lives in Utrecht. He is a famous Dutch boxer.

The appositive in (4a) and the second sentence of (4b) make a similar novel contribution to the discourse (that Jake is a famous boxer).

However, the difference between the CI contribution and the at-issue content is that the CI introduces backgrounded information. In this respect, CIs are similar to presuppositions, which by definition contribute information that is not at-issue, and therefore backgrounded. Presuppositions differ from CIs, however, in that they can *bind* to (dynamically) accessible antecedents, like anaphora (following van der Sandt, 1992). CIs do not allow for such binding, since they signal novel content. This explains why CIs cannot be cancelled, as illustrated in (5) (from Koev, 2014).

(5) a. If Betty slapped Fred, then she regrets that she slapped him.
b. ??If Obama is a socialist, then the President, who is a socialist, will raise taxes on the rich.

In (5a), the verb *regrets* triggers the factive presupposition that Betty slapped Fred, but since the content of this presupposition occurs as the antecedent of the conditional, the presupposition is cancelled, and therefore the entire sentence does not presuppose that she slapped him. In contrast, (5b) shows that a similar 'cancellation' of the conventional implicature (i.e., that the President is a socialist) renders the sentence infelicitous. The projected content of the CI matches the content of the conditional statement, just like in (5a), but the CI contribution cannot be bound by this antecedent. This is because the information that the President is a socialist is signalled to be novel by the use of the appositive construction, while this information is already contributed to the discourse by the antecedent of the conditional.

In summary, converging evidence motivates a unidimensional analysis of CIs and at-issue content. This analysis should account for the fact that both types of content contribute *novel* information, contrasting them with presuppositions. On the other hand, the analysis should also account for the fact that CIs, like presuppositions, contribute *backgrounded* information, and hence project. In what follows, we will derive such an analysis, in which CIs project because they 'piggyback' on an anchor. This piggybacking is due to the fact that CIs provide an elaboration on the referent referred to by their anchor.

4.2.2 CI projection as anchoring

CIs contribute novel information to the discourse that is not part of the at-issue content. As such, CIs place two constraints on the context in which they occur: (i) CIs need to attach their contribution to some part of the discourse, called an *anchor*, and (ii) this anchor must be backgrounded, which means that it must refer to a *specific* referent in the discourse. We motivate both of these requirements below.

(i) CIs attach to an anchor. It has been observed before that supplemental CIs systematically occur with a syntactic anchor, which allows them to connect their content to the main contribution of the discourse (e.g., Del Gobbo, 2003; Nouwen, 2007; Heringa, 2012). Similarly, an antecedent is required for an expressive CI to connect its subjective content to. This way of discourse anchoring contains a clear element of anaphoric binding. However, CIs make a different contribution to the discourse than anaphora proper; whereas anaphora are used to express identity to their antecedent, CIs contribute novel information about the anchor. In the examples in (6) (from Potts, 2003), we have highlighted the anchor of various supplemental and expressive CIs. Note that these examples only include CIs that have a nominal anchor. We will come back to the issue of non-nominal anchors in the discussion in section 4.5.

(6) a. **Edna**, a fearless leader, started the descent.
  b. **Chuck**, who killed a co-worker, is in prison.
  c. As the judge wrote, **Chuck** agreed that the verdict was fair.
  d. Every Democrat advocating **a proposal for reform** says the stupid thing is worthwhile.
  e. **Ame** ga huri - masa - ta.
   rain NOM fall - HON - PERF
   'It rained.'

In most of these sentences, the anchor can be straightforwardly identified via syntactic attachment, as in (6a) and (6b), or referential dependency, as in (6d) and (6e). Interestingly, however, the anchor in (6c) is not the entire proposition that is subordinated by the *as*-clause, but rather its subject. This is motivated by the observation that we can re-write (6c) as follows:

(6) c′. **Chuck**, about whom the judge wrote that he agreed that the verdict was fair, agreed that the verdict was fair.

The *as*-clause in (6c) makes the same contribution as the supplemental clause in (6c′), in which the complement proposition (Chuck agreed that the verdict was fair) is duplicated. Therefore, the CI in (6c) obtains the same anchor as the supplemental clause in (6c′), namely *Chuck*, the subject of the complement proposition.

(ii) The anchor of a CI refers to specific referent. In order to provide novel information that is not at-issue, CIs must attach to an anchor that is backgrounded itself. This means that the anchor must refer to a specific referent that has already been established in the current discourse context. To see why this is the case, consider example (7).

(7) a. John wants to go to a restaurant.
b. John wants to go to a restaurant, which is in the city center.

On the most intuitive reading of (7a), the indefinite *a restaurant* is interpreted non-specifically (i.e., John wants to go to some restaurant, but does not have a specific one in mind). Now, if an NRRC is attached to this indefinite, as in (7b), the non-specific reading disappears completely (i.e., John now wants to go a specific restaurant, which—by the way—is in the city center). This is because the CI triggered by the NRRC signals novel information that is not at-issue, and therefore *non-restrictive*. Any novel information that is contributed to a non-specific referent will, by contrast, be restrictive and hence at-issue, as it helps in determining the referent within the common ground. Thus, the indefinite in (7b) can only be interpreted specifically, since the supplemental clause would otherwise contribute restrictive information (as it would restrict the set of restaurants that John wants to go to). In this case, the supplemental clause would not be an NRRC and therefore it would not trigger a CI.

Now, consider the examples in (6) again. The anchors in (6a-c) are all proper names, and trigger presuppositions that project out of the local context, and hence refer to specific referents. Similarly, it follows from (6e) that the existence of rain, as an object for honoring, is presupposed. Interestingly, however, the indefinite description in (6d) is still ambiguous between a specific and a non-specific reading; the sentence can be interpreted both as referring to a particular proposal for reform (a presupposition trigger), or to *any* proposal for reform. In the latter case, the CI triggered by the epithet *the stupid thing* conveys the speaker's general characterization of Democratic proposals for reform as stupid (Potts, 2003, 21). Critically, in this case the anchor of the CI is non-specific within the global context, while it is specific with respect to the context in which the CI is introduced; the universal quantifier introduces an embedded context in which for any *specific* proposal for reform it holds that if a Democrat advocates it, he says that the stupid thing is worthwhile. In other words, within the local context of the CI trigger, the object referred to by the anchor is specific, since it is introduced in a different, accessible context (namely, in the antecedent of the conditional triggered by the use of *every*). This highlights an interesting difference between the classes of supplemental and expressive CIs. Supplemental CIs always occur within the same context as their anchors, because they syntactically attach to them. As such, the anchor of a supplemental CI must project in order to be specific with respect to the local context. This is illustrated in example (8).

(8) a. It's not the case that the king of France is bald, because he doesn't exist.
b. #It's not the case that the king of France<u>, who lives in Paris,</u> is bald, because he doesn't exist.

The existence presupposition triggered by *the king of France* cannot be cancelled in case a CI is attached to the trigger, as in (8b). In contrast, (8a) shows that the bare anchor does allow such cancellation; this is because the presupposition can be accommodated locally (i.e., within the scope of the negation).

Expressive CIs, by contrast, may attach to a non-projecting anchor, as long as it is introduced in a non-local, accessible context. In this case, the referent referred to by the anchor is at least specific with respect to the context in which the CI is introduced, and therefore the CI is felicitous.

Putting the pieces together: CIs are piggybacking on their anchor. Given that CIs attach to an anchor, which refers to a specific referent in the discourse context, we can explain the projection behavior of CIs. Since the anchor is established within the current context, the CI serves as an elaboration on the referent referred to by the anchor, and will hence be accessible in the discourse wherever the anchor is accessible. This means that the CI content is contributed directly to the context where the referent of the anchor is introduced or accommodated. The CI thus *inherits* the anchor's accommodation site; in other words, the CI is 'piggybacking' on its anchor. Here, our analysis clearly differs from the one proposed by Nouwen (2007), who treats appositives as introducing an implicit variable that is bound by the anchor. We make the stronger claim that appositives, and other CIs, not only introduce a discourse anaphoric dependency to their anchor, but also inherit its accommodation site; they are *projection-anaphoric*.

Projection-anaphoricity explains the tendency of CIs to project to the global discourse context; presuppositional anchors introducing specific referents are often discourse-new, and thus accommodated at the global discourse context (see Poesio and Vieira, 1998). Importantly, projection-anaphoricity also explains why in some cases CIs do *not* project to the global discourse level. This is illustrated in (9) (from Amaral et al., 2007).

(9) Joan is crazy. She's hallucinating that some geniuses in Silicon Valley have invented a new brain chip that's been installed in her left temporal lobe [. . .]. Joan believes that **her chip**<u>, which she had installed last month,</u> has a twelve year guarantee.

In this example, the CI triggered by the NRRC does not project to the global discourse context, since it is clear that the speaker does not want to convey that Joan *actually* had a chip installed last month. Note, however, that the direct anchor of the CI is a possessive construction (*her chip*). This construction triggers a presupposition that is bound (in the sense of van der Sandt, 1992) to the referent introduced by *a new brain chip*, which is also embedded under an epistemic verb (*hallucinate*). Thus, the anchor is not accommodated at the global discourse level, and therefore the CI that is attached to the anchor is not either. Yet, the fact that the anchor is specific within the local context, allows it to felicitously occur with a (non-restrictive) CI.

In what follows, we will show that the analysis of CIs as piggybacking on their anchor can be formalized using the framework of Projective Discourse Representation Theory (Venhuizen et al., 2013b).

4.3 Toward a formal framework of projection

Despite the prevalence of projection phenomena in discourse, semantic formalisms have mostly treated the property of projection as a deviation from standard meaning construction. Van der Sandt (1992), for instance, proposes an account of presupposition projection in terms of anaphora resolution, which he formalizes in Discourse Representation Theory (DRT; Kamp, 1981; Kamp and Reyle, 1993). On this account, projection is carried out by means of a two-stage resolution procedure, in which presuppositions are only resolved after discourse composition has been completed. In order to eliminate such a *post hoc* analysis of projection phenomena, we have recently proposed Projective Discourse Representation Theory (PDRT; Venhuizen et al., 2013b), an extension of DRT in which projection is part and parcel of standard meaning construction. PDRT is a wide-coverage semantic formalism that inherits the DRT analysis of a wide range of linguistic phenomena, including anaphora and tense (see Kamp, 1981), quantification and plurality (see Kamp and Reyle, 1993), attitude reports (see Asher, 1986; Zeevat, 1996; Maier, 2009a), and discourse structure (see Asher and Lascarides, 2003). This wide applicability is underlined by the adoption of PDRT as the formalism of choice in the Groningen Meaning Bank (GMB), a large-scale corpus of automatically-derived deep semantic representations (Basile et al., 2012a; Bos et al., 2015).

4.3.1 Projective Discourse Representation Theory

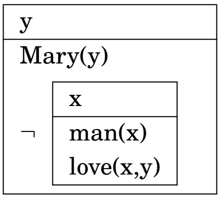
PDRT extends classic DRT by making an explicit distinction between the introduction and interpretation site of semantic content through the use of *projection variables*; all basic structures, called Projective Discourse Representation Structures (PDRSs), are associated with a *label*, and all referents and conditions are assigned a *pointer*. In a PDRS, all semantic content occurs 'in situ', thereby reflecting the linguistic surface form, and hence the introduction site of the content. The interpretation site, on the other hand, is determined by the pointers, which can be bound by the label of an accessible PDRS. Critically, the addition of projection pointers allows us to distinguish between projected and at-issue content, without introducing separate meaning dimensions (cf. Layered DRT; Geurts and Maier, 2003, 2013).

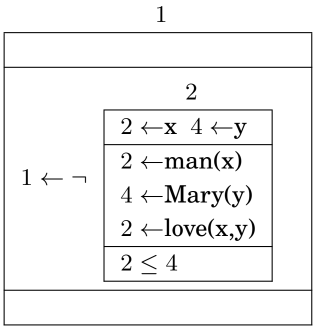
Pointers that are bound by the label of the PDRS in which the content is introduced—the *local* PDRS—indicate at-issue content. Projected content, on the other hand, is either indicated by pointers that are free variables, or by pointers that are bound by the label of some non-local, accessible PDRS. Accessibility is determined in terms of DRS subordination, following Kamp and Reyle (1993): (P)DRS k_1 is *accessible* from (P)DRS k_2 iff k_1 directly or indirectly subordinates k_2. Subordination means that k_2 is part of a condition in k_1, or k_1 serves (directly or indirectly) as the antecedent of k_2. In addition, each PDRS is enriched with a set of 'Minimally Accessible Projection contexts' (*MAPs* for short), which reflect additional constraints on the accessibility of PDRS contexts, in particular those indicated by free pointers (for a similar use of additional accessibility constraints, namely for defining unresolved DRSs, see Reyle, 1993, 1995). Together, the classic accessibility constraints and the MAPs create a partial order over PDRS contexts, which means that each projected context is accessible from at least one (sub-)PDRS that is introduced in the discourse. This assures that the discourse representation remains coherent. Note that the MAPs are an extension to the definition of PDRSs proposed in Venhuizen et al. (2013b), which will prove to be a crucial feature for the formalization of the projection behavior of CIs, described in section 4.4.

4.3.2 Presuppositions in PDRT

To see how presuppositions are represented in PDRT, consider example (10). This example shows a simple sentence containing a presupposition trigger (the proper name *Mary*), with the corresponding DRS in (10a) and PDRS in (10b). As illustrated in this example, PDRSs extend DRSs with projection variables; all PDRS contexts are associated with a label, denoted by the integer on top of the PDRS boxes, and the referents and conditions are all associated with a pointer, denoted by the integer preceding the '←' operator. Finally, the MAPs are shown at the bottom of each PDRS box.

(10) No man loves Mary.

a. 

b. 

In the PDRS in (10b), all at-issue content is associated with a pointer that is bound by the local PDRS (with label 2), meaning that it should be interpreted within that context. In contrast, the presupposition triggered by the proper name *Mary* (namely, that there exists some person named 'Mary') has a pointer that occurs free, since there is no accessible PDRS that has 4 as its label. This indicates that the presupposition still needs to be resolved to some appropriate context. The only information available in this PDRS is that the accommodation site should be accessible from the local PDRS; this is indicated by the constraint '$2 \leq 4$' in the set of MAPs, which means that context 4 is either the same as or higher than local context 2 in the accessibility chain. On top of the semantic constraints provided by the set of MAPs, pragmatic constraints may be employed to resolve the accommodation site of projected content. For instance, provided the pragmatic assumption that global accommodation is preferred over local accommodation (see van der Sandt, 1992; Geurts, 1999), we may conclude that the presupposition in (10) will be resolved to the global context (i.e., the PDRS with label 1); this results in an interpretation that is equivalent to the DRS shown in (10a). Despite this interpretational equivalence, it should be noted, however, that the PDRS provides a *richer* representation than the DRS; in contrast to the DRS resulting from van der Sandt's (1992) analysis of presupposition projection, presuppositions remain discernible from at-issue content in the PDRS (namely, on the basis of their pointer).

Moreover, in contrast to the DRS representation, the PDRS representation directly corresponds to the linguistic surface form, as the presupposition trigger *Mary* syntactically occurs within the scope of the negation. This congruity between form and meaning representation facilitates the compositional construction procedure, since no semantic content needs to be moved within the representation, as in van der Sandt's (1992) account. Instead, at-issue and projected content behave compositionally the same (in the sense of Muskens, 1996), except for their pointers; during composition, the pointers of at-issue content remain bound by the local context, whereas the pointers of projected content are free variables, which may remain unbound, or become bound later on during discourse construction (for a more elaborate description of this compositional construction process, see Venhuizen et al., 2013b, 2014a).[1]

4.4 Implementing the projection behavior of CIs

PDRT enriches DRT with an inherent treatment of projection, which has been shown to provide a parsimonious account of presupposition (Venhuizen et al., 2013b). In what follows, we will show that PDRT can

[1] We have implemented PDRT (and classic DRT) as a Haskell library, called PDRT-SANDBOX. This library incorporates machinery for representing PDRSs and DRSs, translations from PDRT to DRT and first-order logic, composition via different types of merge, and the definition of unresolved structures using Montague Semantics. PDRT-SANDBOX is available at: `http://hbrouwer.github.io/pdrt-sandbox/`.

also account for the projection behavior of conventional implicatures, by implementing the analysis of CIs as piggybacking on their anchor.

4.4.1 CI anchoring in PDRT

The MAPs of a PDRS impose accessibility constraints on projected contexts. Presuppositions pose minimal constraints on their accommodation site, requiring only that it is accessible from their introduction site. This is formalized using the *weak subordination* constraint ($p_1 \leq p_2$), which means that PDRS context p_2 is either the same as p_1 or resolved to some higher context that is accessible from p_1, thereby licencing local, global and intermediate interpretations of presuppositions (in line with van der Sandt, 1992; Geurts, 1999). Critically, MAPs can also place stronger constraints on contextual accessibility: *strict subordination* ($p_1 < p_2$) indicates that p_2 is accessible from p_1, but not identical to it, and *identity* ($p_1 = p_2$) indicates that p_1 and p_2 refer to the same projection context. We can use these latter constraints to account for the projection behavior of CIs.

CIs provide a backgrounded elaboration on the description of the referent referred to by their anchor. As such, they *depend* on the interpretation site of this referent. To capture this dependency in terms of a MAP constraint, identity is required between the interpretation site of the CI and that of its anchor. Moreover, as CIs require their anchor to be specific, they *constrain* their own interpretation site, and thereby that of their anchor, to be non-local with respect to their introduction site; this assures that the referent referred to by the anchor is established with respect to the introduction context of the CI. This constraint can be formalized by requiring the interpretation site of the CI to strictly subordinate its introduction site.

Example (11), which shows the PDRS for example (4a) above, illustrates this PDRT representation of CIs.[2] In this example, like in (10), the presupposition (triggered by the proper name *Jake*) is associated with a free pointer, which is related to its introduction context via weak subordination ($1 \leq 2$). The CI triggered by the nominal appositive (*a fa-*

[2]For reasons of brevity, the CI content and the at-issue content are represented as a single condition.

mous boxer) is also assigned a free pointer. The pointer of the at-issue content (*lives in Utrecht*), by contrast, is bound by the local context. Critically, the CI also introduces the two aforementioned accessibility constraints. First, the interpretation site of the CI strictly subordinates its introduction context ($1 < 3$), and secondly, the interpretation site of the CI is equated with the interpretation site of its anchor ($3 = 2$). In this way, the CI indirectly constrains the interpretation site of the anchor. Since the CI requires identity to the interpretation site of the anchor *and* projection, the presupposition triggered by the anchor can only be accommodated non-locally (thus, the MAP constraint $1 \leq 2$ is strengthened to $1 < 2$).

(11) Jake, a famous boxer, lives in Utrecht.

1

2 ←x 3 ←y
2 ←Jake(x) 3 ←famous_boxer(y) 3 ←y=x 1 ←lives_in_Utrecht(x)
$1 \leq 2$ $1 < 3$ $3 = 2$

It should be noted that the MAP constraints in (11) cannot be satisfied within the current discourse representation. This is the case because the CI requires a non-local context to accommodate to ($1 < 3$), but no such context is available in this representation. Nonetheless, the PDRS shown in (11) is felicitous. This is because PDRSs (like DRSs) are considered a *partial* representation of the discourse; they will be interpreted with respect to some larger model, in which the unresolved presuppositions can be verified (see Kamp and Reyle 1993; see also the analysis of indexicals proposed by Hunter 2013, where an additional global DRS is introduced to which indexical expressions accommodate).

4.4.2 Predicting CI infelicity in PDRT

The MAP constraints introduced by CIs generate straightforward predictions about CI (in)felicity. Example (12), for instance, shows that the constraints predict infelicity when a CI occurs with a non-specific anchor (from McCawley, 1998).

(12) #No climber, a lunatic, survived.

1

$1 \leftarrow \neg$

2

$2 \leftarrow x \;\; 4 \leftarrow y$

$2 \leftarrow \text{climber}(x)$
$4 \leftarrow \text{lunatic}(y)$
$4 \leftarrow y{=}x$
$2 \leftarrow \text{survived}(x)$

$\mathbf{2 < 4 \;\; 4 = 2}$

In this example, the non-specific anchor is locally accommodated. As a consequence, its interpretation site is the same as the local context of the CI. This leads to a contradiction within the accessibility constraints; the constraints introduced by the CI simultaneously require the CI to project out of its local context (strict subordination; $2 < 4$), and its interpretation site to be equal to that of its anchor (identity; $4 = 2$). These constraints ($2 < 4$ and $4 = 2$) cannot be simultaneously satisfied, and therefore the PDRS is infelicitous.

The PDRT analysis also predicts that CIs cannot be cancelled. This is illustrated in (13), which shows the PDRS for example (5b) from above. In this example, the presupposition triggers *Obama* and *the President*, introduced in respectively the antecedent and the consequent of an implication, refer to the same entity. This is reflected by the fact that these conditions affect the same referent and are assigned the same pointer.[3] As such, the contribution made by the CI (that the president is a socialist) is the same as the contribution in the antecedent of the implication (that Obama is a socialist). This equivalence renders the antecedent of the implication void (i.e., not locally informative; cf. van der

[3]In PDRT, uniqueness of referents is determined on the basis of their pointer, whereas in DRT it is determined on the basis of their introduction site. Therefore, referents may be introduced more than once in PDRT, which is not allowed in DRT due to ambiguous binding. See Venhuizen et al. (2014a) for details.

Sandt, 1992). This is because the identity constraint ($6 = 5$) introduced by the CI forces the CI content to project out of the consequent to the interpretation site of its anchor. As a result of this piggybacking, the content contributed by the antecedent is already established within the context in which it is introduced; this is because the MAP constraint in the antecedent ($2 \leq 5$) indicates that the interpretation site of the CI content is accessible from the antecedent's local context. Hence, because CIs piggyback on their anchor, they cannot be cancelled unless their anchor is also cancelled. In (13), the latter is not the case, and the example is therefore infelicitous.

(13) ??If Obama is a socialist, then the President<u>, who is a socialist,</u> will raise taxes.

<table>
<tr><td colspan="3" align="center">1</td></tr>
<tr><td colspan="3"></td></tr>
<tr>
<td>1 ←
2
<table><tr><td>5 ←x</td></tr><tr><td>5 ←Obama(x)
2 ←socialist(x)</td></tr><tr><td>2 ≤ 5</td></tr></table></td>
<td>⇒</td>
<td>3
<table><tr><td>5 ←x 6 ←y</td></tr><tr><td>5 ←President(x)
6 ←socialist(y)
6 ←y=x
3 ←raise_taxes(x)</td></tr><tr><td>3 ≤ 5 3 < 6 6 = 5</td></tr></table></td>
</tr>
<tr><td colspan="3"></td></tr>
</table>

4.4.3 Expressive CIs and speaker-orientedness

In Potts' (2005) redefinition of the class of CIs, a central property of CIs is *speaker-orientedness*. The importance of this property is particularly emphasized in the analysis of expressive CIs, which contribute an opinion (the speaker's, in most cases) to the at-issue content of a sentence. In PDRT, the attribution of subjective attitudes is a pragmatic process, driven by context; both the possible belief-bearers and the polarity of their beliefs are derived by pragmatic inferencing and the application of world knowledge. This is exemplified by (14) (adapted from Kratzer, 1999), where the choice of verb affects to whom the subjective content of the epithet (*that bastard*) is contributed.

(14) a. Father said he would not allow me to *hit* that bastard Webster.
b. Father said he would not allow me to *marry* that bastard Webster.

In (14a), the verb *hit* suggests a negative attitude of the speaker toward the person named 'Webster', and therefore world knowledge will tell us that the speaker is the most likely bearer of the (negative) subjective content triggered by the CI (*that bastard*). In contrast, this same subjective content will most likely be attributed to the father in (14b), since the speaker's attitude toward Webster seems to be positive; the context created by the verb *marry* suggests that the speaker wants to marry Webster.

Pragmatics-driven subjectiveness can be implemented in PDRT by exploiting the unresolved nature of free pointers in the model-theoretic interpretation. In the model-theoretic interpretation of classic DRT (see, e.g., Kamp, 1981; Kamp and Reyle, 1993; Kamp et al., 2011), a DRS is considered a *partial model*, representing the information conveyed by an utterance. In order to determine the truth-conditional interpretation of this DRS, the partial model needs to be embedded in a *total model*. This embedding is done via an *embedding function* that maps the set of discourse referents in the universe of a DRS K, onto elements in the domain of a total model M. This mapping is done in such a way that all conditions of K are verified in M (with respect to some world w). A model-theoretic interpretation of PDRT can be constructed in a similar way. However, the embedding function for PDRSs needs to have proper machinery for dealing with projection variables. Content associated with bound projection variables should be interpreted in its binding context. Content associated with free variables, on the other hand, still has an undetermined interpretation site, and should be resolved to an appropriate context, on the basis of the semantic constraints in the MAPs, and potential additional pragmatic constraints. One possible interpretation site for unresolved content is the global discourse context, which reflects the common ground of the speaker and hearer. Alternatively, unresolved content may be interpreted in a *subjective* discourse context, which provides a means of dealing with speaker-orientedness in PDRT.

Speaker-orientation is formalized in PDRT by assigning each discourse agent (e.g., the speaker, the hearer, and possible other individuals introduced in the discourse) a subjective model M_s, which describes the current state of affairs in the world, as believed by that agent. Critically, this subjective model may or may not coincide with the *actual* state of affairs in the world, which is represented in the truth model M_T. Pragmatic constraints determine whether subjective content associated with a free pointer should be verified with respect to a discourse agent's subjective model M_s, rather than with respect to the truth model M_T. If this is the case, verification of the entire utterance with respect to the truth model M_T entails that this subjective content *only* needs to be verified with respect to the subjective model M_s. Provided this analysis, we can represent the subjective sentence from example (14) as the PDRS shown in (15).

(15) Father$_F$ said he would not allow me$_S$ to hit/marry <u>that bastard</u> Webster.

$$
\begin{array}{c}
1\\
\begin{array}{|l|}\hline
5 \leftarrow \text{x} \;\; 1 \leftarrow \text{p} \\ \hline
5 \leftarrow \text{Father(x)} \\
1 \leftarrow \text{said(x,p)} \\
1 \leftarrow \text{p:} \;
\begin{array}{c}
2\\
\begin{array}{|l|}\hline
\\ \hline
2 \leftarrow \neg \;
\begin{array}{c}
3\\
\begin{array}{|l|}\hline
3 \leftarrow \text{e} \;\; 6 \leftarrow \text{y} \;\; 7_{S/F} \leftarrow \text{z} \\ \hline
3 \leftarrow \text{allow(x,e)} \\
3 \leftarrow \text{hit/marry(e)} \\
3 \leftarrow \text{Agent(e,S)} \\
3 \leftarrow \text{Patient(e,y)} \\
6 \leftarrow \text{Webster(y)} \\
7_{S/F} \leftarrow \text{bastard(z)} \\
7_{S/F} \leftarrow \text{z=y} \\ \hline
3 < 6 \;\; 3 < 7_{S/F} \;\; 7_{S/F} = 6 \\ \hline
\end{array}
\end{array} \\ \hline
\\ \hline
\end{array}
\end{array} \\ \hline
1 < 5 \\ \hline
\end{array}
\end{array}
$$

In this example, there are two relevant discourse agents that may express their subjective beliefs: the speaker and the father, which are associated with subjective models M_S and M_F, respectively. To verify the truth of the entire utterance with respect to a truth model M_T, the subjective content expressed by the CI (*that bastard*), which is associated with a free pointer, should be verified with respect to either M_S (in case of the verb *hit*) or M_F (in case of *marry*). In (16), the subjective model with respect to which the subjective CI content is interpreted, is indicated by means of the subscript (S or F) attached to the free pointer. These 'subjective' projection variables are tuples of pointers and subjective models (e.g., $\langle 7, S\rangle$ or $\langle 7, F\rangle$), that inform the embedding function with respect to which model the subjective content should be interpreted.

In summary, speaker-orientedness is dealt with at the pragmatic, rather than the semantic level in PDRT. This explains the context-dependency of the phenomenon, both on the linguistic context, as illustrated in (14), and on the general discourse context, which includes non-linguistic factors, such as the type of *speaker* (cynical, authoritative, etc.), and the *situation* in which a sentence is uttered (for example, a formal setting or an informal conversation) (see Hoeks and Brouwer, 2014). By accounting for CIs within the semantic representation, while attributing their speaker-orientedness to the pragmatic, interpretational domain, the PDRT analysis of CIs intuitively captures the multi-dimensionality of their contribution. This shows that PDRT provides a robust and flexible framework for the representation and interpretation of different types of linguistic phenomena and their associated levels of information.

4.5 Discussion

Formal analyses of CIs have often focused on differentiating their contribution from at-issue content, as well as from that of other projection phenomena, such as presuppositions. In this chapter, we have proposed an analysis of the projection behavior of CIs, which highlights their correspondence to presuppositions and anaphora, and treats them within the same dimension as at-issue content. In this analysis, formalized in the framework of Projective Discourse Representation Theory, CIs piggyback on a specific anchor because their contribution provides an elaboration on the description of the referent referred to by this anchor.

Concerning Potts' (2005) class of CIs, we have shown that our analysis accounts for the (in)felicity of supplemental and expressive CIs that have a nominal anchor. However, Potts also discusses CIs that do not have a nominal anchor, such as utterance modifiers (e.g., "frankly") and expressive adverbs (e.g., "fucking"). If these expressions do indeed trigger CIs, we might have to extend our analysis to non-nominal anchors. The CI status of these expression has, however, been a subject of debate. Amaral et al. (2007, 725-729), for instance, provide an extensive discussion about utterance level modifiers, like "frankly", arguing that they behave differently from other CIs, both on the theoretical and implementational level; they seem to target the *act* of uttering a statement, rather than the utterance itself. Similarly, Geurts (2007) argues that expressive adverbs, like "fucking", have a semantic dimension of their own, thereby differentiating them from, in particular, supplemental CIs. Hence, the analyses of Amaral et al. (2007) and Geurts (2007) cast doubt on the CI status of utterance modifiers and expressive adverbs. Based on the analysis of CIs as piggybacking on their anchor, we may have to draw a similar conclusion; it remains to be seen whether there indeed exist CIs that have a non-nominal anchor.

A core assumption of our analysis of CIs is that they elaborate the description of a specific referent in the discourse context, as referred to by an anchor. As such, CIs piggyback on their anchor by projecting to its interpretation site. Schlenker (2013), however, argues that there exist CIs that are interpreted *in situ*, and thus do not project to the interpretation site of their anchor. One example, in which he argues this to be the case, is shown in (16).

(16) [Context: someone made a big mistake at the Department.]
If tomorrow I called the Chair, who in turn called the Dean, we would be in deep trouble.

According to Schlenker (2013), the supplemental construction *who in turn called the Dean* is an NRRC that triggers a CI, which is interpreted locally, as indicated by the licensing of the past tense in the relative clause. However, as the relative clause introduces an additional conditional statement, rather than an elaboration on the description of the referent referred to by the anchor *the Chair*, we disagree with Schlenker (2013) on the CI status of this contribution.

Nouwen (2014) also argues for the existence of CIs that have an *in situ* interpretation, which he exemplifies by means of (17).

(17) If a professor, a famous one, writes a book, he will make a lot of money.

According to Nouwen (2014), the indefinite anchor *a professor* can be interpreted non-specifically, while at the same time occurring felicitously with the appositive construction *a famous one*. This means that the CI triggered by the appositive is interpreted *in situ*, namely in the same context as its non-specific, and therefore non-projecting, anchor. However, by virtue of being interpreted locally, the appositive becomes restrictive. As such, it does not provide a backgrounded elaboration on its anchor, and is therefore not a CI.

4.6 Conclusion

We have proposed an analysis of how and why conventional implicatures project; CIs piggyback on their anchor, which refers to a specific referent in the discourse context. This piggybacking is due to the fact that CIs signal novel, backgrounded information that elaborates the description of the referent referred to by the anchor. We have formalized our analysis in the framework of Projective Discourse Representation Theory. This PDRT analysis provides a unified account of the projection behavior of CIs, presuppositions, and anaphora, and treats their contribution within the same dimension as that of at-issue content.

Chapter 5

Aligning discourse semantics with surface structure

Abstract. Traditional semantic formalisms typically do not represent the information status of semantic content (e.g., *given* versus *new*), because it is not considered part of the truth-conditional contribution of an utterance. However, information status crucially determines how linguistic content is related to the current discourse context, and therefore affects the dynamic semantic interpretation. To overcome this shortcoming, we here describe an extension of the widely used Discourse Representation Theory (DRT) framework, called *Projective* Discourse Representation Theory (PDRT). In PDRT, the semantic representations directly reflect the linguistic surface structure of an utterance; the interpretation of semantic content, which may or may not be local, is indicated by means of *projection variables*. The interaction between the projection variables and the surface structure explicitly reflects the way in which information is contributed to the discourse context, that is, the information status. Critically, the incorporation of projection variables affects the formal properties of DRT non-trivially, since the hierarchical structure of the PDRT representations does not reflect their logical structure—just like in natural language. We here show that PDRT extends all formal properties of traditional DRT, using minimal additional machinery. As such, PDRT establishes a significant step toward incorporating non-truth-conditional aspects of meaning into a formal semantic representation.

5.1 Introduction

Discourse Representation Theory (DRT) is a representational framework for dynamic semantics, which was originally developed by Hans Kamp in order to account for discourse anaphoric dependencies and the representation of tense in natural language (Kamp, 1981). Together with Irene Heim's introduction of File Change Semantics (Heim, 1982), DRT established a departure from Montagovian truth-conditional semantics. In contrast to traditional approaches to semantics, dynamic approaches take the discourse context into account in the interpretation of semantic content; anaphoric expressions, for instance, obtain an interpretation by virtue of the context in which they occur, i.e., by binding to an accessible antecedent (Kamp and Reyle, 1993). DRT has proven to be a powerful framework that provides an analysis for a wide range of linguistic phenomena (see, e.g., Geurts and Beaver, 2011, for an overview). Critically, however, these phenomena are mostly limited to truth-conditional aspects of meaning, thereby excluding non-truth-conditional aspects such as information status, which reflects whether information is signalled to be given (e.g., *presupposed* content) or new (*asserted* content). Interestingly, this contrasts with the observation that information status crucially influences the way in which linguistic content interacts with the discourse context. For instance, a definite description, as in "*the man walks*", signals that the discourse context contains a salient referent to which "*the man*" refers, whereas an indefinite noun phrase, as in "*a man walks*", signals the introduction of a new entity into the discourse context. The semantic representations from traditional DRT are not expressive enough to represent this distinction, as they merely reflect the *content* contributed by an utterance, and not the way in which this content is presented, i.e. its *information status*. Any analysis of presuppositions in DRT therefore comes at the cost of losing information about the way in which their content is contributed to the discourse context, and hence fails to capture the fundamental distinction between asserted and presupposed content.

In order to resolve this conflict, we describe an extension of DRT that incorporates information status by providing enriched representations that reflect both the interpretation of linguistic content, and the way in which this content is contributed to the discourse. This extension, called Projective Discourse Representation Theory (PDRT), employs *projection variables* to indicate the interpretation site of linguistic content, so that all content can be represented at the place where it is introduced in the linguistic surface structure. The interaction between the introduction and interpretation site of linguistic content directly reflects its information status; asserted content is associated with projection variables that are bound locally, i.e., at the introduction site, whereas presuppositions are represented using projection variables that are either bound in some higher context (indicating that the content is *projected* to that context) or occur free (indicating that the interpretation site of the presupposition is not yet resolved; see Venhuizen et al. 2013b). Moreover, the addition of projection variables in PDRT allows for representing explicit constraints on the interpretation of semantic content, which can be shown to account for the information status of conventional implicatures (see Venhuizen et al., 2014b). PDRT thus enhances the representational power of the DRT formalism, and allows for explicitly representing different kinds of information status. As such, PDRT presents a departure from semantic analyses that are limited to truth-conditional aspects of meaning, and unifies various approaches that have aimed to incorporate the additional level of information associated with non-truth-conditional content within DRT (see, e.g., van der Sandt, 1992; Krahmer, 1998; Geurts, 1999; Geurts and Maier, 2003, 2013; Hunter, 2013).

An important implication of the addition of projection variables to all linguistic material is that the formal properties underlying DRT are affected non-trivially. In particular, the representation of projected material at the place where it is introduced in the syntactic surface structure violates the traditional DRT notion of accessibility, thereby compromising one of its trademarks: the treatment of anaphora. In the current chapter, we describe the formal and theoretical implications of enhancing DRT with projection variables. We derive the formal properties of the PDRT framework, thereby reestablishing the traditional DRT notions of context accessibility, variable binding, and composition (in the sense of Muskens, 1996).[1] These definitions illustrate that PDRT does not radically change the way in which linguistic meaning is constructed. Since the projection variables are subject to the same constraints as the discourse referents from DRT, the formalization merely extends the notions of variable binding and context accessibility already available in traditional DRT. In other words, PDRT presents a minimal extension of the DRT framework in order to obtain maximal additional representational power.

This chapter is organised as follows. In Section 5.2, we first describe some preliminaries on DRT and the motivation for its extension PDRT. Then, in Section 5.3, we introduce Projective Discourse Representation Structures (PDRSs) and define the notions of accessibility and binding. On the basis of these definitions, the compositional properties of PDRSs are described, using different kinds of merge for asserted and projected content. The resulting structures and their representational power is accordingly illustrated in Section 5.4. Finally, Section 5.5 provides a discussion of the relation of the PDRT framework to other DRT variants, its interpretational consequences and the general implications for semantic theory, providing several directions for future work.

5.2 Toward a formal framework of projection

5.2.1 Discourse Representation Theory

In the dynamic framework of Discourse Representation Theory (DRT; Kamp, 1981; Kamp and Reyle, 1993), the meaning of a discourse is represented by means of recursive units called *Discourse Representation Structures* (DRSs). A DRS consists of a set of *discourse referents* and a set of *conditions* on these referents. Conditions may be either *basic*, reflecting a property or a relation between referents, or *complex*, reflecting embedded contexts introduced by semantic operators such as negation, implication, and modal expressions. DRSs are often visualized using a box-representation consisting of two parts: the set of referents is rep-

[1]The PDRT definitions have also been implemented, alongside with the traditional DRT formulation, as a Haskell library, called PDRT-SANDBOX. This is available at: `http://hbrouwer.github.io/pdrt-sandbox/`

resented in the top of the box, and the set of conditions on these referents are shown in the body. Example (1) shows the DRS of a complex sentence, containing a single complex condition representing an implication, which consists of two embedded DRSs; the first containing two referents and a set of three basic conditions, and the second containing a single basic condition (example adapted from Geach, 1962).

(1) If a farmer owns a donkey, he feeds it.

$$\begin{array}{|c|}\hline \\ \hline \begin{array}{|l|}\hline \text{x y} \\ \hline \text{farmer(x)} \\ \text{donkey(y)} \\ \text{own(x,y)} \\ \hline \end{array} \Rightarrow \begin{array}{|c|}\hline \\ \hline \text{feed(x,y)} \\ \hline \end{array} \\ \hline \end{array}$$

Each (embedded) DRS—each *box*—can be seen as representing a *context*. Taken together, these contexts constitute the logical form of the discourse. Crucially, anaphoric binding of referents is determined on the basis of the accessibility between contexts; in the example above, the antecedent DRS of the implication is accessible from the consequent DRS, and as a result the variables introduced by the pronouns ("*he*" and "*it*") become bound by the referents introduced in the antecedent DRS (x and y).[2]

The analysis of anaphora is DRT's trademark feature, and one of the main motivations behind its development (Kamp, 1981; Heim, 1982), but over the years DRT has been shown to provide a flexible framework that can account for a wide range of other linguistic phenomena as well, including tense (Kamp, 1981), quantification and plurality (Kamp and Reyle, 1993), modal subordination (Roberts, 1989), attitude reports (Asher, 1986, 1989; Zeevat, 1996; Maier, 2009a), and discourse structure (Asher and Lascarides, 2003). Moreover, DRT forms the formal basis of one of the most influential theories of presupposition, namely

[2] In the traditional formulation of DRT, each anaphoric element introduces a discourse referent itself, which is identified with its antecedent via an equality-relation (Kamp, 1981; Kamp and Reyle, 1993). It is straightforward to show that the resulting representation is truth-conditionally equivalent to the representation above, where the equality statement is eliminated by reusing the discourse variables x and y.

the "presupposition as anaphora" account developed by van der Sandt (1992), and further worked out by Geurts (1999). Interestingly, however, this account simultaneously demonstrates the strength and weakness of DRT; it provides an analysis that makes empirically strong predictions, but at the same time pushes the boundaries of the formalism, as it aims to capture a dimension of meaning that is not inherent to DRT. The main issue with the analysis proposed by van der Sandt is that presupposition projection is treated as a deviation from standard meaning construction. As we will show, the analysis of presuppositions can be implemented more parsimoniously in PDRT, where projected and asserted content are treated as two sides of the same coin.

5.2.2 Beyond the surface structure

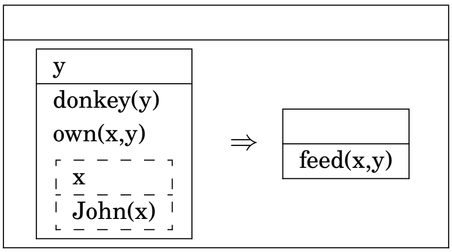
In DRT, *form* determines *interpretation*. That is, the context in which some content appears determines how it is interpreted. Thus, if some content occurs within the (syntactic) scope of, for example, an implication or a negation in the linguistic surface form, it will be interpreted within the logical scope of this operator. This tight correspondence between form and meaning challenges a parsimonious account of *projection*, i.e., the property of being interpreted outside the logical scope of some embedding operator (Langendoen and Savin, 1971). The most paradigmatic class of phenomena exhibiting this property is the class of presuppositions. For example, in (2) the proper name "John" is embedded in the antecedent of an implication, and triggers the presupposition that there exists some person named *John*.

(2) If John owns a donkey, he feeds it.

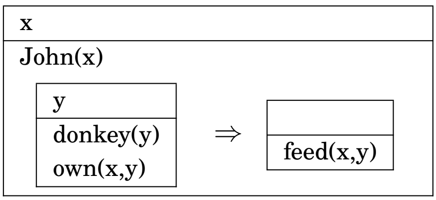
Although this example is almost identical to example (1) in terms of the linguistic surface structure, there is a clear difference in interpretation. Whereas the indefinite "a farmer" in (1) is interpreted in the local context, i.e., as part of the antecedent of the conditional, the presupposition about the existence of "John" in (2) should be interpreted *outside* the scope of the implication. This results in a reading of (2) in which there exists some person named *John*, for whom it holds that if he owns a donkey, then he feeds it. One way of accounting for this interpretation is by making sure proper names always introduce their discourse referents in the global DRS (Kamp, 1981; Kamp and Reyle, 1993), or by simply treating them as constants, which always refer to some specific entity in the discourse context (Muskens, 1996). However, these solutions are difficult to extend to other presupposition triggers, such as definite descriptions and factives, as well as to the related class of conventional implicatures (Potts, 2005).

An important advance in the analysis of presuppositions was made by van der Sandt (1992), who proposed to treat presupposition projection as a variety of anaphora resolution. The account, that was formalized in DRT, has become known as 'binding and accommodation theory', or simply 'binding theory' (see also Geurts, 1999; Beaver, 2002; Bos, 2003). According to van der Sandt, presuppositions behave like anaphora in that they can anaphorically bind to an earlier introduced antecedent (see also Soames, 1979; Kripke, 2009). Unlike anaphora, however, presuppositions can occur felicitously in contexts where no suitable antecedent can be found by creating their own antecedent at an accessible discourse level; this is called *accommodation* (which was first described by Karttunen, 1974, and Stalnaker, 1974, and later named as such by Lewis, 1979). The resolution procedure for presuppositions in binding theory is illustrated in (3).

(3) a. Stage I:



b. Stage II:



As this example shows, the resolution of presuppositions in binding theory involves two stages of processing; in the first stage, presuppositions appear marked at their introduction site, and in the second stage the presupposition is resolved by either binding it to an antecedent or accommodating it to an accessible context. In the example above, the presupposition triggered by "John" does not have a suitable antecedent, which means that it needs to be accommodated. In order to obtain the desired interpretation of (2), the presupposition is accommodated to the global discourse context.

Van der Sandt defines several acceptability constraints that determine the interpretation site of presuppositions, including *semantic* constraints, such as the requirement that bound variables cannot become free as a result of projection ('variable trapping'), and *pragmatic* constraints, such as local and global consistency and informativeness (for more details, see van der Sandt, 1992). Moreover, the analysis assumes a preference for global accommodation over local accommodation. On the basis of these properties of projection, binding theory can account for the projection behavior of various presupposition triggers, including proper names, definite determiners, and possessives. However, despite the empirical robustness of binding theory, the resolution process as well as the obtained representations are not completely satisfactory. Firstly, the two-stage process of deriving the representations is at odds with a compositional construction procedure for DRSs, since presuppositions are only resolved *after* discourse construction has been completed. Therefore, the intermediate representations with unresolved DRSs are ill-defined with respect to formal properties like accessibility and variable binding, which hinders a compositional formalization in terms of Muskens (1996). Secondly, asserted and presupposed content obtain the same status in the resulting representations. This obliterates the information structure of the discourse and precludes any interpretation in which their contributions are distinguished from each other (cf. Kracht, 1994; Krahmer, 1998). Critically, however, these limitations are not inherent to binding theory, but rather to the formalism in which it is implemented; since DRT cannot capture aspects of meaning that go beyond its logical representation, the contribution made by presuppositions and related phenomena cannot be captured in a satisfactory manner.

To overcome these issues, we here describe *Projective* DRT (PDRT), a framework that extends DRT with an inherent treatment of projection via the use of projection variables. The PDRT analysis of presuppositions basically follows van der Sandt's treatment of presuppositions, except that projection does not result in the movement of semantic content within the representation, but is represented by means of variable binding. This eliminates the need for a two-stage resolution algorithm, and thereby provides a more compositional treatment of projection phenomena. Moreover, the application of variables for indicating projection allows for a natural distinction between asserted and presupposed content; presupposed content that cannot be resolved in the current context, i.e., still needs to be accommodated, is represented by means of a free variable, reflecting its potential to be resolved as part of a larger context (e.g., the current common ground). PDRT thus provides a parsimonious treatment of projection phenomena by exploiting a central component of traditional DRT: the binding of variables.

5.2.3 Introducing PDRT

In recent work, we have informally introduced the basic ideas underlying Projective Discourse Representation Theory, and shown how the formalism can be applied to account for the projection behavior of presuppositions and conventional implicatures (Venhuizen et al., 2013b, 2014b). Here, we summarize the basic intuitions behind the framework, and show how the introduction of projection variables affects the traditional DRT notion of accessibility.

5.2.3.1 Representing projected content

A basic structure in PDRT carries more information than a basic structure in DRT; in addition to the structural and referential content of a DRS, a Projective Discourse Representation Structure (PDRS) also makes the information structure of a discourse explicit by keeping linguistic content at its introduction site, and indicating its interpretation site via *projection variables*. Each PDRS introduces a *label* that can be used as an identifier (similar to the context identifiers implicitly assumed by other DRT extensions, e.g., Segmented DRT; Asher and Lascarides, 2003). Moreover, all referents and conditions of a PDRS are associated with a *pointer*, which is used to indicate in which context the material is interpreted. Example (4) shows the PDRS for example (1).

(4) If a farmer owns a donkey, he feeds it.

$$
\begin{array}{c}
1\\
\begin{array}{|c|}
\hline
\\
\hline
\begin{array}{c}2\\ \begin{array}{|l|}\hline 2\leftarrow \text{x}\ \ 2\leftarrow \text{y}\\ \hline 2\leftarrow \text{farmer(x)}\\ 2\leftarrow \text{donkey(y)}\\ 2\leftarrow \text{own(x,y)}\\ \hline \\ \hline\end{array}\end{array}
\Rightarrow
\begin{array}{c}3\\ \begin{array}{|l|}\hline \\ \hline 3\leftarrow \text{feed(x,y)}\\ \hline \\ \hline\end{array}\end{array}\\
\hline
\\
\hline
\end{array}
\end{array}
$$

Projection variables are here represented as natural numbers; the labels introduced by the PDRSs are shown on top of each PDRS, and the pointers associated with referents and conditions are indicated using a leftward pointing arrow '$\leftarrow$'. In (4), all pointers are bound in the local context; that is, the pointers of all referents and conditions coincide with the label of the PDRS in which they are introduced. This indicates that the content should be interpreted locally: it is *asserted*. Now consider the PDRT representation for example (2), shown in (5).

(5) If John owns a donkey, he feeds it.

$$
\begin{array}{c}
1\\
\begin{array}{|c|}
\hline
\\
\hline
\begin{array}{c}2\\ \begin{array}{|l|}\hline 4\leftarrow \text{x}\ \ 2\leftarrow \text{y}\\ \hline 4\leftarrow \text{John(x)}\\ 2\leftarrow \text{donkey(y)}\\ 2\leftarrow \text{own(x,y)}\\ \hline 2\leq 4\\ \hline\end{array}\end{array}
\Rightarrow
\begin{array}{c}3\\ \begin{array}{|l|}\hline \\ \hline 3\leftarrow \text{feed(x,y)}\\ \hline \\ \hline\end{array}\end{array}\\
\hline
\\
\hline
\end{array}
\end{array}
$$

In this example, the proper name "John" introduces a presupposition, which yet needs to be accommodated. This unresolved presupposition occurs *in situ* and obtains a *free* variable pointer in the PDRT representation; the pointer associated with "John" (here, pointer 4) does not coincide with the label of any PDRS that is available in the current representation, indicating that its interpretation site has not yet been determined. This representation thus differs from the DRT representation shown in (3), where the presupposition is explicitly *moved* to outside the scope of the implication.

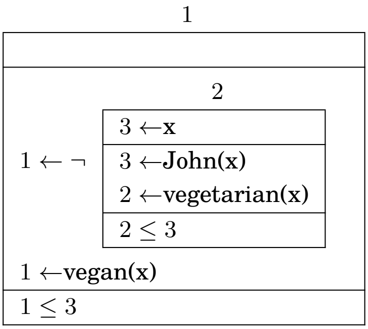
The only information available about the projection site of the presupposed content in (5) is the information derived from the set of 'Minimally Accessible Projection contexts' (MAPs, for short), which are shown in the footer of each PDRS. The MAPs contain constraints on context accessibility, similar to the constraints introduced by Reyle (1993, 1995) for his Underspecified DRSs. In (5), the MAPs indicate that the interpretation site of the projected content (indicated with pointer 4) should be accessible from the context in which the content is introduced (the PDRS labeled 2); this is represented as the constraint $2 \leq 4$ (see Section 5.3.2 for a formal description of MAPs). Note, however, that in this example the MAPs do not fully determine the interpretation site of the projected content; the content may still be interpreted locally, globally, or even in some context outside of the current discourse context, which may become available during discourse construction. Pragmatic heuristics can be employed on top of the MAPs in order to fully determine the interpretation site of projected content; for example, by assuming a preference for global accommodation over local accommodation (e.g., van der Sandt, 1992; Geurts, 1999), or by employing Zeevat's (2002) constraints formulated in Optimality Theory.

In sum, the introduction of projection variables allows for making an explicit distinction between asserted and projected content, in terms of the way in which their content is contributed to the discourse context. Moreover, the use of variables to indicate the interpretation site of semantic content simplifies the construction procedure significantly, as no content needs to be moved within the representation. As we will see in Section 5.3.4 below, the construction procedure for projected content now parallels the construction procedure for asserted content, only differing in terms of how the merge operations affect the projection variables. Integrating projection into the merge operations for PDRS has several advantages. Firstly, it captures the intuitive distinction between asserted and projected content, in terms of how their content is contributed to the unfolding discourse. Secondly, it allows for defining projection as part of the lexical semantics; this means that the fact that a presupposition like 'John' in (5) is associated with a free projection variable is not stipulated, but rather results from the way in which this presupposition is combined with its direct context on the lexical level (see Section 5.3.4).

5.2.3.2 Non-hierarchical variable binding

One of the trademarks of traditional DRT is its treatment of anaphora, which crucially depends on a systematic definition of context accessibility and binding of referents. The addition of pointers and labels to PDRSs affects these definitions non-trivially, since projected content appears *in situ*, while it inherits the binding properties from its interpretation site. This discrepancy is illustrated in example (6).

(6) It is not the case that John is a vegetarian, he is a vegan.



In the PDRS in (6), the variable x in the predicate *vegan* is bound by the discourse referent introduced by the proper name 'John', despite the fact that the referent is structurally introduced in an embedded (non-accessible) PDRS. This is the case because the referent is associated with a free pointer, indicating a presupposition, which means that it has the potential to project to a higher discourse context. The

MAPs are employed to indicate constraints on the projection site of presuppositions; in (6), the constraint $1 \leq 3$ indicates that context 3, i.e., the context to which the referent introduced by the proper name 'John' projects, should be accessible from context 1, i.e., the context in which the anaphoric expression is interpreted. Thus, variables may bind to projected referents whose *introduction* site is *not* hierarchically accessible; the only requirement is that the *interpretation* site of the projected referent is accessible, which is determined using the projection variables and the MAPs. In what follows, we will work out the formal definitions underlying the PDRT framework, and show how it extends the formalization of traditional DRT, while remaining faithful to the basic DRT notions, such as variable binding.

5.3 Projective Discourse Representation Theory

In this section, we describe the basic properties of Projective Discourse Representation Structures (PDRSs) in terms of their syntax and accessibility relations. For reference, the formal definitions of classic DRT are included in Appendix 5.A (variations of which can also be found in, e.g., Kamp and Reyle, 1993; Bos, 2003; Kamp et al., 2011, and references therein).

5.3.1 Syntax of PDRSs

The box representations shown in Section 5.2 are an intuitive way to look at PDRSs, but less useful for formal reasoning about these structures, or for providing a computational implementation. For this, we adhere to the set-theoretical underpinnings of the DRT formalism (see, e.g., Kamp and Reyle, 1993). In set-theoretic terms, a PDRS is a quadruple that consists of a *label*, a set of *projected referents* (i.e., discourse referents associated with a projection variable), a set of *projected conditions* (i.e., PDRS conditions associated with a projection variable), and a set of *MAPs* (representing accessibility relations between projection variables—see below for a more detailed explanation). Formally, a PDRS is defined as follows:

Definition 8 (PDRS)**.** A PDRS P is defined as a quadruple:
$\langle l, \{\delta_1...\delta_n\}, \{\chi_1...\chi_m\}, \{\mu_1...\mu_k\}\rangle$, where:

i. l is a projection variable;

ii. $\{\delta_1...\delta_n\}$ is a finite set of projected referents (also referred to as the *universe*), with $\delta_i = v_i \leftarrow x_i$, such that v_i is a projection variable, and x_i is a discourse referent;

iii. $\{\chi_1...\chi_m\}$ is a finite set of projected conditions, with $\chi_i = v_i \leftarrow \gamma_i$, such that v_i is a projection variable, and γ_i is a PDRS condition (see Definition 9);

iv. $\{\mu_1...\mu_k\}$ is a finite set of MAPs, with $\mu_i = v_1 \leq v_2$ or $\mu_i = v_1 \nleq v_2$, such that v_1 and v_2 are projection variables.

The definition of PDRS conditions basically follows the definition of DRS conditions proposed by Bos (2003). Besides the standard logical operators for negation ($\neg$), disjunction ($\vee$) and implication ($\Rightarrow$), this definition also includes modal operators for logical necessity ($\Box$), and possibility ($\Diamond$), as well as a hybrid condition (:), which associates a variable ranging over possible worlds with a DRS, and can be used to represent sentential complements (see Bos, 2003). The following definitions describes these PDRS conditions:

Definition 9 (PDRS Conditions)**.** PDRS conditions may be either basic or complex, and are defined as follows:

i. $R(x_1, ..., x_n)$ is a basic PDRS condition, with $x_1...x_n$ are discourse referents and R is a relation symbol for an n-place predicate;

ii. $\neg P$, $\Box P$ and $\Diamond P$ are complex PDRS conditions, with P is a PDRS;

iii. $P_1 \vee P_2$ and $P_1 \Rightarrow P_2$ are complex PDRS conditions, with P_1 and P_2 are PDRSs;

iv. $x : P$ is a complex PDRS condition, with x is a discourse referent and P is a PDRS;

v. PDRS conditions are only defined on the basis of clauses i-iv above.

Together, Definitions 8 and 9 define the syntax of PDRSs. Note that this extends the PDRS syntax proposed in Venhuizen et al. (2013b), which did not include the set of Minimally Accessible PDRS-contexts (MAPs). As described above, MAPs introduce constraints on context accessibility, similar to the accessibility constraints for Unresolved DRSs introduced by Reyle (1993, 1995). The MAPs are defined over *PDRS-contexts*, which include all sub-PDRSs, as well as the projected contexts indicated by a free pointer. There are two types of MAPs: $v_1 \leq v_2$ indicates that PDRS-context v_2 is accessible from PDRS-context v_1, and $v_1 \nleq v_2$ indicates that PDRS-context v_2 is *not* accessible from PDRS-context v_1. The first constraint represents *weak subordination*, i.e., v_1 is the same context as v_2 or subordinated by it. The latter constraint indicates that v_1 is not the same context as v_2, nor subordinated by it; it may thus either be the case that v_2 subordinates v_1, or that there exists no subordination relation between v_1 and v_2. Using a combination of these types of MAPs, we can formulate two additional, stronger accessibility constraints: *strict subordination* ($v_1 < v_2$), which can be represented as $v_1 \leq v_2 \ \wedge \ v_2 \nleq v_1$, and *identity* ($v_1 = v_2$), which can be represented as $v_1 \leq v_2 \ \wedge \ v_2 \leq v_1$. We illustrate below in more detail how these stronger constraints can be applied.

The label of a PDRS can be used as an identifier to refer to the PDRS; formally, $lab(P)$ refers to the label of PDRS P. The set of all projection variables of a PDRS P is indicated by $\Pi(P)$, and contains all variables occurring as labels or pointers in P, or as part of the MAPs of P. Formally, this set is defined as follows:

Definition 10 (Projection variables in a PDRS)**.**

i. $\Pi(\langle l, U, C, M\rangle) = \{l\} \cup \bigcup_{u \in U} \Pi(u) \cup \bigcup_{c \in C} \Pi(c) \cup \bigcup_{m \in M} \Pi(m)$;

ii. $\Pi(p \leftarrow x) = \Pi(p \leftarrow R(x_1 \ldots x_n)) = \{p\}$;

iii. $\Pi(p \leftarrow \neg K) = \Pi(p \leftarrow \Diamond K) = \Pi(p \leftarrow \Box K) = \Pi(p \leftarrow x : K) = \{p\} \cup \Pi(K)$;

iv. $\Pi(p \leftarrow K_1 \Rightarrow K_2) = \Pi(p \leftarrow K_1 \vee K_2) = \{p\} \cup \Pi(K_1) \cup \Pi(K_2)$;

v. $\Pi(p_1 \leq p_2) = \Pi(p_1 \nleq p_2) = \{p_1, p_2\}$.

In a similar manner, we can collect the set of projected referents in a PDRS P, represented by $\mathcal{R}(P)$. This includes the projected referents from all universes in P, as well as the projected referents occurring within the conditions of P (these are derived by combining the DRS referents in the conditions with the pointer associated with the condition in which they occur); see Appendix 5.B for the complete definition. The set of all discourse referents in P (i.e., excluding pointers) is referred to by $\Delta(P)$, and can be derived from the set of projected referents as follows:

Definition 11 (Discourse Referents in a PDRS)**.**

$$\Delta(P) = \{x \mid \exists p : p \leftarrow x \in \mathcal{R}(P)\}$$

5.3.2 Binding and Accessibility in PDRT

The definition of free and bound referents in DRT determines when some discourse referent can be anaphorically linked to an earlier introduced referent. This is defined using the notion of *accessibility*, which determines the DRS universes that are accessible from a given context, such that variables introduced in that context can be bound by referents in these universes. In DRT, accessibility is defined based on a subordination relation between DRSs. Informally, the universes that are accessible from a referent introduced in DRS K, are the universe of K itself, and those of any DRS that directly, or indirectly subordinates K (see Definition 23 in Appendix 5.A). Subordination, in turn, is defined as follows (see Definition 24): DRS K_1 *directly* subordinates DRS K_2 if K_2 occurs in a condition in K_1, or K_1 serves as the antecedent of K_2 in an implication, and DRS K_1 *indirectly* subordinates DRS K_2 if K_2 is a sub-DRS of a DRS that directly subordinates K_1.

As described above, the addition of projection variables in PDRT, and the *in situ* representation of projected content critically affect the definitions of binding and accessibility. In contrast to DRT, accessibility in PDRT cannot be simply determined on the basis of context subordination, since the surface structure of a PDRS does not reflect its logical structure—in the same way that the surface structure of a linguistic utterance does not reflect its logical structure. Due to projection, the antecedent of a discourse referent may be introduced in any universe available in the global PDRS, and not just in the universes of the PDRSs that dominate the local context via context subordination, as illustrated in example (6), repeated here as (7).

(7) It is not the case that John is a vegetarian, he is a vegan.

$$
\begin{array}{|l|} \multicolumn{1}{c}{1} \\ \hline \\ \hline 1 \leftarrow \neg \begin{array}{c} 2 \\ \begin{array}{|l|} \hline 3 \leftarrow \mathrm{x} \\ \hline 3 \leftarrow \mathrm{John(x)} \\ 2 \leftarrow \mathrm{vegetarian(x)} \\ \hline 2 \leq 3 \\ \hline \end{array} \end{array} \\ 1 \leftarrow \mathrm{vegan(x)} \\ \hline 1 \leq 3 \\ \hline \end{array}
$$

In this example, the variable x in the predicate *vegan* in PDRS 1 is bound by the projected referent ‘$3 \leftarrow x$’, despite the fact that it appears in an inaccessible universe; according to the traditional DRT notion of accessibility defined above, PDRS context 2 is not accessible from PDRS context 1, since PDRS 1 structurally subordinates PDRS 2. However, accessibility in PDRT is determined based on the *interpretation site* of the content, rather than the context in which it is syntactically introduced. This means that the accessibility of the projected referent should be determined with respect to the context indicated by its pointer, i.e., context 3. As this is a projected context, information about its accessibility constraints should be obtained from the MAPs; here, the constraint ‘$1 \leq 3$’ indicates that the interpretation site of the projected referent is accessible from the interpretation site of the PDRS-condition in which the variable is used, and hence the variable appears bound.

Formalizing this definition of binding in PDRT thus requires an adjusted notion of context subordination, which determines the accessibility relations between all contexts of interpretation; these include contexts introduced by (sub-)PDRS, as well as projected contexts introduced by the free pointers. The accessibility relations between these contexts should take into account the structural subordination constraints from DRT, as well as the additional constraints introduced by the MAPs. This can be formalized using a graph-structure, called the *projection graph*, which contains underspecified accessibility relations for unresolved projected contexts. In what follows, we first describe how to derive the projection graph of a PDRS, and then define binding for both projection variables and projected referents, based on the accessibility constraints that can be derived from the projection graph.

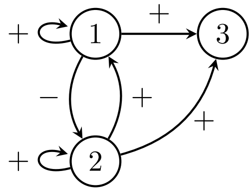
5.3.2.1 Projection graph

A *projection graph* is a partial order over PDRS-contexts, which can be derived from the logical structure of the PDRS and the accessibility constraints in the MAPs. The projection graph is a directed labeled graph (E, V, l), consisting of a set of edges E, a set of vertices V (i.e., PDRS-contexts), and a labeling function l that maps edges to the labels '$+$' and '$-$' (signalling accessibility and inaccessibility, respectively).

The projection graph of a PDRS can be derived directly by traversing the PDRS structure, as shown in Definition 28 in Appendix 5.B. Just like in DRT, a PDRS is accessible from itself and from any other PDRS that it subordinates; i.e. the antecedent of an implication is accessible from its consequent, and the context indicated by the pointer of a condition is accessible from all PDRSs within that condition. Moreover, the projection graph of a PDRS incorporates the additional accessibility constraints provided by the MAPs, as well as the constraint that pointers can only indicate contexts that are accessible from the PDRS in which the pointer is introduced. Importantly, the way in which the projection graph is derived assumes that all projection variables in a PDRS indicate unique discourse contexts; i.e., there cannot be any duplicate uses of projection variables in the PDRS from which the graph is derived, as these cannot be distinguished in the resulting projection graph (in other words, the PDRS must be *pure*; see Definition 18 below).

Example (8) shows the projection graph that can be derived for example (7). The notation used in (8a) defines a projection graph as a set of labeled edges, where an edge between vertices a and b with label l is indicated as $\{\langle a, b\rangle \mapsto l\}$. The graphical representation is shown in (8b); here the spatial ordering loosely reflects the hierarchical structure of PDRS-contexts.

(8) a. $\{\langle 1,1\rangle \mapsto +, \langle 1,2\rangle \mapsto -, \langle 2,1\rangle \mapsto +, \langle 2,3\rangle \mapsto +, \langle 2,2\rangle \mapsto +, \langle 1,3\rangle \mapsto +\}$

b. 

Interestingly, the graph-theoretic properties of the projection graph of a PDRS provide insight into the status of the information represented by it. For instance, the graph in (8) is weakly connected, which means that it is not the case that for every two vertices there is an edge connecting them. This indicates that the PDRS represented by the projection graph still contains undetermined contexts; in this case context 3, which may either be resolved to be the same as context 1, or to some context that subordinates context 1.[3]

Using the projection graph, we can define the accessibility of PDRS-contexts in a PDRS as finding a path p between two vertices, such that all edges in the path indicate a positive accessibility relation. We can define the path-label of a path p as ***pathlab***$(p) = \bigcup_{e \in p} lab(e)$, where $lab(e)$ is e's edge label (Zou et al., 2014). This definition describes the set of labels of all edges that make up a path; note that this set is *unordered*, i.e., duplicates are excluded. Now, accessibility of PDRS-contexts is formally defined as follows (compare: the notion of accessibility in DRT, described in Definition 23 in the appendix). Note that PDRS-contexts (represented by π) refer to both sub-PDRSs available in the global PDRS, and projected contexts indicated by free pointers; a projected referent is part of the universe of a projected context, in case it is introduced in the universe of some sub-PDRS, and its pointer indicates a projected context.

[3]Although the edge from context 3 to context 2 is also undetermined in (8), it can be derived based on the information available in the graph: since $\langle 1,3\rangle \mapsto +$ and $\langle 1,2\rangle \mapsto -$, it must follow that $\langle 3,2\rangle \mapsto -$, since otherwise there would be an indirect positive path from 1 to 2 (based on the transitivity of the accessibility relation), which would make the graph incoherent.

Definition 12 (PDRS accessible universes)**.** The universe of PDRS-context π_j is accessible from PDRS-context π_i in PDRS P with projection graph $\mathcal{G}$, i.e., $\pi_i \leq \pi_j$ (in $\mathcal{G}$), iff:

i. There is a path p from π_i to π_j in $\mathcal{G}$, i.e., $p = \textit{path}(\pi_i, \pi_j, \mathcal{G}) \neq \emptyset$;

ii. p consists only of positive edges, i.e., $\textit{pathlab}(p) = \{+\}$.

Based on the projection graph, we can define variable binding in PDRT. As described above, in order to determine whether some referent is bound, the interpretation site of the referent, as well as the interpretation site of its possible antecedents, must first be determined. This is done on the basis of the projection variables, which in turn may also occur *free* or *bound* in a PDRS. Therefore, we first define variable binding for projection variables, and accordingly describe binding of projected referents in PDRT.

5.3.2.2 Binding of projection variables

The definition for free and bound projection variables in PDRT parallels the DRT definition of free and bound referents. DRS referents can be bound by a referent introduced in the universe of an accessible DRS (see Definition 25 in the appendix). Similarly, projection variables can be bound by the label of some accessible PDRS. This is formalized below.

Definition 13 (Projection Variable Binding)**.** A projection variable v, introduced in some PDRS P_i is bound in global PDRS P (represented as: $\textit{boundpvar}(p, \pi_i, P)$) iff there exists a sub-PDRS P_j in P, such that:

i. P_j is accessible from P_i in the projection graph $\mathcal{G}$ of P, i.e., $P_1 \leq P_j$;

ii. The label of P_j is v, i.e.,$\textit{lab}(P_j) = v$.

Informally, a projection variable v is bound in case it occurs as the label of some PDRS P_j that is accessible from (the label of) the introduction site of v, namely P_i, which has label π_i. Thus, the binding of projection variables in PDRT crucially depends on the introduction site of the projection variable, just like in the DRT definition for bound referents. Note, however, that this introduction site may be a projected context itself, as it may be part of a projected condition. Therefore, the binding of projection variables is not simply determined based on context subordination (as is the case for the binding of referents in DRT), but rather using the PDRT notion of accessibility in terms of the projection graph defined above. The set of all free projection variables in a PDRS P, indicated by $\mathcal{F}_\pi(P)$, can now be derived in a principled manner, by traversing the PDRS and checking for all projected referents in the universes, conditions and MAPs of P whether they occur free (i.e., not bound). Critically, binding is determined for each projection variable relative to the *global* PDRS, such that all possible projected antecedents are taken into account. The full definition is shown in Appendix 5.B, Definition 29.

5.3.2.3 Binding of projected referents

Based on the definition of free and bound projection variables, we can define the binding of referents in PDRT. Since binding crucially depends on the *interpretation site* of the referents, we define binding for *projected referents*, i.e., a discourse referent combined with a pointer, as formalized in Definition 27 in the appendix. The pointer determines the interpretation site of the projected referent, so we define binding based on the existence of an accessibility relation between the pointer of the projected referent and the pointer of its antecedent. This notion of binding of projected referents is formalized as follows.

Definition 14 (Projected Referent Binding)**.** A projected referent $p \leftarrow r$ is bound in global PDRS P (*boundpref*$(p \leftarrow r, l, P)$) iff there exists a PDRS-context $\pi_j \in \Pi(P)$, such that:

i. π_j is accessible from the interpretation site of the projected referent ($p \leq \pi_j$);

ii. $\pi_j \leftarrow r$ is introduced in some universe of P, i.e., there exists some PDRS $P_j \leq P$, such that $\pi_j \leftarrow r \in U(P_j)$.

This definition states that a projected referent $p \leftarrow r$ is bound in case there exists a projected referent $p' \leftarrow r'$ in a universe of the global PDRS, such that $r' = r$ and p' is accessible from p. Based on this notion of binding, it follows that in example (7), the discourse referent x used in the condition $1 \leftarrow vegan(x)$ appears bound by the projected referent $3 \leftarrow x$: since $3 \leftarrow x$ appears in some universe in the global PDRS (namely that of PDRS 2), and there is a positive edge between context 1 and context 3 in the projection graph shown in (8), the occurrence of x in $vegan(x)$ is bound.

The set of all free discourse referents of a PDRS P, represented as $\mathcal{F}_{\mathcal{R}}(P)$, can now be defined straightforwardly by traversing the conditions of P and determining for each referent x_i occurring in a relation-condition ($p \leftarrow R(x_1, ..., x_n)$) or a propositional condition ($p \leftarrow x : K$) whether $p \leftarrow x_i$ is free in P relative to the label of its introduction context l. This is formalized in Definition 30 in Appendix 5.B.

5.3.3 Structural properties

Based on the definitions of free and bound variables in PDRT, we can define several properties of PDRSs. Firstly, a PDRS without any free projected referents is called a *proper* PDRS, just like in DRT (Kamp and Reyle, 1993; Kamp et al., 2011). In general, a PDRS representing a complete sentence or proposition will be proper, meaning that none of the referents remain underspecified.

Definition 15 (Properness)**.** A PDRS P is proper iff P does not contain any free projected referents: $\mathcal{F}_{\mathcal{R}}(P) = \emptyset$.

In PDRT, we have defined free and bound referents, as well as free and bound projection variables. This provides us with an extra level of information about the projective properties of PDRSs. In particular, if a PDRS contains free projection variables, this means that there are still unresolved presuppositions. Thus, we can describe a PDRS without any free projection variables as a *non-presuppositional* PDRS.

Definition 16 (Non-presuppositionality)**.** A PDRS P is non-presuppositional iff P does not contain any free pointers: $\mathcal{F}_{\pi}(P) = \emptyset$.

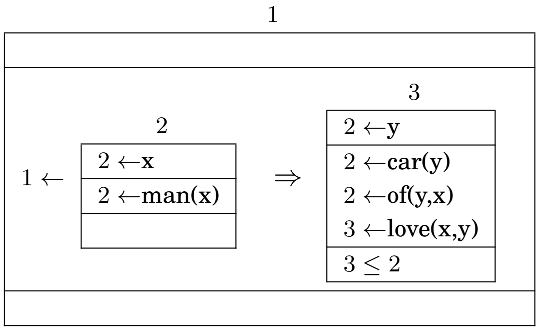
Note that in a non-presuppositional PDRS not all content needs to be asserted; some pointers may be bound by labels of accessible contexts.

Thus, we can also define a property that describes PDRSs with only asserted content, we will call these PDRSs *projectionless*, or *plain*.

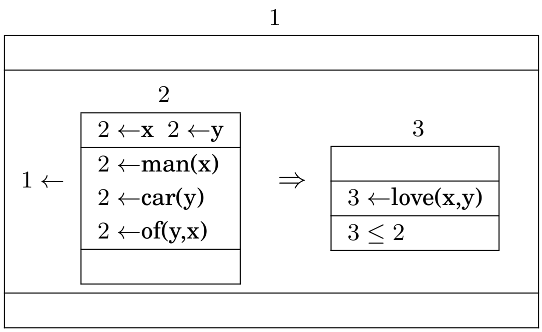
Definition 17 (Plainness)**.** A PDRS P is plain iff all projection variables in P are locally accommodated: For all P', such that $lab(P') \leq lab(P)$, it holds that $\mathcal{F}_\pi(P') = \emptyset$.

Following these definitions, all plain PDRSs are non-presuppositional PDRSs, but not the other way around. Non-presuppositional, non-plain PDRSs are those PDRSs in which some content is accommodated at a higher level, but not presupposed (this is also referred to as *intermediate accommodation*). Interestingly, these properties of PDRSs reflect the structural differences between the sentences that they represent. This is illustrated in (9), which shows two logically equivalent sentences that only differ with respect to the information status of their content; both (9a) and (9b) are non-presuppositional, but only (9b) is plain.

(9) a. Every man loves his car.



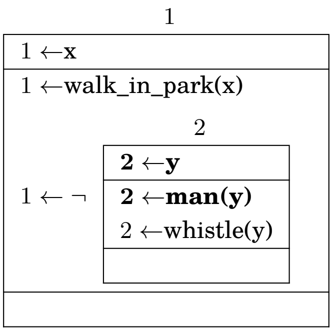
b. Every man who owns a car loves it.



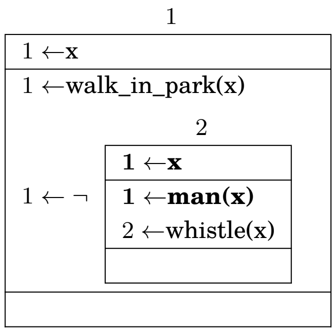
Finally, the property of *purity* refers to the occurrence of duplicate uses of variables. In DRT, this property is defined in order to prevent referents from being bound by multiple instantiations of the same variable. This may happen in case a discourse referent is introduced in multiple accessible universes. In PDRT, discourse referents are allowed to be introduced in multiple universes, as long as they are associated with the same pointer; we call these *shadow referents*. Due to the distinction made in PDRT between the introduction and interpretation site of linguistic content, ambiguities can only arise in case the same discourse referent is *interpreted* in multiple accessible universes. This is illustrated in example (10).

(10) Someone is walking in the park and it's not the case that...

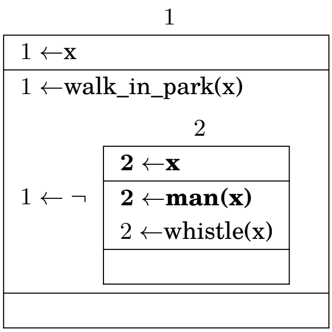
a. **a man** whistles.



b. **he** whistles.



c. **he/a man** whistles.



In (10a), the projected referents introduced by "*someone*" and "*a man*" do not overlap with respect to either their projection variables of their discourse referents, so they explicitly refer to two distinct entities. In (10b), the pronoun "*he*" introduces a *shadow referent*, i.e., a repetition of the projected referent introduced by "*someone*", thereby signalling co-reference. By contrast, the representation in (10c) is ambiguous between an introductory and a co-referential interpretation, because the discourse referent x indicates both PDRS 1 and PDRS 2. As a result, the variable occurring in the PDRS-conditions in PDRS 2 are ambiguously bound by the local introduction of the referent x, and the introduction of "*someone*" in PDRS 1. Therefore, the PDRS in (10c) is not pure.

In a similar manner, a PDRS may be impure with respect to its projection variables, which happens in case a projection variable is used as a label in multiple accessible PDRSs; this may result in ambiguous binding of pointers. The definition of PDRS purity thus consists of two parts; describing impurity with respect to discourse referents, and impurity with respect to projection variables, respectively. This is formally defined as follows (here, $\mathcal{U}(P)$ indicates the union of all universes in P).

Definition 18 (Purity)**.** PDRS P is pure iff:

i. P does not contain any otiose uses of discourse referents (i.e., P does not contain any unbound, duplicate uses of discourse referents): For all P_1, P_2, such that $P_1 < P_2 \leq P$, and $lab(P_1) = p_1$ and $lab(P_2) = p_2$, it holds that: $\{r_1 \mid p_1 \leftarrow r_1 \in \mathcal{U}(P)\} \cap \{r_2 \mid p_2 \leftarrow r_2 \in \mathcal{U}(P)\} = \emptyset$;

ii. P does not contain any otiose uses of projection variables (i.e., P does not contain any unbound, duplicate uses of projection variables): For all P_1, P_2, such that $P_1 < P_2 \leq P$, and $lab(P_1) = p_1$ and $lab(P_2) = p_2$, it holds that: $\{p_1\} \cap (\{p_2\} \cup \mathcal{F}_\pi(P)) = \emptyset$.

5.3.4 Composition in PDRT

As the aim of a semantic formalism is to construct complex meaning representations, we need to define a way to combine structures in order to create larger meaning representations. The combination of basic structures in PDRT basically extends the traditional DRT notion of *merge*. Combining two structures in DRT means creating a novel

DRS that contains all referents and conditions of both conjuncts, and resolves any anaphoric dependencies. To do this, the merge of two DRSs is defined as the union of the sets of referents and conditions from both DRSs. One important condition on this merge, however, is that no accidental binding of variables may occur due to overlapping variables. To appreciate this, consider the following example, where the set-theoretic union of the referent and conditions of two DRSs (indicated by $\oplus$) does *not* yield the desired interpretation.

(11) A dog barks. A cat meows.

$$\begin{array}{|l|}\hline \text{x} \\ \hline \text{dog(x)} \\ \text{bark(x)} \\ \hline \end{array} \oplus \begin{array}{|l|}\hline \text{x} \\ \hline \text{cat(x)} \\ \text{meow(x)} \\ \hline \end{array} = \begin{array}{|l|}\hline \text{x} \\ \hline \text{dog(x)} \\ \text{bark(x)} \\ \text{cat(x)} \\ \text{meow(x)} \\ \hline \end{array}$$

In example (11), the resulting DRS corresponds to the interpretation that there is a single entity x, for which each of the four conditions holds. Clearly, this is not the desired interpretation, on which there are two separate individuals, with two separate conditions each. In order to prevent such undesired bindings, a renaming function $\mathcal{R}$ is applied to the second argument of the merge (K_2). The renaming function basically performs *α-conversion* on K_2 based on a fixed set of variables, namely all discourse referents occurring in K_1, indicated as $\Delta(K_1)$. This function only renames bound variables in K_2, which means that variables that occur free in K_2 may become bound during the merge procedure by some antecedent introduced in K_1. The formal definition for DRS Merge (indicated using the $\bullet$ operator, following van Eijck and Kamp, 1997) is provided below.

Definition 19 (DRS Merge)**.** Given DRSs $K_i = \langle U_i, C_i \rangle$ and $K_j = \langle U_j, C_j \rangle$, the DRS merge between K_i and K_j ($K_i \bullet K_j$) is defined as follows:

$$K_i \bullet K_j = \langle U_i \cup U_{j'}, C_i \cup C_{j'} \rangle$$

where: $K_{j'} = \langle U_{j'}, C_{j'} \rangle = \textit{rename}(K_j, \Delta(K_i))$

Note that in this merge operation the order of the arguments only matters with respect to anaphoric binding; the freely occurring variables from the second argument may have become bound in the resulting DRS. This means that, given two proper DRSs (i.e., without any free variables), DRS merge is α-*symmetrical*; the DRSs resulting from merging both directions are truth-conditionally equivalent and only differ in their variable assignments.

In PDRT, the projection variables should be taken into account when combining structures. In line with Venhuizen et al. (2013b) we can define different types of merge for asserted and projected content.[4] The idea is that the different types of merge reflect the different ways in which content is contributed to the foregoing discourse; asserted content is merely added to the current discourse context, whereas projected content introduces a constraint on the context in which it is introduced, namely the constraint that the projected content can be felicitously bound or accommodated.

5.3.4.1 Assertive Merge

The first type of merge in PDRT is Assertive Merge, an operation between two PDRSs P_1 and P_2 that results in a PDRS in which the asserted content of both P_1 and P_2 remains asserted, i.e., occurring with a pointer that is bound by the label of the local PDRS. The operation is similar to DRS Merge in that it is defined as taking the union of all elements of a PDRS. Just like in the case of DRS Merge, a renaming function makes sure all overlapping bound variables are renamed in P_2, which include both the projection variables of P_1 (i.e., $\Pi(P_1)$) and the discourse referents of P_1 (i.e., $\Delta(P_1)$). The renaming function for PDRSs is worked out in Appendix 5.B, Definition 31. Critically, in order to make sure that all asserted content remains asserted, all occurrences of the label of P_1, $lab(P_1)$, are replaced by the label of the resulting PDRS, i.e., by the label of (the renamed version of) P_2; this is indicated

[4] In Venhuizen et al. (2013b), a third type of merge was defined to account for the projection of Conventional Implicatures (CIs): Implicative Merge. However, this merge variant has become obsolete due to a novel analysis for CIs in PDRT that does not require a special type of merge (see Venhuizen et al., 2014b).

by $P[v_1 \backslash v_2]$, which means that all occurrences of v_1 in P are replaced by v_2. The formal definition for Assertive Merge (indicated using the $+$ operator) is provided below.

Definition 20 (Assertive Merge)**.** Given two PDRSs P_i and P_j, such that $P_i = \langle l_i, U_i, C_i, M_i \rangle$ and $P_j = \langle l_j, U_j, C_j, M_j \rangle$, the assertive merge $P_i + P_j$ is defined as follows:

$$P_i + P_j = \langle l_{j'}, (U_i \cup U_{j'}), (C_i \cup C_{j'}), (M_i \cup M_{j'}) \rangle [l_i \backslash l_{j'}]$$

where: $P_{j'} = \langle l_{j'}, U_{j'}, C_{j'}, M_{j'} \rangle = \textit{rename}(P_j, \Pi(P_i), \Delta(P_i))$

Example (12) illustrates this merge procedure. Here, the two PDRSs that are being merged only contain asserted content; in the resulting PDRS, this content remains asserted. There are no overlapping variables between the two PDRSs, so only the pointers of the content in the left argument of the merge need to be renamed.

(12) A dog barks. A cat meows.

$$
\begin{array}{c} 1 \\ \begin{array}{|l|}\hline 1 \leftarrow \mathrm{x} \\ \hline 1 \leftarrow \mathrm{dog(x)} \\ 1 \leftarrow \mathrm{bark(x)} \\ \hline \\ \hline \end{array} \end{array}
+
\begin{array}{c} 2 \\ \begin{array}{|l|}\hline 2 \leftarrow \mathrm{y} \\ \hline 2 \leftarrow \mathrm{cat(y)} \\ 2 \leftarrow \mathrm{meow(y)} \\ \hline \\ \hline \end{array} \end{array}
=
\begin{array}{c} 2 \\ \begin{array}{|l|}\hline 2 \leftarrow \mathrm{x} \;\; 2 \leftarrow \mathrm{y} \\ \hline 2 \leftarrow \mathrm{dog(x)} \\ 2 \leftarrow \mathrm{bark(x)} \\ 2 \leftarrow \mathrm{cat(y)} \\ 2 \leftarrow \mathrm{meow(y)} \\ \hline \\ \hline \end{array} \end{array}
$$

In line with DRS Merge, Assertive Merge is α-symmetrical given two proper and non-presuppositional PDRSs; that is, if there are no freely occurring projection variables or projected referents, the order of applying the merge does not affect the truth-conditions of the resulting PDRS. However, in case the second argument of the merge contains free variables, these may become bound during the merge procedure. The main contribution with respect to traditional DRT is that not only referents, but also projection variables may become bound during merging. This directly implements van der Sandt's (1992) formalization of presupposition projection as anaphora resolution, since in PDRT presuppositions are resolved in the same way as anaphora, namely by binding to an antecedent during discourse construction. The interpretation site of linguistic content becomes available for binding if during discourse construction it becomes projected; this is the effect of a second type of merge in PDRT, called Projective Merge.

5.3.4.2 Projective Merge

The application of Projective Merge between two PDRSs results in a PDRS in which the asserted content of the first argument of the merge has become projected. This operation is similar to the α-operator introduced by Bos (2003), which is used to separate presupposed from asserted content. In addition, however, Projective Merge defines how the projected content is to be integrated into the discourse context. Projection in PDRT means that the pointer of the projected content indicates some accessible PDRS-context, which may either be instantiated by a higher PDRS, or it may be a projected context in case the pointer cannot be bound by any accessible label. This result can be obtained by defining Projective Merge as an operation that does not affect the pointers of the projected content; the content becomes projected since its pointers are not bound anymore by the local PDRS. Moreover, the set of MAPs of the PDRS resulting from Projective Merge contains an accessibility relation between the resulting PDRS and the context created through projection. This ensures that the interpretation site of the projected content remains accessible from the context that it is merged with, even if this context later becomes projected itself (e.g., in the case of embedded presuppositions, as illustrated in example (14b), below). The formal definition for Projective Merge—indicated using the $*$ operator—is provided below (see Definition 31 for a definition of the renaming function).

Definition 21 (Projective Merge)**.** Given two PDRSs P_i and P_j, such that $P_i = \langle l_i, U_i, C_i, M_i \rangle$ and $P_j = \langle l_j, U_j, C_j, M_j \rangle$, the projective merge $P_i * P_j$ is defined as follows:

$$P_i * P_j = \langle l_{j'}, U_i \cup U_{j'}, C_i \cup C_{j'}, M_i \cup M_{j'} \cup \{l_{j'} \leq l_i\} \rangle$$

where: $P_{j'} = \langle l_{j'}, U_{j'}, C_{j'}, M_{j'} \rangle = \textit{rename}(P_j, \Pi(P_i), \Delta(P_i))$

In contrast to Assertive Merge and traditional DRS Merge, Projective Merge is not an α-symmetrical operation for proper and non-presuppositional PDRSs, since the order of the arguments crucially affects the interpretation of the resulting PDRS; the asserted content of the first PDRS has become projected in the resulting PDRS. This reflects the asymmetrical nature of projected content (and in particular presuppositions), which can be seen as assumptions that have to be fulfilled *before* the asserted content can be evaluated. This results in some interesting interactions between Assertive and Projective Merge, which are illustrated in Section 5.4.2. First, we will briefly discuss how these merge operations contribute to the notion of compositionality in PDRT.

5.3.4.3 Compositional PDRT

The merge operations described so far allow for the combination of *complete* structures, representing for example sentences or propositions. Ideally, however, we could build up structures from even smaller building blocks, such as words. For this, we need a way to express the meaning of unresolved structures that still need to be combined with some additional content in order to form a complete PDRS. Following Muskens's (1996) definition of Compositional DRT, we can define a compositional version of PDRT, in which the unresolved lambda-structures contain PDRSs that are combined using assertive and projective merge. With this machinery, we can deal with projection in a straightforward and intuitive way on the lexical level, so that it becomes an inherent part of meaning composition: presupposition triggers introduce a projective merge to insert their content in their local context, instead of introducing a constant (cf. Kamp and Reyle, 1993; Muskens, 1996) or requiring a post-hoc projection mechanism (such as the algorithm proposed by van der Sandt, 1992, described in Section 5.2.2 above). As described above, projective merge shows a strong correspondence to the α-operator introduced by Bos (2003); this correspondence is also reflected in the resulting lexical semantics. For example, a proper name like "*John*" is a noun phrase that introduces a presupposition about the existence of an entity named 'John'; this can be represented by means of a lambda-term that uses a projective merge to combine the PDRS that introduces ‘John’ with the unresolved structure:

(13) *“John”*: $$\lambda p.\left(\begin{array}{c}1\\ \begin{array}{|l|}\hline 1 \leftarrow \mathrm{x}\\ \hline 1 \leftarrow \text{John(x)}\\ \hline \\ \hline\end{array}\end{array} * p(\mathbf{x})\right)$$

When combined with an unresolved structure requiring a discourse referent, the pointer associated with ‘John’ (1 in the example above) will become a free pointer because the definition of projective merge ensures that the label of the second argument of the merge is distinct from the label of its first argument. Since projected content keeps its own pointer, it appears free and therefore projects. Note, however, that the presupposition may still become bound later on during discourse construction, when it is combined with an antecedent PDRS that matches its label. This treatment of presupposition projection as part of discourse construction is much like the treatment of anaphoric expressions in dynamic semantic formalisms, such as Dynamic Predicate Logic (Groenendijk and Stokhof, 1991). Just like in the case of anaphoric expressions, determining the projection site of a presupposition in PDRT is part of constructing its lexical semantics. This emphasizes a context-dependent view of meaning construction.

A set of lexical items together with their semantics is shown in Appendix 5.C. One of the main advantages of the PDRT representation over the traditional DRT representations is that the use of different types of merge nicely captures the correspondence between projection and non-projecting sibling items. For example, the definite determiner “the” and the indefinite determiner “a” obtain the same lexical semantic representation, except for the fact that the argument “the” is contributed to its direct context using Projective Merge, whereas the argument of “a” uses Assertive Merge. Similarly, a distinction can be made on the lexical level between the factive verb “know” and its non-factive equivalent “believe” (see Appendix 5.C, Table 5.2).

5.4 Harnessing projection using PDRT

This section illustrates how PDRT represents the information status of different types of linguistic expressions. Up to this point, the PDRT examples have mainly concerned existential presuppositions triggered by definite descriptions and proper names. Here, we show that the additional representational level of PDRT can be used to account for various other linguistic phenomena, in particular by exploiting the level of information provided by the MAPs. Moreover, we show how the interaction between Assertive Merge and Projective Merge can account for the emergence of different types of meaning.

5.4.1 Representing information status

As described above, Projective DRT extends the representational power of DRT by explicitly representing the relation between the introduction and interpretation site of linguistic content, via the use of projection variables. As illustrated above, this additional level of information can be used to represent the information status of presuppositions. Critically, the interpretational constraints contributed by the MAPs allow for representing dependencies between projected content and its local discourse context, as well as dependencies between different instances of projected content, as illustrated in (14).

(14) a. John ate the cookies.

$$
\begin{array}{c}
1 \\
\begin{array}{|l|}
\hline
2 \leftarrow x \;\; 3 \leftarrow y \\
\hline
2 \leftarrow \text{John}(x) \\
3 \leftarrow \text{cookies}(y) \\
1 \leftarrow \text{ate}(x,y) \\
\hline
1 \leq 2 \;\; 1 \leq 3 \\
\hline
\end{array}
\end{array}
$$

b. It is John who ate the cookies.

$$
\begin{array}{c}
1 \\
\begin{array}{|l|}
\hline
2 \leftarrow x \;\; 3 \leftarrow y \;\; 4 \leftarrow z \\
\hline
2 \leftarrow \text{person}(x) \\
3 \leftarrow \text{cookies}(y) \\
2 \leftarrow \text{ate}(x,y) \\
4 \leftarrow \text{John}(z) \\
1 \leftarrow x{=}z \\
\hline
1 \leq 2 \;\; 2 \leq 3 \;\; 1 \leq 4 \\
\hline
\end{array}
\end{array}
$$

Example (14a) contains two existential presuppositions, triggered by the use of the proper name "*John*" and the definite description "*the cookies*". In (14b), on the other hand, this latter presupposition is embedded as part of another presupposition, triggered by the *it*-cleft construction, namely that someone ate the cookies. This embedding is represented in the MAPs, which indicate that the interpretation site of the presupposition triggered by "*the cookies*" should be accessible from the interpretation site of the presupposition triggered by the *it*-cleft construction (i.e., $2 \leq 3$). This reflects a hierarchical ordering of presuppositions that states that the mention of the cookies is *more backgrounded* than the observation that someone ate them.

As described above, the MAP constraints for weak subordination ($\leq$) and its negation ($\nleq$) can be used to represent various stronger constraints on accessibility, such as identity (i.e., $v_1 = v_2 \Leftrightarrow v_1 \leq v_2 \wedge v_2 \leq v_1$), and strict subordination (i.e., $v_1 < v_2 \Leftrightarrow v_1 \leq v_2 \;\wedge\; v_2 \nleq v_1$). These stronger constraints can be used, for instance, to represent the contribution made by conventional implicatures, as described by Potts (2005); this is illustrated in example (15) (see Venhuizen et al., 2014b).

(15) Mary, who doesn't like John, hits him.

$$
\begin{array}{c}
1 \\
\begin{array}{|l|}
\hline
3 \leftarrow \text{x} \\
\hline
3 \leftarrow \text{Mary(x)} \\
4 \leftarrow \neg \begin{array}{c} 2 \\ \begin{array}{|l|}
\hline
5 \leftarrow \text{y} \\
\hline
5 \leftarrow \text{John(y)} \\
2 \leftarrow \text{like(x,y)} \\
\hline
2 \leq 5 \;\; 2 \leq 3 \\
\hline
\end{array} \end{array} \\
1 \leftarrow \text{hit(x,y)} \\
\hline
1 \leq 3 \;\; 1 < 4 \;\; 4 = 3 \;\; 1 \leq 5 \\
\hline
\end{array}
\end{array}
$$

This example shows the PDRS for a sentence containing a non-restrictive relative clause (NRRC). The conventional implicature triggered by the NRRC at the same time requires projection from the local context (using strict subordination: $1 < 4$), and introduces a dependency to the projection site of its anchor, "Mary" (using identity: $4 = 3$). These constraints reflect the information status of conventional implicatures, which are considered to provide novel and backgrounded information about the referent referred to by the anchor (for a more detailed explanation, see Venhuizen et al., 2014b).

5.4.2 Merge interactions

As described above, the order of the arguments in a merge operation may in some cases crucially affect the result of the merge; in the case of Projective Merge only the content of the first argument becomes projected, and both Assertive and Projective Merge allow for free variables from the second argument to become bound by content in the first argument (this includes projection variables as well as discourse referents). However, because of the asymmetry between Assertive and Projective Merge, not only the order of the arguments, but also the order of applying merge operations in sequence may affect the resulting PDRS. This effect is demonstrated below, using the three PDRSs in (16). Note that there is a strong dependency between these PDRSs in terms of the projection variables and discourse referents used between them.

(16) $$A = \begin{array}{|l|} \multicolumn{1}{c}{1} \\ \hline 1 \leftarrow \text{x} \\ \hline 1 \leftarrow \text{man(x)} \\ \hline \\ \hline \end{array} \qquad B = \begin{array}{|l|} \multicolumn{1}{c}{2} \\ \hline 2 \leftarrow \text{y} \\ \hline 2 \leftarrow \text{sister(y)} \\ 2 \leftarrow \text{of(y,x)} \\ \hline 2 \leq 1 \\ \hline \end{array} \qquad C = \begin{array}{|l|} \multicolumn{1}{c}{3} \\ \hline \\ \hline 3 \leftarrow \text{love(x,y)} \\ \hline 3 \leq 2 \\ \hline \end{array}$$

In case the PDRSs from (16) are combined as a sequence, using only Assertive or only Projective Merge, the order of applying the operations does not change the result, as long as the sequence of the arguments is not changed; that is, both Assertive Merge and Projective Merge are *associative* operations. This is illustrated in example (17).

(17) a. $A + (B + C) = (A + B) + C =$

$$\begin{array}{|l|} \multicolumn{1}{c}{3} \\ \hline 3 \leftarrow \text{x} \;\; 3 \leftarrow \text{y} \\ \hline 3 \leftarrow \text{man(x)} \\ 3 \leftarrow \text{sister(y)} \\ 3 \leftarrow \text{of(y,x)} \\ 3 \leftarrow \text{love(x,y)} \\ \hline 3 \leq 3 \\ \hline \end{array}$$

b. $A * (B * C) = (A * B) * C =$

$$\begin{array}{|l|} \multicolumn{1}{c}{3} \\ \hline 1 \leftarrow \text{x} \;\; 2 \leftarrow \text{y} \\ \hline 1 \leftarrow \text{man(x)} \\ 2 \leftarrow \text{sister(y)} \\ 2 \leftarrow \text{of(y,x)} \\ 3 \leftarrow \text{love(x,y)} \\ \hline 3 \leq 1 \;\; 3 \leq 2 \;\; 2 \leq 1 \\ \hline \end{array}$$

In (17a), applying Assertive Merge two times in a row results in a PDRS in which all content from the three arguments remains asserted, and all free variables from PDRSs B and C have become bound or renamed. The resulting PDRS corresponds to a non-presuppositional interpretation, i.e., "*A man has a sister and loves her*". In (17b), the two Projective Merge operations result in a representation in which the content from PDRSs A and B has become projected, corresponding to an interpretation that contains two presuppositions, namely that there exists some man and that he has a sister, i.e., "*The man loves his sister*".

On the other hand, if we apply Assertive Merge and Projective Merge in the same sequence, the order of applying the operations crucially affects the resulting representation; in other words, the merge operations are *not inter-associative*. This is due to the fact that Projective Merge results in the projection of its first argument, which means that the projection of the entire foregoing sequence may be affected by adapting the order of applying the operations. This is illustrated in (18).

(18) a. $(A + B) * C =$

$$\begin{array}{c} 3 \\ \begin{array}{|l|} \hline 2 \leftarrow \text{x} \;\; 2 \leftarrow \text{y} \\ \hline 2 \leftarrow \text{man(x)} \\ 2 \leftarrow \text{sister(y)} \\ 2 \leftarrow \text{of(y,x)} \\ 3 \leftarrow \text{love(x,y)} \\ \hline 3 \leq 2 \;\; 2 \leq 2 \\ \hline \end{array} \end{array}$$

b. $A + (B * C) =$

$$\begin{array}{c} 3 \\ \begin{array}{|l|} \hline 3 \leftarrow \text{x} \;\; 2 \leftarrow \text{y} \\ \hline 3 \leftarrow \text{man(x)} \\ 2 \leftarrow \text{sister(y)} \\ 2 \leftarrow \text{of(y,x)} \\ 3 \leftarrow \text{love(x,y)} \\ \hline 2 = 3 \\ \hline \end{array} \end{array}$$

The different results in (18a) and (18b) clearly illustrate that projection differences crucially affect the interpretation of a PDRS. In (18a), the content of both PDRS A and PDRS B becomes projected due to the Projective Merge with C, even though A and B are themselves combined using Assertive Merge. This results in a reading with one presupposition, which combines the content of both PDRS A and PDRS B, i.e., "*The man who has a sister loves her*". But if the arguments are combined in a different order, as in (18b), the resulting PDRS obtains a different interpretation. Because of the MAPs constraining the interpretation site of the content of PDRS B, the result of the Assertive Merge in (18b) is that the presupposition triggered by the Projective Merge becomes locally accommodated, corresponding to the reading "*A man loves his sister*".

Thus, in (18) the order of applying the merge operations matters because the combination of Assertive Merge and the constraints indicated by the MAPs may prevent content from becoming projected. If, however, the Projective Merge operation appears sequentially before the Assertive Merge operation, no such presupposition cancellation can occur, and inter-associativity is reestablished, as shown in (19).

(19) $A * (B + C) = (A * B) + C =$

$$
\begin{array}{c}
3 \\
\begin{array}{|l|}
\hline
1 \leftarrow \text{x} \quad 3 \leftarrow \text{y} \\
\hline
1 \leftarrow \text{man(x)} \\
3 \leftarrow \text{sister(y)} \\
3 \leftarrow \text{of(y,x)} \\
3 \leftarrow \text{love(x,y)} \\
\hline
3 \leq 1 \quad 3 \leq 3 \\
\hline
\end{array}
\end{array}
$$

Here, both of the permutations result in a PDRS with a single presupposition, corresponding to the interpretation " *The man has a sister and loves her*". The availability of natural language examples corresponding to the readings of each of these different PDRSs resulting from the same set of constituents, suggests that these interactions between asserted and projected content are in fact present in the way linguistic meaning is constructed.

5.5 Implications for semantic theory

In this chapter, we have presented Projective Discourse Representation Theory, an extension of traditional DRT that represents the information status of linguistic content by explicitly representing the relation between its semantic interpretation, and how it is presented in the linguistic surface structure. We have shown how the definitions that form the formal backbone of DRT can be extended to account for the additional level of information introduced by the projection variables. With this in place, PDRT has become a full-fledged semantic formalism—the question now arises what can be done with these extended representations, and how they relate to existing extensions of DRT. In what follows, we will first briefly discuss the relation of PDRT to other variations of the DRT framework, and then describe some of its implications in more detail, in particular with respect to the interpretation of the resulting semantic representations. What is more, we will lay out what we believe to be the most important directions for future research.

5.5.1 Position within the DRT family

Projective DRT introduces projection pointers for dealing with projected content within the DRT framework. Because of its intuitive and flexible nature, DRT has already formed the basis for several other extensions aimed at dealing with various linguistic phenomena. Within this family of DRT variations, PDRT unifies the intuitions behind two major directions; (i) the incorporation of an additional level of meaning representation to account for, in particular, the distinct information status of presuppositions (other DRT variants exploiting this principle are, e.g., Presuppositional DRT, as proposed by Krahmer, 1998, Layered DRT, proposed by Geurts and Maier, 2003, 2013, and the treatment of indexicals proposed by Hunter, 2010, 2013), and (ii) the use of specific identifiers for individual discourse contexts and the accessibility relations between them (this concept is used, for example, in the Unresolved DRSs (UDRSs) defined by Reyle, 1993, 1995, but also in the Segmented DRSs (SDRSs) from Segmented DRT, following Asher, 1993; Asher and Lascarides, 2003).

With respect to the class of frameworks dealing with additional levels of meaning, PDRT positions itself right between Presuppositional DRT, defined specifically for dealing with presuppositions, and Layered DRT, which treats different levels of linguistic meaning as separate, non-interacting layers of linguistic content. In contrast, PDRT provides a unified analysis of different projection phenomena within a single dimension, by using projection variables to indicate the interpretation site of linguistic content. As we will see in Section 5.5.2 below, this analysis is highly compatible with the treatment of indexical expressions proposed by Hunter (2010, 2013).

The projection variables of PDRT are used to identify specific discourse contexts, in much the same way that UDRSs and SDRSs obtain an identifier. Critically, however, in PDRT the identifiers are part of the basic structures in the form of variables that behave in the same way as discourse referents in DRT, in the sense that they may appear free or bound, and can be renamed during discourse construction. By contrast, the identifiers used in UDRSs are constants, and therefore do not adhere to the same principles as discourse variables. On the other hand, the identifiers in SDRT are also variables, but in addition obtain a model-theoretic interpretation. This is not the case for the projection variables from PDRT, which serve a representational purpose in explicating the interaction between the linguistic surface structure and the semantic representation (but see Section 5.5.2 for different ways of incorporating the information contributed by the projection variables into the model-theoretic interpretation of PDRSs).

Of particular interest with respect to PDRT is Schlenker's (2011) version of DRT with local contexts, which aims to incorporate the satisfaction-based view on presuppositions as proposed by Heim (1983) into the DRT formalism. Schlenker employs variables to indicate 'local contexts' and uses coindexation for presuppositions to indicate the contexts in which they must be satisfied. Satisfaction here means that the context indicated by the presupposition should entail it. This is where the account crucially differs from PDRT: in Schlenker's local contexts analysis, presuppositions merely impose *constraints* on the interpretation of the context in which they are accommodated. By contrast, projected content in PDRT inherits all referential and interpretational properties from its accommodation site. This means, for example, that the accounts make different predictions about the interpretation of referring expressions that are more descriptive than their antecedent, as illustrated in example (20).

(20) A woman was arrested in Berlin yesterday. <u>The mother of three</u> is suspected of murder.

Here, the underlined definite description "*the mother of three*" simultaneously co-refers with the indefinite description "*a woman*" and provides more information about its referent. On Schlenker's account, the presupposition triggered by the definite description requires the context in which the indefinite "*a woman*" is introduced to entail that this woman is a mother of three. In the PDRT analysis, on the other hand, this information is attached to the description of "*a woman*", and therefore becomes part of the novel contribution of the discourse.

5.5.2 Interpretation of Projected DRSs

One of the strengths of the traditional DRT framework—which we have omitted in this chapter thus far—is that its basic structures have a model-theoretic interpretation, providing them with a truth-conditional semantics. The model-theoretic interpretation can be derived in two ways. Firstly, one can formulate a translation from DRSs into first-order logic. Such translations, as provided, for example, in Bos (2003), assume a variant of DRT that only contains conditions that can be formulated as first-order logic formulas. A second way of deriving a model-theoretic interpretation for DRSs, is by providing a direct interpretation using an *embedding function* that verifies a DRS in a given model. Intuitively, a DRS is considered to be a *partial* model representing the information conveyed by some piece of discourse; determining the truth of DRS K in some world w, therefore means finding a function that embeds K in the *total* model M with respect to w (for a formal definition of the embedding function for DRSs, see, e.g., Kamp and Reyle, 1993; Geurts and Beaver, 2011).

As we have shown extensively in this chapter, PDRT provides a richer semantic formalism than its predecessor DRT, since PDRT incorporates explicit information about how the information in a discourse is structured, in particular regarding the introduction site of projected material. With respect to the model-theoretic interpretation of PDRT, it depends on one's interests whether this informational extension should also be percolated through to the model. On the one hand, PDRSs can be translated into DRSs, as shown in Venhuizen et al. (2013b). This demonstrates that PDRSs—albeit richer representations—inherit all interpretational properties from DRSs, including the translation to first-order logic. However, the translation from PDRT to DRT comes at the cost of losing information about how the information is structured within the discourse (e.g., the status of presupposed material), since during translation all content is moved to the interpretation site indicated by the pointers. So, the additional level of information provided by the projection variables can only be incorporated into the model-theoretic interpretation using the direct interpretation via an embedding function. There are various ways in which the interpretation function can be extended in order to incorporate different aspects of meaning that can be represented in PDRT. Below, we informally describe three different ways of incorporating different dimensions of meaning into the PDRT interpretations.

- **Presupposition failure using three-valued logic.** Since Strawson's (1950) famous argument against Russell (1905), in which he states that sentences containing a definite description that fails to refer (i.e., presupposition failure) should be interpreted as lacking a truth-value, different versions of three-valued logics have been proposed to account for these undefined interpretations. In the context of DRT, Krahmer (1998) has proposed a version of DRT that has a three-valued interpretation (based on *middle Kleene*; Kleene, 1952). PDRT also allows for an interpretation along these lines, since presuppositions are discernible from asserted content in the final representations, namely via the use of free pointers.

- **Anchoring linguistic content to the extra-linguistic context.** In order to account for the interpretational properties of indexical expressions, such as *I* and *here*, Hunter (2013) proposes a DRT analysis in which indexicals, and other presupposition triggers, are accommodated to a special global DRS context, called K_0, which is evaluated with respect to the extra-linguistic context, i.e., the actual world (see also, Hunter and Asher, 2005). Such an analysis is highly compatible with the PDRT approach, since the use of free pointers for presuppositions already suggests their readiness to be bound by the label of some higher (possibly extra-linguistic) context. In fact, the availability of such a context is already assumed by the PDRT analysis of conventional implicatures proposed in Venhuizen et al. (2014b), as CIs require their anchor to project out of the local context via strict subordination, even in case the local context is itself the global context of the discourse, as shown in example (15) above.

- **Speaker-orientedness via subjective models.** Besides providing a special interpretation for presuppositions and indexicals, the embedding function of PDRT may also be extended to provide an adapted treatment for other phenomena, such as conventional implicatures. As already proposed in Venhuizen et al. (2014b), the projection pointers from PDRT may be exploited to incorporate information about the speaker-oriented nature of CIs into the model. CIs are generally interpreted as contributing subjective content that is attributed to the speaker of the utterance. This information may be reflected in PDRT by the use of special pointers that indicate a context in a *subjective model*, which is the state of affairs in the world according to some specific discourse agent, e.g., the speaker, the hearer, or some agent introduced in the discourse (see Venhuizen et al., 2014b, for more details).

5.5.3 PDRT applications and outlook

The representations from PDRT are richer than the traditional DRT representations. This means that they are more expressive, in the sense that more aspects of linguistic meaning can be captured within these representations. In particular, we have focused on representing the difference in information status between asserted and projected content. However, this expressiveness comes at the cost of adding extra variables to the representations, which arguably increases their complexity. An important question is therefore what is gained exactly with this additional level of information. And, from a more philosophical perspective, how far should we go in complicating the semantic representations with the goal of incorporating more detail into the analyses?

Firstly, it should be noted that the variables that Projective DRT employs in order to to indicate the interpretation site of linguistic content are defined using constraints that are already available in traditional DRT; as was demonstrated in the definitions above, projection variables behave just like traditional DRS referents in terms of their binding properties, and the way in which they can be resolved. This means that no additional machinery is required to account for the behavior of these variables. Moreover, the introduction of projection variables results in a simpler construction procedure for representations containing projected content than the one proposed by van der Sandt (1992), since presuppositions do not need to be resolved *post hoc*, as the projection variables can become bound already during discourse construction. The result is a purely *dynamic* analysis of projection, in which projection triggers provide content with 'projection potential' on the lexical level—using Projective Merge—and the context determines where the content will be accommodated by binding the projection variables during discourse construction.

With respect to the applicability of the representations from PDRT, and its position within the broader enterprise of semantic theory, it is important to consider the theoretical as well as the practical perspective. From a theoretical point of view, PDRT opens up the way to model and investigate the semantic properties of different linguistic phenomena. As was already shown by the analysis of conventional implicatures presented in Venhuizen et al. (2014b), formalizing the behavior of specific linguistic phenomena in a semantic framework contributes significantly to the understanding of the semantic properties underlying this behavior. For instance, PDRT could be employed to gain insight into the projection behavior of different kinds of projection triggers; for instance, PDRT could be used to investigate the different contextual constraints underlying the behavior of *weak* and *strong* presupposition triggers. Conversely, as PDRT aims to treat projection as a property that is inherent to the way in which discourse representations are constructed, the PDRT analysis might contribute to the development of a unified analysis of projected content, in line with analyses that aim to explain projection in terms of *at-issueness* (Simons et al., 2010; Tonhauser et al., 2013). Finally, the representations from PDRT may be used to investigate aspects of meaning beyond projection phenomena. The additional level of information available in the representations of PDRT allows for the formalization of different syntactic constructions, and their interaction with linguistic meaning. For example, the MAPs may be employed to explicitly represent the notion of *givenness* as an ordering of projection sites; this has been shown to critically affect the choice of syntactic structure in, e.g., dative alternation Bresnan et al. (see 2007) and genitive alternation (see Rosenbach, 2014).

From a practical point of view, the direct correspondence between the representations of PDRT and the linguistic surface structure make it an attractive semantic formalism for the purpose of natural language generation (see, e.g., Basile and Bos, 2013). Moreover, the formalization of the construction and interpretation procedure make PDRT a suitable formalism for computational applications. As described above, the formal definitions of PDRT have been implemented as part of PDRT-SANDBOX, a widely applicable NLP library (Venhuizen and Brouwer, 2014). PDRT also provides the formal backbone underlying the semantic representations in the Groningen Meaning Bank (GMB; Basile et al., 2012a; Bos et al., 2015). These existing implementations make PDRT a practically useful semantic framework for investigating linguistic phenomena using large-scale computational methods, such as is common in modern day computational linguistics. As a straightforward working example on the applicability of the PDRT analysis for gaining insight into semantic behavior, we are currently gathering data from the GMB on the projection behavior of definite noun phrases (e.g., *accommodation* versus *binding*). We aim to investigate whether we can use the richer semantic representations of PDRT to derive specific semantic features that influence this behavior, in order to obtain more fine-grained insights into the projection behavior of definite descriptions.

In sum, the PDRT analysis provides a rich representational scheme for formalizing and investigating a variety of linguistic phenomena. The addition of projection variables to discourse representation structures has been shown to provide a maximal information gain into the semantic representation, using minimal additional machinery. The formalism employs traditional semantic notions such as variable binding to account for aspects of meaning that were not traditionally considered part of the semantic contribution of an utterance. As such, PDRT opens up new directions for the investigation of linguistic meaning, where non-truth-conditional aspects of meaning (e.g., information status) may contribute to and interact with model-theoretic semantic interpretations.

Appendices

5.A DRT: Definitions

These definitions are based on Kamp and Reyle (1993); Bos (2003); Kamp et al. (2011).

Definition 22 (DRS)**.** A Basic DRS: $\langle\{x_1...x_i\},\{\gamma_1...\gamma_j\}\rangle$ is a tuple, where:

i. $\{x_1...x_i\}$ is a finite set of variables;

ii. $\{\gamma_1...\gamma_j\}$ is a finite set of DRS conditions (which may be either basic or complex);

iii. $R(x_1,...,x_n)$ is a basic DRS condition, with $x_1...x_n$ are variables and R is a relation symbol for an n-place predicate;

iv. $\neg K$, $\Box K$ and $\Diamond K$ are complex DRS conditions, with K is a DRS;

v. $K_1 \vee K_2$ and $K_1 \Rightarrow K_2$ are complex DRS conditions, with K_1 and K_2 are DRSs;

vi. $x : K$ is a complex DRS condition, with x is a variable and K is a DRS.

Definition 23 (DRS accessible universes)**.** The universe of DRS K_j is accessible from (a referent introduced in) DRS K_i ($K_i \leq K_j$) iff:

- $K_i = K_j$;
- $K_i < K_j$ (K_i is subordinated by K_j; see Definition 24).

Definition 24 (DRS Subordination)**.** DRS K_1 is subordinated by DRS K_2 ($K_1 < K_2$) iff:

- $\neg K_1$, $\Box K_1$ or $\Diamond K_1$ is a DRS condition of K_2;
- $x : K_1$ is a DRS condition of K_2 for some x;
- $K_1 \Rightarrow K_3$, $K_3 \Rightarrow K_1$, $K_1 \vee K_3$, or $K_3 \vee K_1$ is a DRS condition in K_2 for some DRS K_3;
- There is a DRS K_3, such that $K_2 \Rightarrow K_1$ is a DRS condition in K_3;
- There is a DRS K_3, such that $K_1 < K_3$ and $K_3 < K_2$.

Definition 25 (DRS Variable Binding)**.** DRS variable x, introduced in DRS K_i, is bound in global DRS K iff there exists a DRS $K_j \leq K$, such that:

i. $K_i \leq K_j$;

ii. $x \in U(K_j)$, where $U(K_j)$ refers to the universe of DRS K_j.

Definition 26 (Free DRS Variables)**.**

i. $\mathcal{F}(\langle U, C\rangle) = (\bigcup_{c \in C} \mathcal{F}(c)) - U$;

ii. $\mathcal{F}(R(x_1, ..., x_n)) = \{x_1, ..., x_n\}$;

iii. $\mathcal{F}(\neg K) = \mathcal{F}(\Diamond K) = \mathcal{F}(\Box K) = \mathcal{F}(K)$;

iv. $\mathcal{F}(K_1 \Rightarrow K_2) = \mathcal{F}(K_1) \cup (\mathcal{F}(K_2) - U(K_1))$;

v. $\mathcal{F}(K_1 \vee K_2) = \mathcal{F}(K_1) \cup \mathcal{F}(K_2)$;

vi. $\mathcal{F}(x : K) = \{x\} \cup \mathcal{F}(K)$.

5.B PDRT: Additional Definitions

5.B.1 Structure

Definition 27 (Projected Referents)**.**

i. $\mathcal{R}(\langle l, U, C, M\rangle) = U \cup \bigcup_{c \in C} \mathcal{R}(c))$

ii. $\mathcal{R}(p \leftarrow R(x_1, ..., x_n)) = \bigcup_{x \in \{x_1,...,x_n\}} \{p \leftarrow x\}$

iii. $\mathcal{R}(p \leftarrow \neg K) = \mathcal{R}(p \leftarrow \Diamond K) = \mathcal{R}(p \leftarrow \Box K) = \mathcal{R}(K)$

iv. $\mathcal{R}(p \leftarrow K_1 \Rightarrow K_2) = \mathcal{R}(p \leftarrow K_1 \vee K_2) = \mathcal{R}(K_1) \cup \mathcal{R}(K_2)$

v. $\mathcal{R}(p \leftarrow x : K) = \{p \leftarrow x\} \cup \mathcal{R}(K)$

Definition 28 (PDRS to Projection Graph)**.**

i. $pg(\langle l, U, C, M\rangle) = \{\langle l, l\rangle \mapsto +\} \cup \bigcup_{v \leftarrow x \in U} \{\langle l, v\rangle \mapsto +\} \cup \bigcup_{c \in C} pg(l, c) \cup \bigcup_{m \in M} pg(l, m)$

ii. $pg(l, v \leftarrow R(x_1, ..., x_n)) = \{\langle l, v\rangle \mapsto +\}$

iii. $pg(l, v \leftarrow \neg P) = pg(l, v \leftarrow \Diamond P) = pg(l, v \leftarrow \Box P) = pg(l, v \leftarrow x : P)$ $= \{\langle l, v\rangle \mapsto +, \langle \boldsymbol{lab}(P), l\rangle \mapsto +, \langle l, \boldsymbol{lab}(P)\rangle \mapsto -\} \cup pg(P)$

iv. $pg(l, v \leftarrow P_1 \vee P_2) = \{\langle l, v\rangle \mapsto +, \langle \boldsymbol{lab}(P_1), l\rangle \mapsto +, \langle l, \boldsymbol{lab}(P_1)\rangle \mapsto -, \langle \boldsymbol{lab}(P_2), l\rangle \mapsto +, \langle l, \boldsymbol{lab}(P_2)\rangle \mapsto -\} \cup pg(P_1) \cup pg(P_2) \cup \{\langle \boldsymbol{lab}(P_1), \boldsymbol{lab}(P_2)\rangle \mapsto -, \langle \boldsymbol{lab}(P_2), \boldsymbol{lab}(P_1)\rangle \mapsto -\}$

v. $pg(l, v \leftarrow P_1 \Rightarrow P_2) = \{\langle l, v\rangle \mapsto +, \langle \boldsymbol{lab}(P_1), l\rangle \mapsto +, \langle l, \boldsymbol{lab}(P_1)\rangle \mapsto -, \langle \boldsymbol{lab}(P_2), l\rangle \mapsto +, \langle l, \boldsymbol{lab}(P_2)\rangle \mapsto -\} \cup pg(P_1) \cup pg(P_2) \cup \{\langle \boldsymbol{lab}(P_1), \boldsymbol{lab}(P_2)\rangle \mapsto -, \langle \boldsymbol{lab}(P_2), \boldsymbol{lab}(P_1)\rangle \mapsto +\}$

vi. $pg(l, v_1 \leq v_2) = \{\langle l, v_1\rangle \mapsto +, \langle l, v_2\rangle \mapsto +, \langle v_1, v_2\rangle \mapsto +\}$

vii. $pg(l, v_1 \not\leq v_2) = \{\langle l, v_1\rangle \mapsto +, \langle l, v_2\rangle \mapsto +, \langle v_1, v_2\rangle \mapsto -\}$

5.B.2 Variable Binding

In the definitions below, *freepvar*(p, l, P) means that projection variable p, introduced in a PDRS labeled l, is *not* bound in PDRS P (that is, *freepvar*$(p, l, P) = \neg$*boundpvar*(p, l, P)). Analogously, *freepref*$(p \leftarrow x, l, P)\}$ means that projected referent $p \leftarrow x$, introduced in a PDRS labeled l, is *not* bound in PDRS P (that is, *freepref*$(p \leftarrow x, l, P)\} = \neg$*boundpref*$(p \leftarrow x, l, P)\}$).

Definition 29 (Free Projection Variables)**.** Given a global PDRS P, we can define the set of free projection variables of PDRS $P' = \langle l, U, C, M \rangle$ as follows:

i. $\mathcal{F}_\pi(\langle l, U, C, M \rangle) = \{p \mid p \leftarrow u \in U \wedge \textit{freepvar}(p, l, P)\} \cup \bigcup_{c \in C} \mathcal{F}_\pi(c, l) \cup \bigcup_{\mu \in M} \mathcal{F}_\pi(m, l)$

ii. $\mathcal{F}_\pi(p \leftarrow R(...), l) = \{p \mid \textit{freepvar}(p, l, P)\}$

iii. $\mathcal{F}_\pi(p \leftarrow \neg P_1, l) = \mathcal{F}_\pi(p \leftarrow \Diamond P_1, l) = \mathcal{F}_\pi(p \leftarrow \Box P_1, l) = \mathcal{F}_\pi(p \leftarrow x : P_1, l) =$
$\{p \mid \textit{freepvar}(p, l, P)\} \cup \mathcal{F}_\pi(P_1)$

iv. $\mathcal{F}_\pi(p \leftarrow P_1 \Rightarrow P_2, l) = \mathcal{F}_\pi(p \leftarrow P_1 \vee P_2, l) = \{p \mid \textit{freepvar}(p, l, P)\} \cup \mathcal{F}_\pi(P_1) \cup \mathcal{F}_\pi(P_2)$

v. $\mathcal{F}_\pi(p_1 \leq p_2, l) = \{\textit{freepvar}(p_1, l, P)\} \cup \{\textit{freepvar}(p_2, l, P)\}$

Definition 30 (Free Projected Referents)**.** Given a global PDRS P, we can define the set of free projected referents of PDRS $P' = \langle l, U, C, M \rangle$ as follows:

i. $\mathcal{F}_\mathcal{R}(\langle l, U, C, M \rangle) = \bigcup_{c \in C} \mathcal{F}_\mathcal{R}(c, l)$

ii. $\mathcal{F}_\mathcal{R}(p \leftarrow R(x_1, ..., x_n), l) = \{p \leftarrow x_i \mid x_i \in \{x_1, ..., x_n\} \wedge \textit{freepref}(p \leftarrow x_i, l, P)\}$

iii. $\mathcal{F}_\mathcal{R}(p \leftarrow \neg P_1, l) = \mathcal{F}_\mathcal{R}(p \leftarrow \Diamond P_1, l) = \mathcal{F}_\mathcal{R}(p \leftarrow \Box P_1, l) = \mathcal{F}_\mathcal{R}(P_1)$

iv. $\mathcal{F}_\mathcal{R}(p \leftarrow x : P_1, l) = \{p \leftarrow x \mid \textit{freepref}(p \leftarrow x, l, P)\} \cup \mathcal{F}_\mathcal{R}(P_1)$

v. $\mathcal{F}_\mathcal{R}(p \leftarrow P_1 \Rightarrow P_2, l) = \mathcal{F}_\mathcal{R}(p \leftarrow P_1 \vee P_2, l) = \mathcal{F}_\mathcal{R}(P_1) \cup \mathcal{F}_\mathcal{R}(P_2)$

5.B.3 Composition

Definition 31 (PDRS Renaming)**.** Given the function: $\textit{rename}(P, \Pi, \Delta)$, let:

$$\Pi_T = \bigcup_{i=1}^{|\Pi(P)|} \{(p_i, p'_i)\} \text{ where } p_i \in \Pi(P) \text{ and } p'_i \in \textit{newpvars}(\Pi(P))$$
$$\Delta_T = \bigcup_{i=1}^{|\Delta(P)|} \{(x_i, x'_i)\} \text{ where } x_i \in \Delta(P) \text{ and } x'_i \in \textit{newdrefs}(\Delta(P))$$

where *newpvars* is a function that takes a set of projection variables and returns a set of the same length containing novel projection variables, and *newdrefs* does the same for discourse referents.

Now, define $\textit{rename}'(P)$ as follows:

i. $\textit{rename}'(\langle l, U, C, M\rangle) = \langle l', U', C', M'\rangle$
where: $l' = \textit{trpvar}(l, l)$
$U' = \bigcup_{p \leftarrow r \in U} (\textit{trpvar}(p, l) \leftarrow \textit{trdref}(r, p, l))$
$C' = \bigcup_{c \in C} \textit{rename}'(c, l)$
$M' = \bigcup_{p_1 \leq p_2 \in M} \textit{trpvar}(p_1, l) \leq \textit{trpvar}(p_2, l)$
$\cup \bigcup_{p_1 \not\leq p_2 \in M} \textit{trpvar}(p_1, l) \not\leq \textit{trpvar}(p_2, l)$

ii. $\textit{rename}'(p \leftarrow \textit{Rel}(x_1, \ldots, x_n), l) = p' \leftarrow \textit{Rel}(x'_1, \ldots, x'_n)$

iii. $\textit{rename}'(p \leftarrow \neg P_1, l) = p' \leftarrow \neg(\textit{rename}'(P_1))$

iv. $\textit{rename}'(p \leftarrow \Box P_1, l) = p' \leftarrow \Box(\textit{rename}'(P_1))$

v. $\textit{rename}'(p \leftarrow \Diamond P_1, l) = p' \leftarrow \Diamond(\textit{rename}'(P_1))$

vi. $\textit{rename}'(p \leftarrow P_1 \vee P_2, l) = p' \leftarrow (\textit{rename}'(P_1) \vee \textit{rename}'(P_2))$

vii. $\textit{rename}'(p \leftarrow P_1 \Rightarrow P_2, l) = p' \leftarrow (\textit{rename}'(P_1) \Rightarrow \textit{rename}'(P_2))$

viii. $\textit{rename}'(p \leftarrow x : P_1, l) = p' \leftarrow (x' : \textit{rename}'(P_1))$

where: $p' = \textit{trpvar}(p, l)$
$x'_i = \textit{trdref}(x_i, p, l)$

The translate function *trpvar*(p, l) determines for a projection variable p whether it is free relative to its local context l, given global PDRS P. If this is not the case, p is translated based on the set of tuples of projection variables Π_T.

$$\textit{trpvar}(p, l) = \begin{cases} p & \text{if } \textit{freepvar}(p, l, P) \\ p' & \text{otherwise, where: } (p, p') \in \Pi_T \end{cases}$$

Similarly, *trdref*(x, p, l) determines whether a discourse referent x is free relative to pointer p and local context l, given global PDRS P. If this is not the case, x is translated based on the set of tuples of discourse referents Δ_T.

$$\textit{trdref}(x, p, l) = \begin{cases} x & \text{if } \textit{freepref}(p \leftarrow x, l, P) \\ x' & \text{otherwise, where: } (x, x') \in \Delta_T \end{cases}$$

5.C PDRT: Lexical semantics

Tables 5.1 and 5.2 shows the lexical semantics for a set of English lexical items, following the lexical DRT semantics proposed by Bos (2003). These lexical items employ a neo-Davidsonian event semantics, with thematic roles from VerbNet (Kipper et al., 2008).

Table 5.1: PDRT representations for a set of lexical items

<table>
<tr><th>Item</th><th>PDRT Semantics</th></tr>
<tr><td>$\mathrm{a}_{\langle et,\langle et,t\rangle\rangle}$</td><td>$\lambda p.\lambda q.((\overset{1}{\begin{array}{|l|}\hline 1 \leftarrow \mathrm{x} \\ \hline \\ \hline \\ \hline \end{array}} + p(\mathbf{x})) + q(\mathbf{x}))$</td></tr>
<tr><td>$\mathrm{the}_{\langle et,\langle et,t\rangle\rangle}$</td><td>$\lambda p.\lambda q.((\overset{1}{\begin{array}{|l|}\hline 1 \leftarrow \mathrm{x} \\ \hline \\ \hline \\ \hline \end{array}} + p(\mathbf{x})) * q(\mathbf{x}))$</td></tr>
<tr><td>$\mathrm{dog}_{\langle e,t\rangle}$</td><td>$\lambda x.(\overset{1}{\begin{array}{|l|}\hline \\ \hline 1 \leftarrow \mathrm{dog}(x) \\ \hline \\ \hline \end{array}})$</td></tr>
<tr><td>$\mathrm{John}_{\langle et,t\rangle}$</td><td>$\lambda p.(\overset{1}{\begin{array}{|l|}\hline 1 \leftarrow \mathrm{x} \\ \hline 1 \leftarrow \mathrm{John(x)} \\ \hline \\ \hline \end{array}} * p(\mathbf{x}))$</td></tr>
<tr><td>$\mathrm{someone}_{\langle et,t\rangle}$</td><td>$\lambda p.(\overset{1}{\begin{array}{|l|}\hline 1 \leftarrow \mathrm{x} \\ \hline 1 \leftarrow \mathrm{person(x)} \\ \hline \\ \hline \end{array}} + p(\mathbf{x}))$</td></tr>
<tr><td>$\mathrm{happy}_{\langle et,et\rangle}$</td><td>$\lambda p.\lambda x.(\overset{1}{\begin{array}{|l|}\hline 1 \leftarrow \mathrm{e} \\ \hline 1 \leftarrow \mathrm{happy(e)} \\ 1 \leftarrow \mathrm{Patient}(\mathrm{e},x) \\ \hline \\ \hline \end{array}} + p(x))$</td></tr>
<tr><td>$\mathrm{his}_{\langle et,\langle\langle et,t\rangle,\langle et,t\rangle\rangle\rangle}$</td><td>$\lambda p.\lambda q.\lambda r.((q(\lambda x.(\overset{1}{\begin{array}{|l|}\hline 1 \leftarrow \mathrm{y} \\ \hline 1 \leftarrow \mathrm{male}(x) \\ 1 \leftarrow \mathrm{of}(\mathrm{y},x) \\ \hline \\ \hline \end{array}})) + p(\mathbf{y})) * r(\mathbf{y}))$</td></tr>
</table>

Table 5.2: PDRT representations for a set of lexical items (cont.)

<table>
<tr><th>Item</th><th>PDRT Semantics</th></tr>
<tr><td>bark
$\langle\langle et,t\rangle,\langle et,t\rangle\rangle$</td><td>$\lambda p.\lambda q.(p(\lambda x.(\overset{1}{\begin{array}{|l|}\hline 1 \leftarrow \text{e} \\ \hline 1 \leftarrow \text{walk(e)} \\ 1 \leftarrow \text{Theme}(\text{e},x) \\ \hline \\ \hline \end{array}} + q(\text{e}))))$</td></tr>
<tr><td>love
$\langle\langle et,t\rangle,\langle\langle et,t\rangle,\langle et,t\rangle\rangle\rangle$</td><td>$\lambda p.\lambda q.\lambda r.(q(\lambda x.(p(\lambda y.(\overset{1}{\begin{array}{|l|}\hline 1 \leftarrow \text{e} \\ \hline 1 \leftarrow \text{love(e)} \\ 1 \leftarrow \text{Experiencer}(\text{e},x) \\ 1 \leftarrow \text{Stimulus}(\text{e},y) \\ \hline \\ \hline \end{array}} + r(\text{e}))))))$</td></tr>
<tr><td>know
$\langle\langle et,t\rangle,\langle\langle et,t\rangle,\langle et,t\rangle\rangle\rangle$</td><td>$\lambda p.\lambda q.\lambda r.(q(\lambda x.($
$\overset{1}{\begin{array}{|l|}\hline 1 \leftarrow \text{e} \quad 1 \leftarrow \text{k} \\ \hline 1 \leftarrow \text{know(e)} \\ 1 \leftarrow \text{Agent}(\text{e},x) \\ 1 \leftarrow \text{Theme}(\text{e,k}) \\ 1 \leftarrow \text{k}\text{: } p(\lambda y.(\overset{2}{\begin{array}{|c|}\hline \; \\ \hline \; \\ \hline \; \\ \hline \end{array}})) * p(\lambda z.(\overset{3}{\begin{array}{|c|}\hline \; \\ \hline \; \\ \hline \; \\ \hline \end{array}})) \\ \hline \\ \hline \end{array}} + r(\text{e}))))$</td></tr>
<tr><td>believe
$\langle\langle et,t\rangle,\langle\langle et,t\rangle,\langle et,t\rangle\rangle\rangle$</td><td>$\lambda p.\lambda q.\lambda r.(q(\lambda x.($
$\overset{1}{\begin{array}{|l|}\hline 1 \leftarrow \text{e} \quad 1 \leftarrow \text{k} \\ \hline 1 \leftarrow \text{believe(e)} \\ 1 \leftarrow \text{Agent}(\text{e},x) \\ 1 \leftarrow \text{Theme}(\text{e,k}) \\ 1 \leftarrow \text{k}\text{: } p(\lambda y.(\overset{2}{\begin{array}{|c|}\hline \; \\ \hline \; \\ \hline \; \\ \hline \end{array}})) + p(\lambda z.(\overset{3}{\begin{array}{|c|}\hline \; \\ \hline \; \\ \hline \; \\ \hline \end{array}})) \\ \hline \\ \hline \end{array}} + r(\text{e}))))$</td></tr>
</table>

Part III

Data-driven Formal Semantics

Chapter 6

The Groningen Meaning Bank*

Abstract. As a preliminary to data-driven semantic analysis, a resource is required that provides semantic annotations for a large amount of data. In this chapter, we describe the development of such a resource: the Groningen Meaning Bank (GMB). The GMB project aims at annotating a wide variety of linguistic phenomena, and integrating them into a single representational framework, namely, the structures provided by (Projective) Discourse Representation Theory. In this chapter, we describe the various levels of annotation available in the GMB, as well as the concept of *meaning banking*: creating a semantic resource and improving the existing annotations in a bootstrapping fashion. As the GMB is part of an ongoing research project, we here evaluate its current status and describe some directions for future work.

6.1 Introduction

The development of linguistic resources in the form of syntactic treebanks, such as the Penn TreeBank (Marcus et al., 1993), has inspired

*For a more elaborate treatment of the information presented in this chapter, see: Bos, Johan, Valerio Basile, Kilian Evang, Noortje J. Venhuizen, and Johannes Bjerva (2015). The Groningen Meaning Bank. In Ide, Nancy and James Pustejovsky, editors, *Handbook of Linguistic Annotation – Part Two: Case studies*, Text, Speech and Language Technology. Springer. (forthcoming).

much research in statistical parsing methods. In the field of semantics, on the other hand, such data-driven approaches are much less prevalent, due to the limited availability of semantically annotated resources. Most of the existing semantic corpora focus on specific semantic phenomena, such as semantic roles (PropBank; Palmer et al., 2005), semantic frames (FrameNet; Baker et al., 1998), or discourse structure (the Penn Discourse Tree-Bank; Miltsakaki et al., 2004; Prasad et al., 2008). A notable exception is OntoNotes (Hovy et al., 2006), which combines syntactic annotations (in the style of the Penn Treebank), predicate-argument structure (based on PropBank), word senses, and co-reference annotation. In this chapter, we present the Groningen Meaning Bank (GMB), a large collection of *texts* (rather than isolated sentences) that not only combines various levels of linguistic annotation, but also provides a '*deep*' level of formal meaning representation that integrates the annotations from these layers into a single semantic formalism. The different levels of syntactic and semantic annotation in the GMB include, e.g., named entities, word senses, tense, events, thematic roles, rhetorical relations, presuppositions, and coreference relations. The semantic formalism used to represent the deep semantic annotations is based on the foundations of Discourse Representation Theory (DRT; Kamp, 1981; Kamp and Reyle, 1993); this framework is particularly suitable for providing a formal backbone for analyzing these phenomena, as it consists of a *representational* dimension on top of the model-theoretic semantic interpretations, which can be exploited to incorporate various non-truth-conditional aspects of linguistic meaning, such as rhetorical structure (Asher and Lascarides, 2003) and information status (Venhuizen et al., 2015).

Critically, semantic phenomena often involve discourse-level analysis, in contrast to syntactic parsing, which can mostly be done on the sentence-level. Therefore, the units of analysis in the GMB are *texts*, rather than isolated sentences, as is common in syntactic treebanks. Besides providing a means of resolving discourse-level phenomena, such as cross-sentential anaphoric dependencies, the analysis of larger texts also enables one to resolve sentence-level ambiguities based on the larger discourse context. For instance, the discourse context may determine whether a specific noun is treated as an animate or as an inanimate entity, which, in turn may aid the syntactic parser in determining the right argument structure.

In order to obtain large-scale, high quality annotations, the GMB employs an annotation method that combines automatic semantic analysis with human annotations; we call this *human-aided machine annotation*. With this method, existing natural language processing software performs the bulk of the annotations, which are then verified by humans via two main platforms: a wiki-like interface, called the "GMB Explorer", which allows expert linguists to directly modify the annotations, and a crowd-sourcing platform, called "Wordrobe", which is a collection of 'Games with a Purpose' for different levels of linguistic annotation. The annotations obtained from these sources are used for bootstrapping the automatic analyses; by re-training the NLP software on the corrected data, the tools as well as the analyses of novel texts are improved. This way, both quantity and quality of annotation is ensured.

In section 6.2, we describe the various levels of analysis incorporated in the GMB in more detail, and section 6.3 describes the process of collecting and improving the linguistic annotations. Finally, section 6.4 presents an evaluation of the current state of the GMB, and provides some directions for future research. For a more detailed and elaborate discussion of the information presented in this chapter, the interested reader is referred to Bos et al. (2015). An earlier description of the GMB and its development can be found in Basile et al. (2012a).

6.2 Levels of annotation

The aim of the Groningen Meaning Bank is to incorporate many layers of linguistic annotation into a single semantic formalism. Critically, the semantic representations in the GMB depend upon a variety of underlying analyses, including part-of-speech (POS) tagging, syntactic analysis and the resolution of scopal properties and co-reference. Below we describe all levels of analysis that are represented in the GMB, divided over three levels of annotation: the token-level, the sentence-level and the discourse-level.

6.2.1 Tokens

The smallest units of linguistic annotation in the GMB are *tokens*, representing the individual words and punctuation characters in a text. Therefore, the first step of analysis in the GMB is identifying these tokens, and accordingly tagging these tokens using various syntactic and semantic analyses.

Tokenization. The separation of texts into word tokens and sentence tokens is done using a variant of the IOB tagging scheme, which is annotated on the character level (Evang et al., 2013). This scheme indicates for each character whether it occurs inside a token (I), outside of a token (O), at the beginning of a token (T), or at the beginning of a sentence (S). This provides a flexible tagging scheme that can account for, e.g., discontinuous tokens, such as hyphenated words at line-breaks, and contractions, such as in the case of English negations like "*did*|*n't*".

Syntactic categories. After the tokenization, all tokens are annotated with part-of-speech (POS) categories, which critically influence the higher-level syntactic and semantic analyses. The set of POS categories in the GMB is based on the tag-set introduced in the Penn TreeBank (Marcus et al., 1993) and later adopted and extended in the CCGBank (Hockenmaier and Steedman, 2007).

Based on the POS annotations, each word is assigned a syntactic category from the syntactic framework of Combinatory Categorical Grammar (CCG; Steedman, 2001). The advantage of using a categorical grammar for the development of the GMB is that it is lexically driven; all combinatory information is stored in the categories assigned to the word tokens, such that they can be mapped onto unique semantic types. This way, the lexical semantic representations can be derived directly from the CCG categories. There are two types of categories in CCG: *base* categories and *functor* categories, the latter being functions composed out of base categories (using backward and forward slashes indicating the direction of the function). Table 6.1 shows the CCG categories used in the GMB, together with their associated semantic types. On the sentence level, the lexical CCG categories are combined using a fixed set of

Table 6.1: Overview of CCG categories and semantic types

Category	**Description**	**Semantic type**
S	sentence	$\langle\langle e,t\rangle,t\rangle$
NP	noun phrase	$\langle\langle e,t\rangle,t\rangle$
N	noun	$\langle e,t\rangle$
PP	prepositional phrase	$\langle e,t\rangle$
$b\!:\!\beta/a\!:\!\alpha$	functor (with argument a right)	$\langle \alpha,\beta\rangle$
$b\!:\!\beta\backslash a\!:\!\alpha$	functor (with argument a left)	$\langle \alpha,\beta\rangle$

combinatory rules: forward application ($>$), backward application ($<$), composition ($\vee$) and type raising ($*$) (see Figure 6.1 below for an example derivation from the GMB).

Named Entities. In addition to the syntactic categories, all tokens are annotated with several aspects of lexical meaning. Firstly, Named Entity (NE) tagging categorizes tokens that refer to some named entity into different classes of entities, i.e., distinguishing persons, locations, organizations, geo-political entities (GPE's), etc. As a tagging scheme for Named Entities, we use a simplified version of Sekine et al.'s (2002) Extended Named Entity classification (for the complete overview of NE tags, see Bos et al., 2015).

One of the main issues in NE tagging is the observation that named entities often exhibit a complex hierarchical structure. For instance, the expression "*New York Times*" as a whole indicates an organization (or artifact, depending on the context), but part of the expression, i.e., "*New York*", indicates a location (this, in turn, could even be separated into two separate named entities, i.e., "*New*" and "*York*", which arguably make different contributions to the identification of the location). In the GMB, embedded named entities are currently not explicitly tagged; all tokens obtain the NE tag associated with the outer NE ('organisation', in the example above). Critically, however, this analysis assumes that complex named entities introduce a single entity into the discourse model, which means that any embedded entity is not available for anaphoric reference. Therefore, the analysis cannot account for the anaphoric interpretation of "the city" in the following example from the GMB:

> The New York City Fire Department has released thousands of pages of oral histories and hours of radio transmissions from the September 11, 2001, terrorist attacks on the World Trade Center. [...] The New York Times newspaper and families of the victims of the attacks had sued for the release of the material. *The city*... (`doc. 20/0099`)

In this example, the definite description "the city" refers to the city of New York. However, all mentions of "New York" in this text are part of embedded named entities; both "the New York City Fire Department" and "the New York Times" introduce organisations. Under the 'flat' analysis of embedded named entities, "New York" is therefore not an available antecedent for "the city" in this text.

A possible way to resolve the issue of embedded named entities is based on the 'Structured Named Entities' approach suggested as part of the Quaero program (Grouin et al., 2011; Rosset et al., 2012): all NEs are separated into referential expressions (called *types*) and functional expressions (called *components*), such that the components act as modifiers of the types. For example, the title "*President*" in the compound expression "*President Obama*" serves as a component modifying the person-type "*Obama*", via a relation that can be dubbed "*title*". If we extend this idea and assume that types themselves can serve as components, the structure becomes compositional and we can treat "*New York Times*" as a structured NE: an expression of type 'organization', consisting of two sub-types, i.e., the location-type "*New York*" and the organisation-type "*Times*", such that the first modifies the latter via a "*from*"-relation. Interestingly, such an analysis highlights the correspondence between embedded named entities and implicit relations, which will be discussed in more detail in section 6.2.2 below.

Animacy and Word Senses. The two other layers of lexical meaning present in the GMB are animacy-tagging and word sense-labelling. Animacy is a property of nouns, which denotes whether its referent is animate or not, or whether it refers to a human-like entity (e.g., a robot). This property has been shown to play a role in various linguistic phenomena, including argument realization (Dell'Orletta et al., 2005), dative alternation (Bresnan et al., 2007), and co-reference resolution (Orasan and Evans, 2007; Lee et al., 2013). The classification used in the GMB is the one proposed by Zaenen et al. (2004). Finally, all lexical items that are POS tagged as noun, verb, adjective or adverb are associated with a word sense from WordNet (version 3.1; Fellbaum, 1998), expressed as a synset identifier.

6.2.2 Sentences

The annotation layers described above can all be seen as operating on the word-level; despite the obvious influence of the direct linguistic context on these layers, the properties can all be assigned to a single token. Now, if we move to the sentence-level, we find various aspects of meaning that are constituted by relations *between* tokens.

Thematic roles. The primary example of a sentence-level annotation is the annotation of thematic roles associated with verbs, which indicate the relation between the verb and its arguments. The thematic roles used in the GMB are adopted from VerbNet (Kipper et al., 2008). Interestingly, the CCG category assigned to the verb helps in determining which thematic roles may be associated with it (Bos et al., 2012); for example, CCG category $(S\backslash NP)/NP$ signals a transitive verb requiring two arguments, e.g., [Agent, Patient]. This means that we can associate each token with an ordered (possibly empty) set of roles, whose size is determined by the CCG category of the token, i.e., the number of arguments and adjuncts. So, despite the fact that the thematic role assignment is considered a sentence-level phenomenon (as it describes the relation between multiple tokens), it is annotated on the token level in the GMB. This keeps the annotations simple, since the relation between the verb and its argument remains implicit (it is derived from the syntax), and flexible, since the arguments of the verb and the thematic roles associated with them, are determined on separate layers.

Implicit relations. The implicit semantic relations annotated in the GMB are signaled, for example, by noun-noun compounds (e.g., a "*fishing zone*" describes a zone for fishing), possessive constructions (e.g., "*his wife*" refers to the wife of him) and temporal modifiers ("*the meeting Wednesday*" indicates the meeting on Wednesday). The implicit relations are based on English prepositions, and derived from the Google Ngram Viewer (Michel et al., 2011). Just like the thematic roles, these relations can be represented on the token-level, as properties of the tokens that trigger the implicit relations.

Scope interactions. Another type of semantic relation annotated in the GMB is the way in which the scopal properties of semantic operators are affected by specific linguistic constructions. In most cases, the arguments that are affected by a scope-bearing operator are determined by the order in which the arguments occur in the linguistic surface structure of the sentence. For instance, in the sentence "*A man loves many women*", the quantifier "*many*" only takes scope over the argument "*women*". Critically, however, this default ordering may in some cases be affected by the use of specific linguistic constructions, for example modifying constructions triggered by prepositions, such as "*A teacher of every student came to the party*". In this example, the quantifier "*every*" may be interpreted both as taking wide scope (corresponding to the reading that every student has a teacher that came to the party) and narrow scope (corresponding to the reading that there is some teacher that teaches all students, who came to the party). In the GMB, scope-ordering is annotated on the token level as a property of prepositions, which may either be scope-preserving, in which case they are tagged as 'Default', or scope-inverting, in which case they are tagged as 'Inverting' (see Evang and Bos, 2013).

Co-reference. When two or more separate expressions denote the same entity, they are considered to be *co-referential*. The property of co-reference is arguably a discourse-level, rather than a sentence-level property, since co-reference relations often reach beyond the sentence-boundaries. Just like the semantic relations described above, however, co-reference is annotated at the token level in the GMB: co-referential expressions are annotated with their direct antecedent, which may have been introduced in the same, or any preceding sentence. The expressions that are currently being annotated for co-reference in the GMB are pronouns, definite noun phrases, and proper names. Pronouns are the most paradigmatic cases of co-referential expressions; they predominantly occur in contexts in which they co-refer with some antecedent expression. Definite noun phrases and proper names, on the other hand, often occur without any antecedent, in which case they are annotated with the co-reference tag 'new'. Depending on the referential expression that is used, the set of expressions that can serve as possible antecedents differs: pronouns and proper names generally refer to entities (introduced by nouns), definite descriptions may also refer to more abstract entities introduced, for example, by verbs (e.g., "*[protestors] demonstrated [...] <u>the</u> demonstration*"). Two important issues that need to be taken into account when annotating co-reference are the availability of multiple or complex antecedents, and the occurrence of bridging inferences; below we describe these issues in some more detail.

In some cases, an expression has multiple correct antecedents, which all denote the same entity (e.g., "*The president [...] Obama [...] <u>he</u>*"); we call this a 'co-reference chain'. Similarly, the antecedent of a referential expression may consist of multiple tokens (e.g., "*President Barack Obama [...] <u>he</u>*"). The annotation convention used in the GMB for these cases is to annotate the closest antecedent-token of the co-reference chain as the correct antecedent. This decision does not affect the final representations; since co-reference is considered a transitive property in the semantic analysis, multi-word expressions and co-reference chains are interpreted as introducing a single entity into the discourse model (which is represented as a Discourse Representation Structure; see Section 6.2.3 below).

While co-reference indicates that the referent of a referring expression is identical to some previously introduced entity, 'bridging' indicates a different relation between a referring expression and its antecedent. For instance, in a text about "*an ox*", a subsequent mention of "*the owner*"is interpreted as referring to the owner *of* the ox. Bridged inferences are currently not explicitly annotated in the GMB, but treated as regular cases of co-reference. In the next chapter, we will investi-

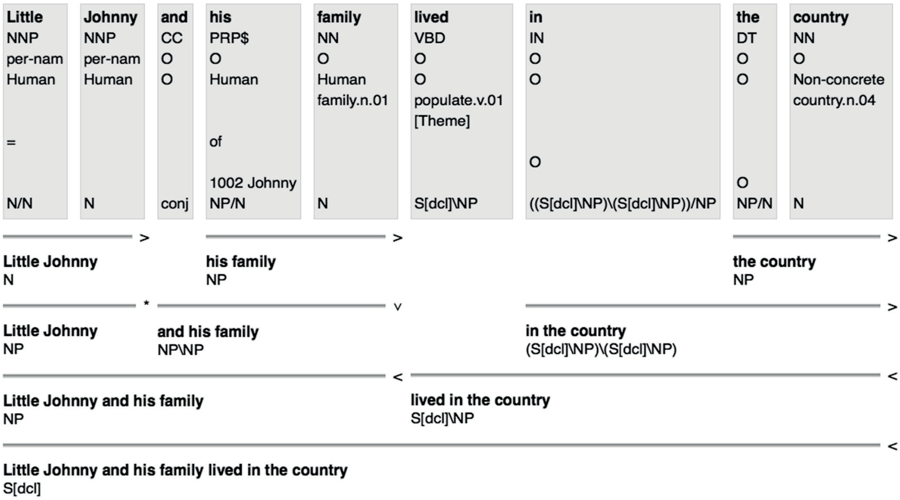


Figure 6.1: CCG derivation from the GMB (document 89/0732) with token- and sentence-level tags; from top to bottom: POS, NE tags, animacy, word senses, thematic roles, implicit relations, scope, co-reference, and CCG category.

gate whether the distinction between co-reference and bridging can be derived automatically, either based on features from the expressions involved (e.g., number and animacy classification), or based on external resources, such as WordNet.

Tokens and sentences: the complete picture. Since all sentence-level phenomena described in this section are annotated on the token-level, we can integrate these different levels of annotation into a single representation. Figure 6.1 shows an example sentence from the GMB, annotated with its token- and sentence-level annotations, as well as the full CCG derivation.

6.2.3 Discourse

The main aim of the GMB is integrating a variety of linguistic phenomena into a single semantic formalism. The levels of annotation described thus far serve as preliminaries to an integrated representation of the meaning of a complete discourse, based on Discourse Representation Theory. Here, we describe the motivation for the formalism, and the representation of discourse-level phenomena, such as rhetorical structure, and events.

Semantic formalism. The semantic representations in the GMB are based on Discourse Representation Theory (DRT; Kamp, 1981; Kamp and Reyle, 1993). The main advantages of using DRT in the context of developing the GMB, are its high coverage, the existence of tools for automatic analysis, and the readability of the resulting representations. We will briefly describe each of these properties below.

With respect to the coverage of linguistic phenomena, DRT has been shown to provide a powerful formalism that can account for various phenomena, including anaphora, quantification, tense, attitude reports and ellipsis (see Kamp et al., 2011, for an overview). Moreover, the version of DRT applied in the GMB is based on the extension Projective Discourse Representation Theory (PDRT), which provides a parsimonious analysis of projection phenomena, such as presuppositions and conventional implicatures (Venhuizen et al., 2013b, 2015). In PDRT, projection is represented by means of variable binding; all content is associated with a projection variable, which can be bound in the local context (for asserted content), or some higher, accessible context (for projected content). This way, projected content does not need to be moved from its interpretation site, preserving the alignment between the semantic representations and the linguistic surface structure.

From a practical point of view, DRT is interesting because it is one of the few semantic formalisms for which high-coverage automatic semantic analyzers exist, most notably *Boxer* (Bos, 2008), the application used in the GMB. Boxer produces representations based on CCG derivations, and incorporates several relevant extensions of DRT, including Projective DRT and Segmented DRT (see below). Finally, the property of providing readable representations is especially interesting for our current purpose, as the GMB consists of automatically derived semantic analyses that need to be verified by human annotators.

Table 6.2: Overview of rhetorical relations used in the GMB

	Relation	**Description**
Coordinated	*Continuation*	common topic
	Narration	common topic, occur in sequence
	Precondition	inverted *Narration*
	Result	cause
Subordinated	*Background*	extra information (general)
	Elaboration	extra information (specific elem.)
	Commentary	opinion
	Explanation	inverted *Result*

Rhetorical structure. For representing the rhetorical structure of a discourse, the GMB uses representations based on Segmented Discourse Representation Theory (SDRT; Asher, 1993; Asher and Lascarides, 2003). The formal semantics of DRT is combined with a detailed semantics for rhetorical relations, producing a theory that incorporates both semantic and pragmatic features in order to account for dependencies between the discourse structure and semantic interpretation.

The first step of annotating rhetorical structure is separating texts into units corresponding to single propositions; these are called discourse segments. These segments are accordingly connected via rhetorical relations. In SDRT, a basic distinction is drawn between two types of discourse relations: coordinating discourse relations and subordinating discourse relations, reflecting different properties of textual coherence, for example the temporal order of the described events, or the communicative intentions of the speaker. Table 6.2 summarizes the discourse relations currently used in the GMB.

Event Structure. The representation of events in the GMB is based on a neo-Davidsonian event semantics. This means that events are first-order entities, characterized by one-place predicate symbols. Events are associated to their arguments and modifiers via an inventory of thematic roles, derived from VerbNet (Kipper et al., 2008). This analysis of events results in consistent representations, with predicates

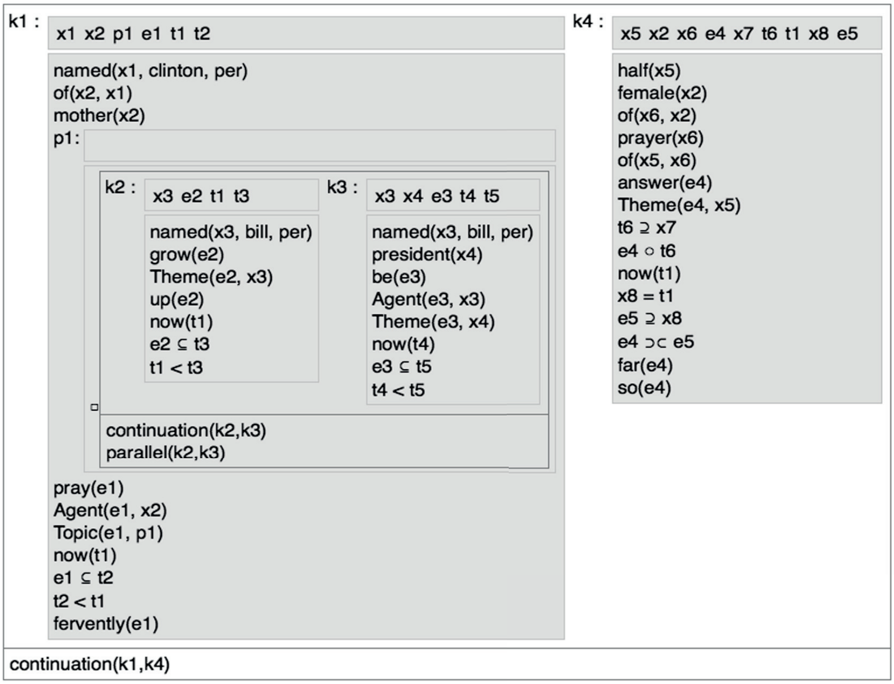


Figure 6.2: DRS representation from the GMB for the discourse: “Clinton’s mother prayed fervently that Bill would grow up and be president. So far, half of her prayer has been answered.” [GMB document 19/0727].

of a fixed arity, which can be derived in a compositional manner.

Discourse: the complete picture. Figure 6.2 shows a DRS from the Groningen Meaning Bank, including the neo-Davidsonian event structure, rhetorical relations, and the local representation of projected content following PDRT. Note that in the current version of the GMB, the projection variables from PDRT are only represented internally.

6.3 Meaning banking

The process of collecting and automatically analyzing linguistic data, correcting the analyses, and improving the automatic tools, is what we

Table 6.3: Overview of sub-corpora in the GMB. The statistics are as of November 1, 2014 (excluding RTE and Tatoeba sub-corpora, which are only for development purposes).

Subcorpus	**Documents**	**Sentences**	**Tokens**
Aesop's fables (fables)	224	950	23,105
Basic jokes (jokes)	122	443	7,533
CIA World Factbook (CIA)	514	4,430	112,535
Manually annotated sub-corpus Open ANC (MASC)	35	291	6,991
Voice of America (VOA)	9,207	57,159	1,238,659
Total	10,102	63,269	1,388,823

call *meaning banking*. In this section, we describe each of these steps in the process.

6.3.1 Data collection and automatic analysis

The GMB aims to provide high-quality annotations that can be used for theoretical purposes, as well as for improving statistical machine learning methods. This implies that the data for analysis should be diverse, representative, and freely available for distribution. On the other hand, state-of-the-art natural language processing tools are relatively limited in terms of their coverage of different genres, languages and modalities (e.g., written vs. spoken language). Therefore, the GMB currently focuses on English written texts, which are available from the public domain. It covers various genres, divided over different sub-corpora, including newswire and informative texts (represented by the *VOA* and *CIA* sub-corpora), fiction (the *fables* sub-corpus), and a miscellaneous genre containing, e.g., jokes and emails (from the *jokes* and *MASC* sub-corpora). Table 6.3 provides an overview of the different sub-corpora and their distribution over the corpus.

The tool-chain of NLP tools applied in the GMB for automatically deriving the semantic analyses consists of the following elements: tokenization using the statistical tokenizer Elephant (Evang et al., 2013), POS and Named Entity tagging using the taggers included with the

C&C tools (Curran et al., 2007), morphological analysis using Morpha (Minnen et al., 2001), animacy tagging using a dedicated classifier (Bjerva, 2014), syntactic analysis using the C&C parser (Curran et al., 2007), and finally semantic analysis using Boxer (Bos, 2008).

6.3.2 Collecting linguistic annotations

The annotations made by the NLP tool-chain are necessarily imperfect due to the way in which these tools are developed. Machine learning methods always need to make a trade-off between quality and quantity, since obtaining high-quality annotations for large amounts of data is expensive. Wide-coverage methods therefore often need to rely on statistical regularities, which means that the analysis is easily subject to overgeneralization. In order to improve the annotations, as well as the tools used for obtaining them, we collect data annotated by human annotators via two main sources: the GMB Explorer interface and Wordrobe, which will be briefly described below. In addition to these two sources of input, the GMB annotations are improved using various existing NLP tools, as well as externally collected annotations (e.g., the annotations provided by the Open ANC project for their MASC subcorpus).

Explorer interface. The documents of the GMB, together with their annotations, are available via an online wiki-like platform, called the GMB Explorer[1] (Basile et al., 2012b). The platform allows for navigation and search through the documents of the GMB, visualization of various levels of annotation, and manual correction of the existing annotations. The latter feature is especially aimed at expert linguists, who are able to directly assess and improve the provided annotations. The Explorer moreover allows users to see and evaluate the result of their corrections, as documents can be reprocessed directly based on the NLP tool-chain and the provided annotations. In case a document contains an erroneous annotation that cannot be corrected, or if there is some other problem with the document, users can report an issue which will be directly send to the developing team. Currently, manual corrections

[1] `http://gmb.let.rug.nl/explorer/explore.php`

can only be made to the token- and sentence-level annotations; the discourse annotations are not yet available for manual editing. One reason for this is that the discourse-level annotations are highly influenced by lower-level annotations. In order to prevent conflicting information at the token/sentence- and discourse-level, our first aim is to create a high standard for the lower-level annotations, which can then be used to provide a good reference point for the discourse-level annotations.

Wordrobe. In order to obtain annotations from non-experts in linguistic annotation, we use a crowd-sourcing technique called *Game with a Purpose*. To this end, we developed 'Wordrobe',[2] a collection of games each targeting a specific level of linguistic annotation, e.g., Named Entity tagging, co-reference resolution, and word sense disambiguation; see Table 6.4 for an overview. In all games, a question consists of a piece of text from the GMB with one or more highlighted elements, and a set of annotation choices. The question and choices both contain as few linguistic terminology as possible. Players are asked to select one choice, and indicate their confidence in the answer using a betting system. The correctness of the answer is evaluated based on a system of agreement: the more players agree with the same answer, the more points can be gained by selecting that answer. This means that player's points are subject to constant change, since the agreement score for any question may change as more players answer the same question. In addition to the dynamic points-system, players are awarded with static *achievements*, which are obtained by finishing sets of questions (called *drawers*).

In order to evaluate the quality of the annotations obtained from Wordrobe, we have conducted an study that focuses on one of the games, namely the word sense disambiguation game called 'Senses' (Venhuizen et al., 2013a). Using different aggregation methods, the quality of the annotations was measured using precision and recall for questions that obtained answers from six different players, relative to a gold standard annotation. The highest F1-score was 0.86 (representing the harmonic mean between the precision—0.88—and recall—0.83), which was ob-

[2] `http://wordrobe.housing.rug.nl`

Table 6.4: Current set of Wordrobe games

Game	Task	Possible choices
Animals	Animacy classification	*human*, *animal*, ..., *abstract*
Bridges	Information structure	*explicit*, *implicit*, or *new*
Burgers	Compound noun disamb.	prepositions (Google Ngrams)
Names	Named entity tagging	*person*, *organisation*, ..., *time*
Pointers	Anaphora resolution	possible antecedents in text
Roles	Thematic role labelling	VerbNet relations
Senses	Word sense disamb.	WordNet 3.1 synsets
Twins	Homonym disamb.	*noun* or *verb*

tained using the least conservative method of aggregation, i.e., relative majority. Precision got up to 0.97 using the most conservative method of aggregation, i.e., unanimity, albeit with a considerable loss of recall (to 0.35). Overall, the results showed that the annotations obtained using Wordrobe are of a relatively high quality, despite the difficulty of the task at hand, which is well-known for word sense disambiguation, as demonstrated by the various SensEval/SemEval tasks organized over the last 16 years (Kilgarriff and Rosenzweig, 2000). In order to perform a more general evaluation of the annotations obtained from Wordrobe, we are currently conducting a large-scale experiment in which the annotations are compared to other crowdsourced annotations obtained via the CrowdFlower platform[3].

6.3.3 Bootstrapping the analyses

The annotations gathered from the different sources need to be incorporated into the GMB in order to improve the annotations, as well as the tools used for deriving the annotations. Besides providing annotations for different levels of linguistic information, the corrections obtained from the various sources may in some cases be contradictory. This means that corrections should be stored and evaluated as separate annotation decisions. Moreover, the corrections should not be dependent upon a specific annotation format, since the tools used for automatic

[3] `http://www.crowdflower.com`

analysis may change. Therefore, the GMB employs a bootstrapping approach that treats annotations as stand-alone facts, which are called *Bits of Wisdom*, or BOWs for short. Each type of linguistic annotation is associated with its own type of BOW, containing an identifier to the annotated expression (usually in the form of character offsets), information about the annotation layer (e.g., 'POS'), and the annotation itself (e.g., 'NNP').

The various sources of BOWs differ with respect to the quality of their contributions; for example, the annotations obtained from the expert linguists in the Explorer interface may be considered more reliable than the crowd-sourced annotations from Wordrobe. As a means of adjudicating between these different BOWs, we currently employ a combination between weighted evaluation and recency; the BOWs based on less consistent sources like Wordrobe are derived using an aggregation method (currently, relative majority, but see Venhuizen et al., 2013a, for alternative proposals), and in the case of conflicting BOWs, the most recent BOW is applied to the annotation. However, all BOWs are permanently stored, so that they can be reevaluated and reapplied at each processing cycle. Our aim is to incorporate a more elaborate 'judge component' in the future, which can make informed decisions about the choice of BOW using statistical methods.

The BOWs are applied at various stages of the toolchain in order to improve the automatically derived annotations in the GMB. In the future, we also aim to employ the BOWs to improve the NLP tools themselves, which will aid future efforts in automatic annotation. Figure 6.3 shows an overview of the toolchain employed in the GMB, together with the bootstrapping mechanism that incorporates BOWs from different sources into the toolchain.

6.4 Evaluation and future perspectives

Previous work on semantically annotated corpora has mainly aimed at analyzing specific semantic phenomena; e.g. predicate-argument relations (PropBank; Palmer et al., 2005), semantic frames (FrameNet; Baker et al., 1998), or discourse relations (the Penn Discourse TreeBank; Prasad et al., 2008). One notable exception is OntoNotes (Hovy

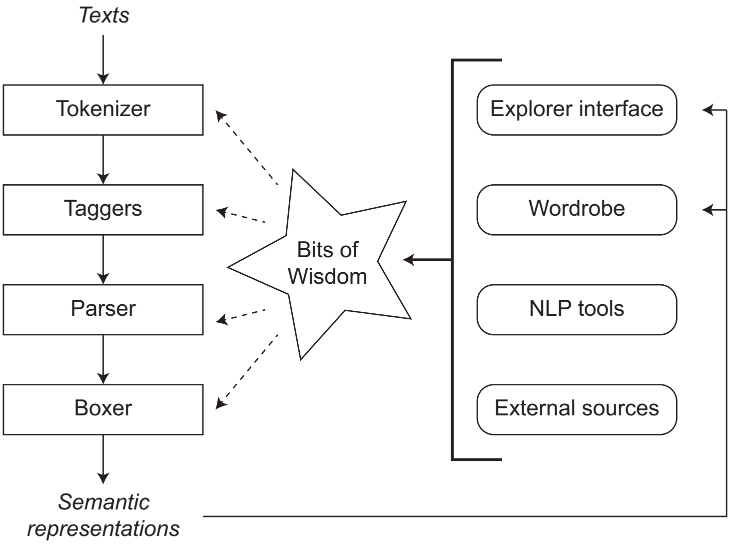


Figure 6.3: Graphical representation of the GMB toolchain and bootstrapping mechanism

et al., 2006), which combines various levels of annotation, including syntactic analysis based on the Penn TreeBank, predicate-argument structure based on PropBank, word senses, and co-reference. The advantage of the GMB over these resources is that the semantic representations in the GMB combine all levels of annotation into a single analysis. Moreover, the bootstrapping method of combining automated and human annotations results in a large-scale resource with high quality annotations.

The amount of Bits of Wisdom available for different levels of analysis in the GMB gives an indication of the quality of the obtained annotations. Table 6.5 shows an overview of the number of BOWs per source. Moreover, it shows the number and percentage of *effective BOWs*, which represent the BOWs that actively contribute to the annotation; in the case of multiple or conflicting BOWs, only the BOW selected by the judge component is considered ‘effective’. The high total number of BOWs obtained from the different sources indicates that the automatic analyses in the GMB are subject to a considerable amount of verification. We can

Table 6.5: BOW statistics per source, as per November 7, 2014.

Source	#BOWs	#Effective	%Effective
Explorer: Manual	44,000	39,279	89%
Explorer: Script	134,335	104,744	78%
Wordrobe	7,018	4,639	66%
External (MASC)	13,351	9,626	72%
Total	198,704	158,288	79%

look at the high percentage of effective BOWs in two ways: on the one hand, it indicates that the data-collection effectively targets the weak spots of the natural language processing tools used for obtaining the automatic annotations. On the other hand, it highlights the fact that these tools still need to be improved significantly in order to approach a gold standard annotation. We aim to achieve this by re-training the tools on the manually corrected data, which in time should lead to an overall smaller percentage of effective BOWs (since the tools would approximate the gold standard annotation).

As the aim of the GMB is to both provide a resource for semantic analysis, and to improve existing semantic annotation tools, directions for future work are also two-fold; firstly, we aim to improve the tools used for obtaining the semantic analyses in the GMB, and thereby the analyses themselves, using the BOWs obtained via the different sources. In particular, the current method of choosing between competing BOWs, which is based on recency, is rather uninformed. In the future we aim to evaluate the quality of the different sources of BOWs, and apply them accordingly. The second main direction for future work is applying the GMB data for data-driven semantic analysis. In the next chapter, we describe such an effort, which uses co-reference annotation from the GMB to predict the information status of different phenomena.

Finally, the GMB is subject to constant extension and improvement. As described above, we aim to incorporate several additional levels of semantic analysis, such as the annotation of bridging relations. Moreover, the GMB is currently being extended to include multi-lingual texts and translations between them. This project, called the Parallel Meaning

Bank, opens up the way to exploring semantic alignment between different languages, and resolving ambiguities in one language based on annotations obtained from another language (Bos, 2014). These different applications show that meaning banking opens up new and promising directions in computational linguistic research.

Chapter 7

The information status of referential expressions: A corpus study

Abstract. When investigating the information status of projection phenomena in discourse, the class of referential expressions is especially interesting because of its heterogeneity; despite the common property of referring to an entity, referential expressions seem to differ considerably with respect to the information status associated with this entity, i.e., whether it is *given*, *new* or *inferred* in the discourse context. A prerequisite for a formal semantic treatment of these expressions, therefore, is to determine to what extent referential expressions differ in terms of the information status of their referent, and how this choice is affected by different types of linguistic features, including the expression itself, its context, and the way in which it is represented as part of the semantic representation. To investigate this, we used a crowdsourcing platform to collect manual annotations of the information status of referential expressions, and fitted a multinomial logistic regression model with a set of token-based, contextual and deep semantic features. The model shows that the type of referential expression (e.g., name versus definite noun phrases) is indeed one of the strongest predictors for information status, along with a superficial contextual feature describing whether the lemma was already introduced in

the discourse context. The results moreover indicate that information status should be interpreted on a gradient scale, where *given* and *new* represent the two extremes, and *inferred* constitutes the less well-defined middle area.

7.1 Introduction

Formal semantic analyses often take a categorical stance with regard to the semantic behavior of different linguistic expressions. This means that differences in interpretation generally result in the identification of separate meaning categories. Proper names and definite descriptions, for instance, have been argued to differ in terms of their *information status*; that is, whereas proper names are generally used to refer to fixed referents in the discourse, definite descriptions may get their reference by virtue of the context in which they occur. As such, traditional semantic approaches have treated proper names and definite descriptions as distinct semantic phenomena; the former are interpreted using semantic constants while the latter are interpreted using discourse variables. Critically, the separation of these phenomena into different meaning classes on the basis of their information status forgoes a unified semantic analysis; i.e., an analysis that does not only highlight the differences between semantic phenomena, but also accounts for their similarities and interactions.

A semantic formalism that does not presume an inherent distinction between expressions on the basis of information status is Projective Discourse Representation Theory (PDRT; Venhuizen et al., 2013b, 2015, see Part II of this thesis). PDRT is an extension of Discourse Representation Theory (Kamp, 1981; Kamp and Reyle, 1993) that explicates the notion of information status, i.e., how linguistic content is related to the discourse context, without treating it as a segregating aspect of linguistic meaning. More specifically, the PDRT analysis introduces variables to represent differences in information status, which can be resolved in different ways depending on the discourse context. This means that there is no *a priori* distinction in the way different referential expressions are analyzed within PDRT, since their information status is determined by the context that they interact with.

In this chapter, we investigate differences in information status within the class of referential expressions, including proper names, definite descriptions, indexical expressions, possessive constructions, and pronouns. We aim to determine whether referential expressions differ in terms of the information status of their referent and what linguistic features affect this property. We will investigate features from three different levels of linguistic analysis: the *token level*, containing properties of the referential expression itself, the *context level*, describing properties of the sentential as well as the discourse context in which the expression occurs, and the *deep semantic level*, reflecting how the expression is embedded as part of the formal semantic representation of the context in which it occurs. We use the Groningen Meaning Bank (GMB; Bos et al., 2015, see Chapter 6) corpus to collect instances of the different types of referential expressions, which are then manually annotated for information status using an online crowdsourcing platform. Traditionally, the information status of referential expressions is described using three main classes: *given*, *new*, and *inferred* (Prince, 1981). The first class describes expressions that refer to entities that were already explicitly introduced in the discourse, and the second class describes expressions that refer to entities that were not introduced before. The third class consists of expressions that refer to an entity that is related to a known discourse referent, but not identical to it. As such, this class has also been described as 'bridging inferences' (Clark, 1975), 'inferrables' (Prince, 1981), and 'indirect anaphora' (Chafe, 1976). An example is given in (1) (GMB document 14/0694), where the entity referred to by "the owner" is taken to be *inferred*, as it refers to the owner *of the "Ox"* that was introduced in the previous sentence.

(1) A Heifer saw an Ox hard at work harnessed to a plow, and tormented him with reflections on his unhappy fate in being compelled to labor. Shortly afterwards, at the harvest festival, <u>the owner</u> released the Ox from his yoke, but bound the Heifer with cords and led him away to the altar to be slain in honor of the occasion.

The information status of referential expressions has been studied from a psychological (e.g., Clark and Haviland, 1977; Burkhardt, 2006), formal (e.g., Bos et al., 1995; Asher and Lascarides, 1998; Piwek and Krahmer, 2000; Irmer, 2009), as well as an empirical perspective (see, e.g., Fraurud, 1990; Poesio and Vieira, 1998; Gardent et al., 2003). These investigations focus on how the different classes of information status are processed, represented, and categorized, respectively. Another line of work aims to automatically identify the information status of referential expressions (e.g., Postolache et al., 2005; Nissim et al., 2004; Nissim, 2006). None of these studies, however, focuses on determining which specific features from different levels of linguistic analysis affect the information status of different types of referential expressions. The goal of the current study, therefore, is to investigate which linguistic features can be used to predict whether a referential expression is *given*, *new*, or *inferred* with respect to its discourse context. We will statistically assess the contribution of the different features using multinomial logistic regression modeling (cf. Bresnan et al., 2007).

7.2 Data collection

In order to investigate which contextual features influence the information status of referential expressions (based on the three classes *given*, *new*, and *inferred*), we collect instances of the different expressions from a large corpus of semantically annotated texts (the Groningen Meaning Bank), and derive a range of different linguistic features for each of these expressions. Accordingly, we collect annotations on the information status of these expressions via the crowdsourcing platform CrowdFlower. The resulting data will form the basis for the multinomial logistic regression model presented in Section 7.3.

7.2.1 Collecting data from the Groningen Meaning Bank

The Groningen Meaning Bank (GMB; Bos et al., 2015) is a large corpus of texts, with different levels of linguistic annotation, including token-based annotations (e.g., syntactic categories, named entity tagging, animacy classification, and word senses), sentence-level annotations (e.g., thematic roles, syntactic scope), and discourse-level annotations (e.g., co-reference annotation and rhetorical structure); see Chapter 6 for a more detailed overview. Critically, the annotations in the GMB also incorporate deep semantic representations, based on Projective Discourse Representation Theory (Venhuizen et al., 2013b, 2015). This level of deep semantic representation is especially useful for the current purpose of deriving features regarding the information status of referential expressions, as it allows for a formal definition of information structural properties, such as the number of accessible discourse referents.

We identified a set of referential expressions from the GMB, focusing on proper names, definite descriptions, and possessive constructions. We did not incorporate pronouns in this study, because a preparatory analysis revealed that pronouns in the GMB were predominantly used in contexts in which their referent is *given* (which can be ascribed to the fact that the GMB consists largely of newspaper articles; see Table 6.3 in the previous chapter). Including pronouns would therefore highly bias our data, since the information status of these expressions can be determined solely based on the fact that they are pronouns, disregarding any other features. We only selected referential expressions that occurred in the beginning of texts—within the first 70 tokens—in order to keep the texts reasonably short for the annotation procedure (see Section 7.2.2). Moreover, text-inital expressions were omitted, as they are *new* by necessity. In total, we collected 991 referential expressions, for which we derived a collection of features from the different levels of analysis represented in the GMB.

For the task of classifying referential expressions according to their information status, we can select features similar to those used for co-reference resolution, with the exception that the current task is based on a single referential expression along with its context, instead of comparing two referential expressions for co-reference. This means that we cannot use features that depend on the properties of a specific potential antecedent, such as its grammatical role (which has been shown to be an important feature for co-reference resolution; see, e.g., Mitkov, 2002). Instead, we collect *token-based features* that describe properties of the referential expression itself, and a set of *contextual features* that describe properties of the entire preceding context (e.g., the availability of potential antecedents). In addition, we also collect a set of *deep semantic features* that describe the properties of the referential expression as part of the semantic representational structures from Projective Discourse Representation Theory. Below, we will briefly describe the features selected in each of these classes; an overview of all features is provided in Table 7.1.

(i) Token-based features. The first set of features represents token-based features that describe properties of the referential expression itself. *Sentence* indicates the number of the sentence in which the referential expression occurs, counted from the beginning of the text, and *Position* indicates the position of the referential expression in the sentence, counted from the beginning of the sentence to the head noun of the referential expression. The part-of-speech tag and animacy classification associated with the head noun are reflected by *POS* and *Animacy*, respectively (for more information about the token-level annotations in the GMB, see section 6.2 in the previous chapter). Finally, this set includes two properties of the referential expression as a whole: *NPlength* indicates the number of tokens in the NP (counting from the first token up to the head noun), and *NPtype* indicates to which of the four referential classes the expression belongs: definite noun phrases (e.g. "*the president*"), indexical expressions (e.g., "*this week*"), names (e.g., "*Mr. Zhvania*"), or possessive constructions (e.g., "*[Britain 's] participation*").

(ii) Contextual features. The second set of features represents properties of the sentential and discourse context in which the referential expression occurs. *Role* represents the semantic role of the referent (again, see section 6.2). *Arg1* describes whether the introduction of the referent is modified by some other referent or event. If this is the case, the semantic analysis in the GMB indicates that there is binary semantic relation between the referent and the modifying entity, and the referent is the *first* argument of this relation. Analogously, *Arg2* determines whether the referent occurs as the *second* argument of binary predicate in semantic representation, i.e., reflecting wether the referent itself modifies another entity. These features are illustrated in (2):

Table 7.1: Overview of features collected from the GMB. The most frequently selected (i.e., median) values are underlined, and m indicates the mean value for integer features.

Feature	Description	Values
Token-based features		
Sentence	Number of sentence	1, 2, 3, … ($m = 1.8$)
Position	Position in sentence	1, 2, 3, … ($m = 14.2$)
POS	POS-tag of head token	*NN(S)*, *NNP(S)*, …
Animacy	Animacy of referent	*Inanimate* or *Animate*
NPlength	Number of tokens in NP	1, 2, 3, … ($m = 2.0$)
NPtype	Type of referential expression	*Definite* or *Indexical* or *Name* or *Possessive*
Contextual features		
Role	Thematic role of referent	*None*, *Agent*, *Patient*, …
Arg1	Occurs as 1st argument of a binary predicate	*False* or *True*
Arg2	Occurs as 2nd argument of a binary predicate	*False* or *True*
LemmaOcc	Same lemma was already used in text	*False* or *True*
LemmaFreq	Number of occurrences of same lemma	0, 1, 2, … ($m = 0.3$)
LemmaRec	Token distance to previous occurrence of lemma	0, 1, 2, … ($m = 5.3$)
SuitAnt	Number of compatible antecedents	0, 1, 2, … ($m = 1.6$)
Deep semantic features		
PDRSdepth	Level of DRS embedding	1, 2, 3, … ($m = 2.9$)
PSites	Number of accessible projection sites in PDRS	1, 2, 3, … ($m = 2.8$)
Ants	Number of accessible referents in PDRS	0, 1, 2, …, 20, 21, … ($m = 20.4$)
xAnts	Number of accessible entities in PDRS	0, 1, 2, …, 7, 8, 9, … ($m = 10.0$)
Projected	Pointer of referent is not bound in local PDRS	*False* or *True*
Presup	Pointer of referent is released (free)	*False* or *True*
EmbPresup	Pointer of referent is released and subordinated	*False* or *True*

(2) United Nations investigators have begun questioning top Syrian officials about <u>the assassination</u> of former Lebanese Prime Minister Rafik Hariri.

The target referential expression in this example ("*the assassination*") occurs as the first argument of the binary predicate "*of*", which means that it is modified by the introduction of "*former Lebanese Prime Minister Rafik Hariri*". Moreover, the expression also occurs as the second argument of a binary predicate, namely "*about*", where it modifies the event introduced by "*questioning*". Hence, the referential expression "*the assisination*" has both *Arg1* and *Arg2* set to *True*.

Another aspect of the context that may affect the information status of referential expressions, is the availability of other referential expressions in the context that in some way match with the target expression. One way to specify such matching antecedents is on the basis of the lemma of the target expression; *LemmaOcc* describes whether or not the lemma of the head noun occurs in the foregoing text, and *LemmaFreq* and *LemmaRec* define the frequency of the lemma, and the distance to the closest mention, respectively. Another way of finding matching antecedents is to determine whether there are other referential expressions in the context that describe the same type of entity. This is captured by the *SuitAnt* feature, which determines the number of referential expressions in the context that match with the target expression in number (i.e., singular versus plural) and animacy.

(iii) Deep semantic features. The final set of features is extracted from the formal semantic representations underlying the analyses in the GMB; the semantic structures from PDRT (called PDRSs), which are derived using the semantic annotation system Boxer (Bos, 2003). We used PDRT-SANDBOX[1] (Venhuizen and Brouwer, 2014, see Appendix A) to derive features for predicting information status from Boxer's PDRS output. The PDRS features primarily describe the level of embedding of the target expression in the formal semantic representation, which affects the possible ways in which the information con-

[1]Available at: `http://hbrouwer.github.io/pdrt-sandbox/`

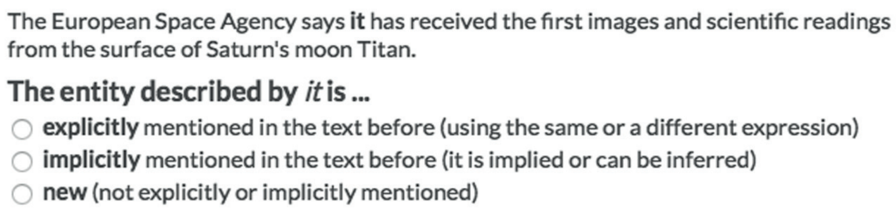
tributed by the referential expression can be resolved. Firstly, *PDRS-depth* represents the number of PDRS-embeddings from the introduction site of the referent, introduced by the target expression, to the outer PDRS. *PSites*, in turn, indicates the number of accessible PDRSs from the introduction site of the referent, taking into account constraints on accessibility (see Section 5.3.2 in Chapter 5). The number of accessible referents in the PDRS is represented by *Ants*, and the number of accessible referents that refer to entities (as opposed to events, or points in time) is indicated by *xAnts* (since Boxer's convention is to use variables starting with x as referring to entities).

In order to arrive at its semantic analyses, Boxer employs a rule-based mechanism to determine the presuppositional status of referential expressions (i.e., whether they are *bound* or *accommodated*; Bos, 2003, pp. 200–204) . As such, the presuppositional status of an expression as determined by Boxer is a complex interaction between contextual features and formal semantic rules. This means that we can use this information for predicting information status, as it can be derived directly (i.e., in a rule-based manner) from other, more superficial features. The three features that were derived from Boxer's analysis of presuppositions are *Projected*, *Presup* and *EmbPresup*. In case the referent is associated with a projection pointer that is not bound by the local PDRS, the referent is analyzed as projected (*Projected* = *True*), and in case this variable occurs free, it is analyzed as presuppositional (*Presup* = *True*). Moreover, presuppositional expressions may be embedded within other presuppositional expressions, in case the projection site indicated by the pointer of the target expression is such that another projection context is accessible from it (*EmbPresup* = *True*).

7.2.2 Annotating information status using CrowdFlower

CrowdFlower[2] is an online crowdsourcing market-place that provides a platform for collecting data from contributors from all over the world. In a similar way to other crowdsourcing platforms, such as Amazon's Mechanical Turk, CrowdFlower collects data by distributing small tasks over a large group of contributors ("the crowd"). These tasks may have

[2] http://www.crowdflower.com


The European Space Agency says **it** has received the first images and scientific readings from the surface of Saturn's moon Titan.

The entity described by *it* is ...

- **explicitly** mentioned in the text before (using the same or a different expression)
- **implicitly** mentioned in the text before (it is implied or can be inferred)
- **new** (not explicitly or implicitly mentioned)

Figure 7.1: Example of a CrowdFlower question

various objectives; for instance, sentiment analysis, data categorization (e.g., of images) or content rating, but they all have in common that they are short tasks that can be performed by a large pool of contributors. In the current study, contributors were presented with shorts texts from the Groningen Meaning Bank. Each text contained a highlighted referential expression, and was accompanied by a multiple-choice question aimed at determining the information status of this expression. That is, contributors had to indicate whether the referent of the highlighted expression was *given* ("explicitly mentioned in the text before"), *inferred* ("implicitly mentioned in the text before"), or *new* in the given text. An example question is shown in Figure 7.1 (note that this is one of the control questions, with a pronoun as target expression; see below), and the instructions that were given to the contributors are provided in Appendix 7.A.

Contributors were paid 0.06\$ per set of 6 questions. Each question was answered by at least 7 contributors. We used the most conservative way of selecting contributors in order to obtain the highest quality answers (on a 3-point Likert scale, ranging from "Highest speed" to "Highest quality"). In general, the contributors were satisfied with job, as reflected by an overall $4/5$ satisfaction rating. For the 991 questions we obtained a total number of 7199 annotations from 133 contributors. CrowdFlower implements a quality control system based on *control questions* (called "test questions" by CrowdFlower): questions that have a pre-determined gold-standard answer, and that are indistinguishable from the target questions. We employed a set of 100 control questions, of which half had pronouns as target expressions, an-

Table 7.2: Summary of the CrowdFlower data. The 'Annotations' are the individual answers of all contributers, and 'Questions' represents the aggregated answer per question. 'Trust' reflects the average trust (or confidence score) for the questions.

Answer	**Annotations**		**Questions**		**Trust**
given	1947	(27.1%)	269	(29.9%)	0.79
inferred	771	(10.7%)	42	(4.2%)	0.50
new	4481	(62.2%)	653	(65.9%)	0.87
Total	7199		991		0.83

notated with the gold-standard answer *given*, and the other half had indefinite noun phrases as target expressions, which were annotated with the gold-standard answer *new*. The control questions were randomly inserted throughout the job, in such a way that each page of 6 questions contained at least one control question. In order to participate in the job, contributors had to pass a "Quiz Mode" consisting of control questions for which they needed to obtain a minimum accuracy of 70%, and they had to maintain this accuracy throughout the job. The average contributor accuracy on our job was 0.95%.

Annotations were aggregated such that we obtain a single data point per question. The aggregate result of a question is chosen based on the answer with the greatest 'confidence score' or 'trust', which indicates the level of agreement between contributors (weighted by their individual accuracy acc); the trust of a given answer a to question q is calculated as follows:

$$trust(a|q) = \frac{\sum_{c \in C_a} acc(c)}{\sum_{c \in C_q} acc(c)} \tag{7.1}$$

where C_a represents the set of contributors that responded to question q with answer a, and C_q is the total set of contributors that answered the question. The average confidence score over all questions was 0.83 (on a scale from 0 to 1). Table 7.2 presents a summary of the collected data. One thing that immediately stands out is the low number of *inferred* answers overall; we will address this observation in Section 7.3.2.

Table 7.3: Overview of CrowdFlower data per *NPtype*

Answer	**NPtype**							
	Definite NP		Indexical		Name		Possessive	
given	137	(36%)	2	(22%)	143	(30%)	14	(12%)
inferred	24	(6%)	1	(11%)	14	(3%)	3	(3%)
new	222	(58%)	6	(67%)	322	(67%)	103	(86%)
Total	383		9		479		120	

Table 7.3 presents a breakdown of the aggregated CrowdFlower data per type of referential expression (based on the feature *NPtype*). Firstly, this table shows the under-representation of indexical expressions in our data-set, constituting only 1% of all annotated referential expressions. Interestingly, the distribution of answers shows that each type of referential expression is associated with each of the three different answers, which means that all different types of referential expressions are acceptable with each of the three types of information status. Moreover, for each of the different types of referential expressions, the ordering of the different types of information status corresponds to the ordering observed in the overall data set (shown in Table 7.2): the highest percentage of occurrences is annotated as *new*, followed by *given* and finally *inferred*. Critically, however, the different types of expressions do show a difference in the distribution of the individual answers; whereas possessives show a very high bias toward *new* as opposed to *given* (86% versus 12%), this preference is much smaller for definite noun phrases (58% versus 36%). In the next section, we will evaluate the statistical contribution of the *NPtype* feature, as well as of the other features shown in Table 7.1, in predicting the information status of referential expressions.

7.3 Modeling the data

We employ multinomial logistic regression to investigate if and how the selected features can be used to predict the information status of referential expressions.[3] A logistic regression model allows for statistically assessing the contribution made by individual features in predicting a categorical variable (in this case, information status), due to the transparent nature of the resulting model, which is defined as a non-linear combination of the *intercept* and the values of each of the predictors, corrected by their *weight* (or: *slope*). More specifically, a logistic regression model can be described using the following equation:

$$P(Y) = \frac{1}{1 + e^{-(b_0 + b_1 X_1 + b_2 X_2 + \ldots + b_n X_n)}} \tag{7.2}$$

Here, $P(Y)$ is the probability of Y occurring, b_0 is the Y intercept, each X_i represents a feature, and b_i is the weight associated with the corresponding feature X_i. The multinomial variant of this logistic regression model, then, performs multiple comparisons using the above equation in order to predict categorical variables with more than two categories; in our case, we aim to predict three categories (*given*, *inferred* and *new*), which boils down to two main comparisons: *inferred* versus *given*, and *new* versus *given* (i.e., taking *given* as our reference category). For completeness, we will also report the third comparison (*inferred* versus *new*) in order to be able to compare all individual differences between the categories.

7.3.1 Finding the best model

In order to find the best combination of features, we first ran an all-subsets regression analysis on all 20 predictors from Table 7.1. In this analysis, each of $1,048,575$ models (the power set of our 20 predictors) is compared to the null model (a model that always predicts the most frequent information status category, i.e., *new*) based on the Akaike Information Criterion (AIC; Akaike, 1974). The AIC of a model is defined at its deviance (-2 times LL, the log-likelihood ratio), penalized by the number of parameters (coefficients) that it contains ($2k$), yielding $AIC = -2LL + 2k$. Lower values of AIC indicate better models, and

[3]All statistical analyses were carried out using R version 3.2.0 (R Core Team, 2014) on an Apple Mac Pro. The Multinomial Logit Models were estimated using the `mlogit` package (Croissant, 2013).

models with fewer parameters are favored. As an additional estimate of goodness of fit, we also report McFadden's Pseudo R^2 (McFadden, 1977). For a given model, McFadden's Pseudo R^2 is defined as $1 - LL/LL_{null}$, where LL refers to the log-likelihood ratio of the model, and LL_{null} to that of the null model. Critically, values of McFadden's Pseudo R^2 that indicate good fit are typically lower than that of R^2 in ordinary regression; McFadden's Pseudo R^2 values of 0.2 to 0.4 are considered to indicate excellent fit (McFadden, 1977, pp. 34–35).

The best model resulting from the all-subsets regression ($AIC = 941.70$; $LL = -444.84$; $R^2 = 0.42$) is:

Model 1: *Position* + *Animacy* + *NPlength* + *NPtype* + *Arg1* + *Arg2* + *LemmaOcc* + *SuitAnt* + *PDRSdepth* + *PSites*

This model significantly improves upon the null model ($\chi^2 = 635.99$, $p < 0.0001$) in predicting the information status of referential expressions in terms of the *given*/*inferred*/*new*-distinction. Table 7.4 shows the confusion matrix for Model 1, with the rows indicating the observed choices in the aggregated CrowdFlower data, and the columns indicating the predictions made by Model 1. In this matrix, the numbers on the diagonal indicate correctly predicted values. As is immediately clear from this table, the model never predicts the answer *inferred*; this is not surprising given the small percentage of *inferred* referents overall (4.2%). Otherwise, the model performs well with an overall accuracy of 0.84 ($Acc_{M1:given} = 0.72$; $Acc_{M1:inferred} = 0$; $Acc_{M1:new} = 0.95$), which is a 27% improvement upon the accuracy of the null model ($Acc_{M0} = 0.66$). The $F1$-score of the model, which describes the harmonic mean of its precision and recall, is strongly affected by the lack of *inferred* answers ($F1_{M1} = 0.56$), but it shows an improvement over the null model ($F1_{M0} = 0.27$) and good results for the individual classes *given* ($F1_{M1:given} = 0.77$) and *new* ($F1_{M1:new} = 0.90$). The overall performance of the model is comparable to results obtained by other studies that aim to automatically identify the information status of referential expressions (in particular, Nissim, 2006, presents a decision tree model, which achieves an overall accuracy of 79.5%, and an $F1$-score of 0.67).

Table 7.4: Confusion matrix Model 1

	Predicted			
Observed	*given*	*inferred*	*new*	Total
Aggregated answer = *given*	213	0	83	296
Aggregated answer = *inferred*	13	0	29	42
Aggregated answer = *new*	34	0	619	653
Total	260	0	731	991

Table 7.5 shows the results of the regression model for the comparison between *new* and *given* (the other comparisons of this model–*inferred* versus *given* and *inferred* versus *new*–are shown in Tables 7.10 and 7.11 in Appendix 7.B). This table shows the contribution made per feature to predicting whether the target expression is *new* as opposed to *given*, and whether this contribution is significant. Out of the total of 12 binary predictors[4], 9 are significant ($p < 0.05$) for the comparison at hand.

To assess how the different features contribute to the overall model, we can take a look at their odds ratios. An odds ratio reflects in which direction and how much the odds of predicting one category over another change due to the value of a predictor. This is calculated by dividing the odds of *new* being selected instead of *given* after a single unit change in the predictor $P_{after}(new)/P_{after}(given)$ by the odds before the unit change $P_{before}(new)/P_{before}(given)$; an odds ratio of 1 thus reflects a $50/50$ odds. An odds ratio that is smaller than 1 indicates a decrease in the likelihood of the expression being *new*; in other words, it increases the likelihood of *given*. For instance, *Animacy* has an odds ratio of 0.32, indicating that animate entities are less likely to be categorized as *new* (in other words, animate entities are more likely to be categorized as *given*); more specifically, the odds of non-animate entities to be categorized as *new* are $1/0.32 = 3.125$ times more likely than for animate entities. Other features that decrease the likelihood of *new* are *Lem-*

[4]Categorical predictors with more than two categories (such as *NPtype*) are transformed into binary parameters by treating one of the categories as the baseline category (in this case, the most frequently occurring category *Definite*), and comparing each of the other categories to this baseline category.

Table 7.5: Results Model 1: *new* vs. *given*. The left part of this table lists the regression coefficient (B) for each predictor, along with its standard error (SE) and a significance rating for its p-value (0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1). The right part lists the odds ratio for each predictor (Ratio), as well as the lower bound (Lower) and the upper bound (Upper) of its 95% confidence interval (CI).

New vs. Given			**95% CI for odds ratio**		
	B	**(SE)**	**Lower**	**Ratio**	**Upper**
Intercept	−1.54	(0.64)*			
Position	0.03	(0.02)*	1.00	1.03	1.06
Animacy	−1.13	(0.32)***	0.17	0.32	0.61
NPlength	1.06	(0.22)***	1.88	2.90	4.46
NPtype (vs. Definite)					
Indexical	0.54	(0.88)	0.30	1.72	9.71
Name	2.54	(0.37)***	6.11	12.67	26.28
Possessive	0.75	(0.45).	0.88	2.12	5.11
Arg1	1.38	(0.26)***	2.41	3.99	6.60
Arg2	0.49	(0.24)*	1.03	1.63	2.59
LemmaOcc	−4.30	(0.30)***	0.01	0.01	0.02
SuitAnt	−0.17	(0.06)**	0.74	0.84	0.96
PDRSdepth	0.08	(0.15)	0.81	1.08	1.45
PSites	−0.28	(0.11)*	0.60	0.75	0.94

maOcc (i.e., the odds for *new* drastically decrease—namely $1/0.01 = 100$ times—if the lemma of the target expression occurs in the foregoing context), *SuitAnt* (i.e., the more suitable antecedents in the discourse context, the lower the likelihood of *new* as opposed to *given*), and *PSites* (i.e., the higher the number of accessible projection sites, the higher the likelihood of *given* as opposed to *new*). All other features have an odds ratio that is larger than 1, which indicates an *increase* in the likelihood of the referential expression being *new* (as opposed to *given*). This effect is small but significant for *Position* (i.e., the later in the sentence, the more likely it becomes that the expression is *new*), and even stronger for *NPlength* (i.e., the more tokens the referential expression consists of, the more likely it is to be *new*). For *NPtype*, only the comparison between definite noun phrases and names results in a significant effect on the distinction between *new* and *given*, showing that names are

Table 7.6: Confusion matrix Model 2

Observed	**Predicted** *given*	*inferred*	*new*	Total
Aggregated answer = *given*	194	0	102	296
Aggregated answer = *inferred*	7	0	35	42
Aggregated answer = *new*	19	0	634	653
Total	220	0	771	991

more likely to be categorized as *new* (as opposed to *given*) than definites. Finally, both *Arg1* and *Arg2* also significantly increase the odds of *new* (i.e., modified or modifying expressions have an increased the likelihood of being *new* as opposed to *given*).

To summarize, referential expressions referring to animate entities, whose lemma was already mentioned before, which have more suitable antecedents matching in number and animacy, and have a high number of accessible projection sites are more likely to be tagged as *given*. Conversely, referential expressions that refer to named entities, that occur later in the sentence and consist of multiple tokens, and that are both modified by and modifiers of other referential expressions are more likely to be tagged as *new*. Of these 9 significant predictors, *LemmaOcc* is the strongest predictor for the *given* answer, and *NPtype:Name* is the strongest predictor for the *new* answer. In fact, the model consisting of only these two features performs reasonably well ($AIC = 1043.96$; $LL = -511.98$; $R^2 = 0.33$), with a similar overall accuracy (0.84) and $F1$-score (0.56) as Model 1. The confusion matrix is shown in Table 7.6, and the contribution of the individual features is shown in Table 7.12 (comparing *new* to *given*), Table 7.13 (comparing *inferred* to *given*), and Table 7.14 (comparing *inferred* to *new*) in Appendix 7.B.

Model 2: *NPtype* + *LemmaOcc*

Critically, the confusion matrix of this simplified model shows that the *inferred* category is still severely underrepresented in the predicted results (in fact, it is never predicted). In the next subsection, we will take a closer look at the *inferred* referents, and investigate whether they should be considered a separate category, or if they can be conflated with either the *given* or the *new* referents.

7.3.2 The case of inferred referents

As shown in Table 7.2, of the total number of 991 annotated questions, only 42 were tagged with the answer *inferred* (cf. Adams, 1979), which constitutes 4, 2% of the data. Moreover, the *inferred* answers obtained the lowest trust score overall (0.50). Interestingly, however, the percentage of *inferred* answers in the overall (unaggregated) number of annotations is less dramatic, as these constitute 10.7% of the total number of 7199 data points. This discrepancy suggests that although the *inferred* answer is reasonably often selected by the contributors, it gets the majority vote only in a very small percentage of the questions (see the aggregation procedure described above). This is confirmed by the observation that out of the total of 456 questions that have at least one *inferred* answer, the vast majority (94%) has three or less *inferred* answers (on a minimum of 7 answers per question). After aggregation, the distribution of *given* answers (29.1%) and *new* answers (61.6%) among these 456 questions matches the distribution of the overall data.

In order to investigate the effect of the *inferred* answers on our results, we derived three more models using all-subsets regression: Model 3 was fitted on the data that excludes all questions that obtained *inferred* as the aggregated answer, Model 4 was fitted on the data that conflates these *inferred* referents with the *given* referents, and Model 5 was fitted on the data that conflates the *inferred* referents with the *new* referents. The best results are obtained by Model 3, which excludes the *inferred* referents ($AIC = 594.00$; $LL = -277.00$; $R^2 = 0.53$), followed by the model that conflates *inferred* with *new* (Model 5: $AIC = 645.69$; $LL = -304.85$; $R^2 = 0.50$), and finally the model that conflates *inferred* with *given* (Model 4: $AIC = 714.19$; $LL = -337.09$; $R^2 = 0.47$). Interestingly, all of these model provide a better fit to the data than Model 1 and Model 2. The winning model contains the following features, whose individual contributions are shown in Table 7.7 (the results of Models 4 and 5 can found in Tables 7.15 and 7.16 in Appendix 7.B):

Table 7.7: Results Model 3 (*inferred* referents excluded). The left part of this table lists the regression coefficient (B) for each predictor, along with its standard error (SE) and a significance rating for its p-value (0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1). The right part lists the odds ratio for each predictor (Ratio), as well as the lower bound (Lower) and the upper bound (Upper) of its 95% confidence interval (CI).

New vs. Given			**95% CI for odds ratio**		
	B	**(SE)**	**Lower**	**Ratio**	**Upper**
Intercept	−2.51	(0.74)***			
Position	0.04	(0.01)*	1.01	1.04	1.07
Animacy	−1.07	(0.34)**	0.18	0.34	0.67
NPlength	0.98	(0.25)***	1.64	2.67	4.34
NPtype (vs. Definite)					
Indexical	0.93	(0.89)	0.44	2.53	14.42
Name	2.71	(0.40)***	6.81	15.03	33.16
Possessive	0.48	(0.45)	0.66	1.61	3.92
Role (vs. None)					
Agent	1.21	(0.45)**	1.38	3.35	8.12
Cause	0.40	(0.81)	0.30	1.49	7.35
Experiencer	1.23	(0.41)**	1.52	3.42	7.69
Patient	1.77	(0.68)**	1.56	5.89	22.24
Theme	1.23	(0.50)*	1.27	3.41	9.15
Other	1.97	(0.84)*	1.38	7.18	37.32
Arg1	1.63	(0.28)***	2.95	5.08	8.75
Arg2	1.24	(0.36)***	1.71	3.45	6.97
LemmaOcc	−2.45	(1.10)*	0.01	0.09	0.74
LemmaFreq	−0.96	(0.73)	0.09	0.38	1.58
LemmaRec	−0.05	(0.03).	0.91	0.96	1.01
SuitAnt	−0.16	(0.06)*	0.75	0.85	0.96
PSites	−0.25	(0.08)**	0.67	0.77	0.91

Model 3: *Position* + *Animacy* + *NPlength* + *NPtype* + *Role* + *Arg1* + *Arg2* + *LemmaOcc* + *LemmaFreq* + *LemmaRec* + *SuitAnt* + *PSites*

Model 3 consists of essentially the same set of features as Model 1, with the exclusion of the non-significant feature *PDRSdepth*, and the addition of *Role*, *LemmaFreq* and *LemmaRec* (although the latter two do

Table 7.8: Confusion matrix Model 3

Observed	**Predicted** *given*	*new*	Total
Aggregated answer = *given*	214	82	296
Aggregated answer = *new*	33	620	653
Total	247	702	949

not reach significance; $ps > 0.05$). The features that were also selected in Model 1 contribute similarly to the choice between *new* and *given* in Model 3. In addition, the results show that referential expressions that are associated with a semantic role have an increased likelihood of being tagged as *new* as opposed to *given* relative to expressions that are not associated with any role ('None'); this effect is significant for all roles except 'Cause'. Table 7.8 shows the confusion matrix for Model 3. The accuracy remains high (0.88) and is now consolidated with an equally high $F1$-score of 0.86 (based on a precision score of 0.88, and a recall of 0.84).

7.4 Discussion

We employed multinomial logistic regression modeling to investigate how linguistic features contribute to the information status of referential expressions. More specifically, using a crowdsourcing platform we have collected manual annotations of the information status of referential expressions. Referential expressions were annotated to be either *given*, *inferred*, or *new*, relative to their discourse context. We have then fitted several regression models to predict these annotations of information status on the basis of a set of token-based, contextual and deep semantic features. We found that the distinction between *given* and *new* referential expressions can be successfully accounted for by means of two features: *LemmaOcc* and *NPtype* (see the results of Model 2 in section 7.3.1); the former increases the likelihood of a referential expression being *given* if its lemma occurs in preceding discourse. The latter increases the likelihood of a referential expression being *new* if it is a name or a possessive construction as compared to a definite noun phrase (thereby confirming the aforementioned intuition that names names and definite noun phrases behave differently with respect to their context).

Our models proved less successful in predicting *inferred* referential expressions. The best model obtained by all-subsets regression on the set of 20 features from Table 7.1 never predicted the answer *inferred*. Moreover, when repeating the all-subsets regression on different datasets that either (i) excluded the *inferred* referents (i.e., referents that obtained the aggregated answer *inferred*) altogether, (ii) treated all '*inferred* referents as *given* referents, or (iii) treated all *inferred* referents as *new* referents, we found that option (i) obtained the best results, and that option (iii) performed better than option (ii). This seems to suggest that *inferred* referents indeed constitute a separate category from *given* and *new*, since otherwise the results of conflating the *inferred* referents with either the *given* or the *new* referents would have given the best result. Interestingly, the results show that based on the current set of features, *inferred* referents are more resemblant to *new* referents than to *given* referents, since the model obtained by option (iii) performs better than the model obtained by option (ii). This is in line with previous results on the distinction between bridging inferences and coreference (Nissim, 2006).

In the collected annotation data, the *inferred* answer was very infrequent. This may have several reasons. Firstly, it may be due to the task itself. The instructions (shown in Appendix 7.A) were designed in such a way that contributors would have enough examples of each of the possible answers, without an explicit formulation of the conditions under which one of the answers was correct. This was done in order to give the contributors the freedom to evaluate the questions, as well as the interpretation of the different types of information status, according to their own intuitions. However, the degree to which participants actually followed the instructions is difficult to quantify. This is in particular true with regard to the *inferred* answer, because the control questions used to evaluate the participants only contained questions that had either *given* as the correct answer (in the case of pronouns), or *new* as the correct answer (in the case of indefinite noun phrases). Moreover, in

Table 7.9: Comparison of distributions of definite descriptions

Current study *NPtype = Definite*		**Poesio & Vieira (1998)**		**Gardent e.a. (2003)**	
given	35.8%	coreferential	45.5%	coreferential	16.9%
inferred	6.3%	bridging	8.7%	bridging	4.7%
new	58.0%	larger situation	24.1%	first mention	78.4%
		unfamiliar	21.7%		

addition to being unable to assess the performance of the contributors on *inferred* answers due to the absence of control questions for these answers, this choice of control questions may have biased contributors toward the *given* and *new* answers, despite the indistinguishability of target and control questions.

Interestingly, however, the general lack of *inferred* answers is consistent with the results of previous corpus studies investigating the information status of referential expressions, which focused mainly on definite descriptions. Of particular interest in this context are the studies by Poesio and Vieira (1998) (English data), and Gardent et al. (2003) (French data). Table 7.9 shows an overview of the answer-distribution for definite noun phrases in the current study (*NPtype = Definite*), as compared to the results found by Poesio and Vieira (1998) and Gardent et al. (2003). These distributions are remarkably similar, suggesting that the lack of *inferred* answers in our annotation data is not a task-confound.

Another, and arguably more interesting, explanation for the lack of *inferred* answers is that the distinction between *given*, *inferred*, and *new* that we pursue is not such an intuitive distinction. The three categories of information status can be seen as part of a gradient representing the salience of the described information (similar to the 'Givenness Hierarchy'; Gundel et al., 1993); *given* and *new* represent the two outer limits of this scale, and *inferred* can be interpreted as the grey area between these extremes. Some support for this explanation can be found by looking at the questions that obtained the *inferred* answer in the aggregated data:

(3) Germany's highest court has ruled that random data profiling for terror suspects is legal only when the country faces a specific threat to security or lives. Germany's Constitutional Court ruled Tuesday that the general threat of terror since September 11, 2001, does not warrant random profiling.

(4) Pakistani forces have targeted militants in the northwest for a third day, launching airstrikes that they say killed at least nine suspected insurgents. Helicopter gunships Saturday pounded militant hideouts in the Orakzai tribal region, where many Taliban militants are believed to have fled to avoid an earlier military offensive in nearby South Waziristan.

In (3) the entity referred to by the referential expression "*the general threat*" is clearly not *new* because it is strongly related to "*a specific threat*" in the previous sentence. However, the choice between *inferred* and *given* is less clear, as it strongly depends on the interpretation of the notions '*inferred*' and '*given*'. That is, when interpreted strictly, the referent described by "*the general threat*" is not the same as the referent described by "*a specific threat*" and should therefore be tagged as *inferred*. On a more lenient interpretation of givenness, however, one could argue that (some kind of) "*threat*" was already explicitly introduced, and that "*the general threat*" can therefore be tagged as *given*. In (4), in turn, the referential expression "*Taliban*" is clearly not explicitly mentioned in the preceding discourse (i.e., *given*), but here the choice between *inferred* and *new* is less obvious; depending on one's world knowledge about the relation between the Pakistan and the Taliban, a mention of the former may or may not lead to an inferred interpretation of the latter.

The intuition that *given* and *new* constitute two extremes on a gradient salience scale is also supported by the features that were shown to affect the information status of referential expressions. In section 7.3.1, we have shown that the strongest predictors for distinguishing *new* from *given* are the superficial features *LemmaOcc* and *NPtype*. This suggests that the distinction between *new* and *given* referents can be predicted based on superficial features from the linguistic surface representation; it does not require deeper semantic features, since superficial contextual cues provide sufficient information for determining which of the two extremes is the case. Critically, however, the prediction of *inferred* referents seems to require deeper semantic features that go beyond the superficial token-based properties of referential expressions. In particular, it requires a richer notion of *world knowledge*, which may be captured by semantic features describing the ontological status of expressions relative to each other (e.g., using WordNet Fellbaum, 1998).

7.5 Conclusions

We have investigated how different types of linguistic features contribute to the information status of referential expressions by fitting multinomial logistic regression models to crowdsourced data. Our results indicate that referential expressions that are *given* in the foregoing discourse and referential expressions that are *new* can be successfully distinguished on the basis of two superficial features: one indicating whether the lemma of the referential expression was already used in the discourse before, and another indicating the type of the referential expression (e.g., name, definite noun phrase, or possessive construction). In addition, our results indicate that it is less straightforward to determine whether a referential expression is *inferred* in the discourse context, which can be attributed to the observation that *inferred* referents constitute a gray area of referential expressions in between the two extremes *given* and *new*. Future work should aim at improving upon predicting *inferred* referential expressions by incorporating additional deep semantic features, in particular those encoding a notion of world knowledge, e.g. based on an ontological database like WordNet (Fellbaum, 1998). Moreover, the spectrum of expressions covered by the analysis could be increased, for instance by including conventional implicatures, and to incorporate different genres of texts in order to investigate whether and how genre affects the information status of different expressions.

Appendices

7.A CrowdFlower instructions

What do we want to know?

You are given a text with a word in bold face that refers to an entity (person, organisation, etc), or a set of entities. We want to know what the information status of this entity is: was it already **explicitly** or **implicitly** introduced in the preceding text, or is it **new**?

NB: We do not want to know if the highlighted *word* was already used, but rather if the *entity* described by this word was already introduced before.

Explicitly mentioned before

These are entities that have already been introduced in the preceding text. Note that these may include full or partial repetitions of phrases, as well as a complete rephrasing of the description of the entity. For instance:

- “President Obama wants to [...] **Obama** decided to [...]”
- “Barack Obama has [...] **The president** [...]”
- “Two persons died in the crash. **The first victim** [...]”

Implicitly mentioned before

These entities have not been mentioned directly in the text, but they are not introduced out of the blue either. In some cases, the existence of the entity is implied by the introduction of another entity (based on world knowledge). In other cases, the highlighted entity is indirectly related to an earlier introduced event or referent. For example:

- “A bus was [...] **the driver**”; it is implied that “the driver” mentioned here is the driver *of* the bus that was introduced before (since buses in general have drivers); therefore it is already implicitly introduced.

- "A car bomb that killed 10 people [...] **the incident**"; here, "the incident" refers to the bombing event described before, but wasn't explicitly mentioned.

- "Venezuela's defense ministry decided to [...] **Minister General Raul Baduel** said [...]"; the mentioned "Minister" is implied to be part of an earlier introduced entity, namely "Venezuela's defense ministry". Therefore, it's previous introduction is implicit.

New

Entities that are neither explicit nor implicit in the preceding text are considered new. Examples:

- "John was happy. **Mary** gave him a book."

- "Nuclear envoys from North and South Korea have met in **Beijing** to [...]"

- "[...] a convention in San Francisco, **California**."

7.B Regression models

Tables 7.5 (see Section 7.3.1), 7.10 and 7.11 show the results of Model 1, which is the best model found by all-subsets multinomial logistic regression with the 20 predictors from Table 7.1. Tables 7.12, 7.13, and 7.14 show the results of Model 2, the model resulting after feature selection on Model 1. Tables 7.7 (see Section 7.3.2), 7.15, and 7.16 show the results of Models 3, 4 and 5, respectively, which are obtained from all-subsets multinomial logistic regression with the 20 predictors from Table 7.1, on different variants of the data: excluding all questions that obtained "implicit" as the aggregated answer (Model 3), conflating "implicit" questions with "explicit" questions (Model 4), and conflating "implicit" questions with "new" questions (Model 5).

The left part of each table lists the regression coefficient (B) for each predictor, along with its standard error (SE) and a significance rating for its p-value (0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1). The right part lists the odds ratio for each predictor (Ratio), as well as the lower bound (Lower) and the upper bound (Upper) of its 95% confidence interval (CI).

Table 7.10: Results Model 1: *inferred* vs. *given*

Implicit vs. Explicit			**95% CI for odds ratio**		
	B	**(SE)**	**Lower**	**Ratio**	**Upper**
Intercept	−1.70	(0.97).			
Position	0.00	(0.03)	0.96	1.00	1.06
Animacy	−0.18	(0.47)	0.33	0.84	2.10
NPlength	0.70	(0.29)*	1.14	2.00	3.54
NPtype (vs. Definite)					
Indexical	0.87	(1.29)	0.19	2.38	30.07
Name	0.80	(0.56)	0.75	2.23	6.65
Possessive	−0.32	(0.76)	0.16	0.73	3.20
Arg1	0.77	(0.40).	0.99	2.16	4.73
Arg2	−0.00	(0.38)	0.47	1.00	2.11
LemmaOcc	−2.41	(0.47)***	0.04	0.09	0.22
SuitAnt	−0.09	(0.10)	0.75	0.91	1.11
PDRSdepth	−0.61	(0.28)*	0.32	0.54	0.93
PSites	0.22	(0.18)	0.87	1.24	1.76

Table 7.11: Results Model 1: *inferred* vs. *new*

Implicit vs. New			**95% CI for odds ratio**		
	B	**(SE)**	**Lower**	**Ratio**	**Upper**
Intercept	−0.16	(0.89)			
Position	−0.02	(0.02)	0.93	0.98	1.02
Animacy	0.95	(0.44)*	1.08	2.58	6.15
NPlength	−0.37	(0.24)	0.43	0.69	1.10
NPtype (vs. Definite)					
Indexical	0.32	(1.17)	0.14	1.38	13.79
Name	−1.74	(0.50)***	0.07	0.18	0.47
Possessive	−1.07	(0.66)	0.09	0.34	1.26
Arg1	−0.61	(0.37).	0.26	0.54	1.12
Arg2	−0.49	(0.36)	0.30	0.61	1.24
LemmaOcc	1.88	(0.50)***	2.45	6.58	17.71
SuitAnt	0.08	(0.10)	0.89	1.08	1.31
PDRSdepth	−0.69	(0.26)**	0.30	0.50	0.84
PSites	0.50	(0.17)**	1.19	1.65	2.28

Table 7.12: Results Model 2: *new* vs. *given*

New vs. Explicit			**95% CI for odds ratio**		
	B	**(SE)**	**Lower**	**Ratio**	**Upper**
Intercept	1.23	0.14***			
NPtype (vs. Definite)					
Indexical	−0.14	0.83	0.17	0.87	4.43
Name	1.11	0.22***	1.97	3.03	4.65
Possessive	1.49	0.39***	2.07	4.43	9.48
LemmaOcc	−4.40	0.28***	0.01	0.01	0.02

Table 7.13: Results Model 2: *inferred* vs. *given*

Implicit vs. Explicit			**95% CI for odds ratio**		
	B	**(SE)**	**Lower**	**Ratio**	**Upper**
Intercept	−1.09	0.24***			
NPtype (vs. Definite)					
Indexical	0.40	1.25	0.13	1.49	17.20
Name	0.00	0.38	0.47	1.00	2.13
Possessive	0.18	0.70	0.30	1.20	4.74
LemmaOcc	−2.24	0.45***	0.04	0.11	0.26

Table 7.14: Results Model 2: *inferred* vs. *new*

Implicit vs. New			**95% CI for odds ratio**		
	B	**(SE)**	**Lower**	**Ratio**	**Upper**
Intercept	−2.33	0.22***	0.06	0.10	0.15
Indexical	0.53	1.10	0.20	1.71	14.81
Name	−1.11	0.36**	0.16	0.33	0.67
Possessive	−1.31	0.63*	0.08	0.27	0.92
LemmaOcc	2.16	0.49***	3.31	8.66	22.61

Table 7.15: Results Model 4 (*inferred* conflated with *given*)

New vs. Explicit			**95% CI for odds ratio**		
	B	**(SE)**	**Lower**	**Ratio**	**Upper**
Intercept	−1.90	(0.62)**			
Position	0.04	(0.01)**	1.01	1.04	1.07
Animacy	−1.09	(0.30)***	0.19	0.34	0.60
NPlength	0.64	(0.19)***	1.30	1.90	2.76
NPtype (vs. Definite)					
Indexical	0.52	(0.78)	0.37	1.68	7.76
Name	2.24	(0.33)***	4.88	9.35	17.90
Possessive	0.77	(0.40).	0.98	2.15	4.72
Role (vs. None)					
Agent	0.97	(0.39)*	1.22	2.63	5.71
Cause	0.48	(0.72)	0.40	1.62	6.66
Experiencer	1.13	(0.36)**	1.54	3.09	6.24
Patient	1.41	(0.62)*	1.21	4.08	13.73
Theme	0.65	(0.43)	0.83	1.92	4.45
Other	1.74	(0.75)*	1.33	5.71	24.62
Arg1	1.34	(0.24)***	2.38	3.81	6.10
Arg2	1.16	(0.32)***	1.72	3.19	5.92
LemmaOcc	−1.87	(1.07).	0.02	0.15	1.25
LemmaFreq	−1.06	(0.72)	0.08	0.35	1.43
LemmaRec	−0.05	(0.03).	0.90	0.95	1.00
SuitAnt	−0.17	(0.06)**	0.75	0.84	0.95
PSites	−0.22	(0.07)**	0.70	0.81	0.93

Table 7.16: Results Model 5 (*inferred* conflated with *new*)

New vs. Explicit			**95% CI for odds ratio**		
	B	**(SE)**	**Lower**	**Ratio**	**Upper**
Intercept	−1.76	(0.69)*			
Position	0.02	(0.01).	1.00	1.02	1.05
Animacy	−0.97	(0.32)**	0.20	0.38	0.71
NPlength	0.82	(0.23)***	1.46	2.28	3.57
NPtype (vs. Definite)					
Indexical	0.76	(0.86)	0.40	2.14	11.50
Name	2.28	(0.37)***	4.68	9.74	20.27
Possessive	0.53	(0.45)	0.70	1.70	4.09
Role (vs. None)					
Agent	1.02	(0.42)*	1.21	2.78	6.38
Cause	0.59	(0.79)	0.38	1.81	8.48
Experiencer	1.26	(0.39)**	1.65	3.54	7.61
Patient	1.46	(0.64)*	1.22	4.29	15.03
Theme	1.07	(0.47)*	1.15	2.92	7.41
Other	1.76	(0.76)*	1.32	5.82	25.66
Arg1	1.50	(0.26)***	2.69	4.50	7.53
Arg2	1.22	(0.34)***	1.72	3.38	6.63
LemmaOcc	−4.22	(0.30)***	0.01	0.01	0.03
SuitAnt	−0.16	(0.06)**	0.75	0.85	0.96
PSites	−0.22	(0.07)**	0.69	0.80	0.93

Chapter 8

Conclusion

Projection refers to the indifference of linguistic content to the syntactic scope of entailment-cancelling operators, such as negation, modals, and implication. Projection behavior is traditionally associated with *presuppositions*, which have been the object of study in the linguistic literature for over a century, dating back to the works by Frege (1892) and Russell (1905). The work presented in this thesis is part of a recent development in the study of linguistic meaning (see, e.g., Simons et al., 2010; Tonhauser et al., 2013) that aims to provide a unified analysis of the different types of expressions that have been associated with the property of projection, including (different types of) presuppositions, anaphoric expressions, and conventional implicatures (as defined by Potts, 2005). Critically, any unified analysis should be able to account for the similarities between the different types of expressions—as demonstrated by the shared property of projection—as well as for the observed differences between them in terms of the contribution they make to the unfolding discourse context. In this thesis, I presented an analysis of projection phenomena in which this requirement is met by defining the differences and similarities between these phenomena in terms of the *information status* of their contribution. This analysis was implemented as part of the formal semantic framework Projective Discourse Representation Theory, and formed the basis for an empirical investigation of the information status of referential expressions. This formal and empirical analysis of projection phenomena provides the starting point for a more elaborate, data-driven and unified approach to projection in discourse.

Below, I will summarize the most important results from the individual parts of this thesis and indicate directions for future work.

8.1 Part I: Categorizing projection phenomena

The first part of this thesis presents an overview of the different phenomena that have been associated with the property of projection, and introduces the notion of *information status*, which refers to whether linguistic content is *given*, *backgrounded*, or *foregrounded* relative to the current discourse context. In Chapter 2, I presented a categorization of different types of projected content, including anaphoric expressions, strong and weak presuppositions, and conventional implicatures, based on their acceptability in the different contexts constituted by these three types of information status. This categorization explains the shared property of projection based on the observation that each of these different types of expressions can describe common ground information that is either *given* or *backgrounded*. Critically, this analysis also explains the relation between the different types of projected content and asserted content without introducing multiple dimensions of meaning; assertions are argued to be felicitous only in contexts in which their contribution is *foregrounded*, whereas presuppositions and conventional implicatures pose less restrictive constraints on their context.

8.2 Part II: A semantic formalism for projection

In the second part of this thesis, this unified analysis of projected and asserted content is formalized as part of Projective Discourse Representation Theory (PDRT): an extension of traditional Discourse Representation Theory (Kamp, 1981; Kamp and Reyle, 1993) that provides a parsimonious treatment of the projection behavior of different types of phenomena. In Chapter 3, I introduced PDRT as an extension of van der Sandt's (1992) idea to treat presupposition projection as anaphora resolution. In PDRT, all semantic content is associated with 'projection variables', which explicitly represent the relation between the introduction site of the semantic content and where this content is interpreted. This way, differences in information status become an explicit part of the semantic representations. It was shown that the resulting formal representations account for the different contributions made by projected and asserted content, without assuming a two-stage resolution procedure for presupposed content (see van der Sandt, 1992), or multiple dimensions of meaning (as in, for instance, Layered DRT; Geurts and Maier, 2003).

In Chapter 4, I employed this unidimensional approach to projection to formulate an analysis of the projection behavior of conventional implicatures (CIs; Potts, 2005). CIs were argued to require an *anchor* that refers to a specific entity in the discourse context. The contribution made by the CI, then, can be formulated as an elaboration on the description of the referent referred to by its projecting anchor. I showed that this analysis of CIs can be formalized in the semantic representations from PDRT, by treating CIs as 'piggybacking' on their projecting anchor; they introduce an anaphoric dependency on the interpretation site of their anchor, while at the same time requiring their anchor to project. It was shown that extending PDRT with explicit constraints on projection behavior enriches the representation of the information status of semantic content and thereby increases the representational power of the formalism.

The practical implications of extending traditional Discourse Representation Theory with a notion of information status by means of projection variables were described in Chapter 5. It was argued that basic DRT notions such as variable binding are affected by the introduction of projection variables in a non-trivial manner. In order to establish PDRT as a robust, widely applicable semantic framework, the basic definitions regarding the syntax, accessibility constraints, and composition of PDRSs were formally worked out. This formalization showed that PDRT extends all representational properties of traditional DRT, and in addition captures differences in information status in a parsimonious way. As a proof of concept, the formal definitions underlying the PDRT framework were implemented as an NLP library called "PDRT-SANDBOX"; this implementation incorporates all structural and compositional aspects of PDRT, as well as its translation to traditional DRT and first-order logic (see Appendix A).

8.3 Part III: A data-driven analysis of projection phenomena

In the third and final part of this thesis, the unified analysis of projection phenomena was employed in a data-driven approach to formal semantics. Chapter 6 describes the development of the Groningen Meaning Bank, a large corpus of public domain texts annotated with the semantic structures from PDRT. I presented the different levels of annotation underlying these semantic analyses, together with the concept of 'meaning banking'; automatically deriving deep semantic representations for large amounts of texts and bootstrapping these analyses by collecting linguistic annotations from different sources. In order to provide a satisfactory trade-off between the quality and quantity of linguistic annotations, different annotation platforms were developed for collecting annotations from expert as well as non-expert linguists. A preliminary analysis of the data collected via a 'Game with a Purpose' called *Wordrobe*, showed that high quality annotations can be obtained using crowdsourcing methods, provided that the task at hand can be explained using minimal linguistic terminology (Venhuizen et al., 2013a).

In Chapter 7, I presented a data-driven analysis of the information status of referential expressions, investigating which token-based, contextual, and deep semantic features affect how the contribution of a referential expression is related to the discourse context. The results indicated that the type of a referential expression (e.g., 'name' versus 'definite description') indeed affects the likelihood of the expression being categorized as *given* or *new* with respect to the discourse context. It was shown that this choice was affected even more strongly by the superficial contextual feature describing whether the lemma of the expression was already introduced in the text before. Moreover, the results indicated that *inferred* referential expressions were less straightforward to predict. I argued that this class should be interpreted as part of the middle spectrum of an information status gradient that ranges from *given* to *new*. This result supports the hypothesis that information status is not a categorical property. In terms of the categorization presented in Chapter 2, this means that the categories representing *given*, *backgrounded*, and *foregrounded* information should be considered abstractions over the information status gradient. Moreover, the identification of superficial contextual features as the most important factors for predicting the information status of referential expressions, paves way toward an empirically driven formalization of the information status of different types of semantic content.

8.4 Future directions in data-driven formal semantics

In this thesis, I combined a formal analysis of projection with a data-driven account of the phenomena exhibiting this property. As a result, the directions for future work can be described for both aspects of the analysis. On the one hand, the formal semantic representations can be extended as to incorporate a wider range of phenomena (e.g., including indexical expressions, following Hunter, 2010). On the other hand, the data-driven analysis must explore a wider range of expressions, features and data-sets, in order to obtain a better understanding of the behavior of projection phenomena in discourse. Critically, however, it is the combination of these different strands of research that presents the most interesting challenge for future work (Pulman, 2007).

One way of combining formal analyses and data-driven approaches is exemplified by the integration of information from linguistic resources like FrameNet (Baker et al., 1998) into the formal structures from DRT and Segmented DRT (Bos and Nissim, 2008; Irmer, 2009, 2013). This appears to be a promising direction for future work, where results from the natural language processing community can be used to inform and enrich formal semantic representations. Conversely, the development of semantically annotated corpora, such as the Groningen Meaning Bank (Bos et al., 2015, see Chapter 6), as well as the implementation of semantic formalisms in applications like Boxer (Bos, 2003) and PDRT-SANDBOX (Venhuizen and Brouwer, 2014, see Appendix A), inspire the utilization of formal semantic analyses in large-scale computational investigations of natural language. I believe that this data-driven approach to formal semantics will prove to be crucial for a robust and wide-coverage analysis of linguistic meaning.

Appendix

Appendix A

PDRT-SANDBOX*

Abstract. We introduce PDRT-SANDBOX, a Haskell library that implements Projective Discourse Representation Theory (PDRT; Venhuizen et al., 2013b), an extension of Discourse Representation Theory (DRT; Kamp, 1981; Kamp and Reyle, 1993). The implementation includes a translation from PDRT to DRT and first-order logic, composition via different types of merge, and unresolved structures based on Montague Semantics (Muskens, 1996), defined as Haskell functions.

A.1 Introduction

The semantic property of projection, traditionally associated with presuppositions, has challenged many structure-driven formal semantic analyses. Linguistic content is said to project if it is interpreted outside the scope of an operator that syntactically subordinates it. In semantic formalisms, this behaviour has often been treated as a *deviation* from standard meaning construction, despite the prevalence of expressions exhibiting it (van der Sandt, 1992; Geurts, 1999; Beaver, 2001). By contrast, we have proposed a formalism that *centralizes* the property of projection as a strategy for integrating material into the foregoing con-

*Appendix adapted from: Venhuizen, Noortje and Harm Brouwer (2014). PDRT-SANDBOX: An implementation of Projective Discourse Representation Theory. In Rieser, Verena and Philippe Muller, editors, *Proceedings of the 18th Workshop on the Semantics and Pragmatics of Dialogue (DialWatt - SemDial 2014)*, pages 249–251, Edinburgh text. This formalism is called Projective Discourse Representation Theory (PDRT; Venhuizen et al., 2013b), and is an extension of the widely used framework Discourse Representation Theory (DRT; Kamp, 1981; Kamp and Reyle, 1993). In PDRT, all linguistic material is associated with a *pointer* to indicate its interpretation site. In this way, an explicit distinction is made between the surface form of an utterance, and its logical interpretation. The formalism can account for various projection phenomena, including presuppositions (Venhuizen et al., 2013b) and Potts' 2005 conventional implicatures (Venhuizen et al., 2014b), and has already been integrated into the Groningen Meaning Bank (Basile et al., 2012a).

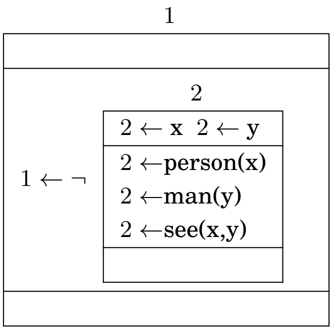
Critically, adding projection pointers to all linguistic material affects the formal properties of DRT non-trivially; the occurrence of projected material at the interpretation site results in non-hierarchical variable binding, and violates the traditional DRT notion of context accessibility, thereby compromising the basic construction mechanism. Here, we present an updated construction mechanism as part of a Haskell library called PDRT-SANDBOX that implements PDRT, as well as standard DRT. The implementation incorporates definitions for building and combining structures, translating Projective Discourse Representation Structures (PDRSs) to Discourse Representation Structures (DRSs) and first-order logic (FOL) formulas, and dealing with unresolved structures via lambda abstractions (Muskens, 1996). Moreover, it allows for various input and output representations, and is highly modular, thereby providing a full-fledged toolkit for use in other NLP applications.

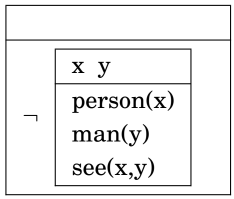
A.2 Projective Discourse Representation Theory

PDRSs carry more information than DRSs; in addition to the structural and referential content of a DRS, a PDRS also makes the information structure of a discourse explicit by keeping linguistic content at its introduction site, and indicating the interpretation site via a projection variable. That is, each PDRS introduces a *label* that can be used as an identifier, and all of its referents and conditions are associated with a *pointer*, which is used to indicate in which context the material is *interpreted* by means of binding it to a context label.

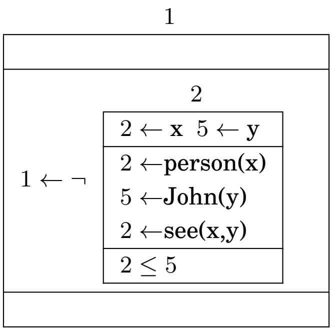
Examples (1) and (2) show two PDRSs and their corresponding DRSs. An important addition to the PDRS definitions described in Venhuizen et al. (2013b), is the introduction of *Minimally Accessible Projection contexts* (MAPs) in the footer of each PDRS. These MAPs pose minimal constraints on the accessibility of projection contexts, creating a partial order over PDRS contexts (Reyle, 1993, 1995).

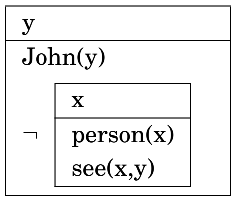
(1) Nobody sees a man.

a. 

b. 

(2) Nobody sees John.

a. 

b. 

In the PDRS in (1a), all pointers are bound by the label of the PDRS in which the content is introduced, indicating *asserted material*. As shown in (1b), this representation is identical to the standard DRT representation of this sentence, except for the addition of labels to PDRSs and pointers to all referents and conditions. In (2), on the other hand, the proper name "John" triggers a presupposition about the existence of someone called 'John'. The pointer associated with the referent and condition describing this presupposition indicates *projected material*; it occurs free, as it is not bound by the label of any accessible PDRS. This means that no antecedent has been found yet. In the corresponding DRS in (2b) the presupposition is accomodated at the most global accommodation site. Note that in contrast to the DRT representation, the accommodation site of the presupposition is not determined in the PDRS; (2a) only stipulates that the accommodation site should be accessible from the introduction site of the presupposition. This flexibility of interpretation increases the compositionality of PDRT, since more context may become available later on in which the presupposition becomes bound. In combination with MAPs, this property can also be exploited to account for the projection behaviour of conventional implicatures (Venhuizen et al., 2014b).

A.3 Playing in the sandbox

We implemented the formal definitions for the construction and manipulation of the structures of PDRT and standard DRT in a Haskell library called PDRT-SANDBOX. For a full description of all definitions, see Venhuizen et al. (2014a). The library provides the following core features:

- **Definitions for building and combining (P)DRSs.** The binding and accessibility definitions in DRT and PDRT are fully worked out, and applied as conditions on combining (*merging*) structures and resolving them. Two different types of merge are defined for PDRT: *projective merge* and *assertive merge* (Venhuizen et al., 2013b).

- **Translations.** PDRSs can be translated to DRSs, FOL-formulas, and flat (non-recursive) representations called P-Tables.

- **Lambda abstractions.** Unresolved structures obtain Montague-style representations, following Muskens (1996). The implementation exploits Haskell's lambda-theoretic foundations by formalising unresolved structures as Haskell functions, thereby profiting from all existing associated functionality.

- **Various input and output formats.** As (P)DRS output format, the standard "boxes" representation is available, as well as a linear representation of the boxes, a set-theoretic representation, and the internal syntax for (P)DRSs. The latter two are also recognised as input formats, along with the Prolog syntax from Boxer (Bos, 2003).

A.4 Conclusion

PDRT-SANDBOX is a full-fledged NLP library for constructing and manipulating the discourse structures from DRT and PDRT, which can be used as part of a larger NLP architecture. One direction would be combining the implementation with a syntactic parser, resulting in a tool-chain similar to the one created by the C&C tools and Boxer (Curran et al., 2007). Furthermore, the representations produced by PDRT-SANDBOX may be applied in a separate model checker, QA system, or any other NLP tool that uses deep semantic representations. PDRT-SANDBOX is freely available (under the Apache License, Version 2.0) at: `http://hbrouwer.github.io/pdrt-sandbox/`

Bibliography

Abusch, Dorit (2002). Lexical alternatives as a source of pragmatic presuppositions. In *Proceedings of SALT*, volume 12, pages 1–19.

Abusch, Dorit (2010). Presupposition triggering from alternatives. *Journal of Semantics*, 27(1):37–80.

Adams, Douglas Noel (1979). *The Hitchhiker's Guide to the Galaxy*. Pan Books.

Akaike, Hirotsugu (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.

Amaral, Patricia, Craige Roberts, and E. Allyn Smith (2007). Review of 'The Logic of Conventional Implicatures' by Chris Potts. *Linguistics and Philosophy*, 30(6):707–749.

AnderBois, Scott, Adrian Brasoveanu, and Robert Henderson (2010). Crossing the appositive/at-issue meaning boundary. In *Proceedings of SALT*, volume 20, pages 328–346. CLC Publications.

Ariel, Mira (1988). Referring and accessibility. *Journal of Linguistics*, 24(1):65–87.

Asher, Nicholas (1986). Belief in discourse representation theory. *Journal of Philosophical Logic*, 15(2):127–189.

Asher, Nicholas (1989). Belief, acceptance and belief reports. *Canadian journal of philosophy*, 19(3):327–361.

Asher, Nicholas (1993). *Reference to abstract objects in discourse*. Kluwer, Dordrecht.

Asher, Nicholas and Alex Lascarides (1998). Bridging. *Journal of Semantics*, 15(1):83–113.

Asher, Nicholas and Alex Lascarides (2003). *Logics of Conversation*. Cambridge University Press.

Bach, Kent (1999). The myth of conventional implicature. *Linguistics and Philosophy*, 22(4):367–421.

Baker, Collin F., Charles J. Fillmore, and John B. Lowe (1998). The Berkeley FrameNet project. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics*, volume 1, pages 86–90. Association for Computational Linguistics.

Basile, Valerio and Johan Bos (2013). Aligning formal meaning representations with surface strings for wide-coverage text generation. In *Proceedings of the 14th European Workshop on Natural Language Generation*, pages 1–9, Sofia, Bulgaria. Association for Computational Linguistics.

Basile, Valerio, Johan Bos, Kilian Evang, and Noortje J. Venhuizen (2012a). Developing a large semantically annotated corpus. In Calzolari, Nicoletta, Khalid Choukri, Thierry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk, and Stelios Piperidis, editors, *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, pages 3196–3200, Istanbul, Turkey. European Language Resources Association (ELRA).

Basile, Valerio, Johan Bos, Kilian Evang, and Noortje J. Venhuizen (2012b). A platform for collaborative semantic annotation. In Daelemans, Walter, Mirella Lapata, and Lluís Màrquez, editors, *Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 92–96, Avignon, France. Association for Computational Linguistics.

Basile, Valerio, Johan Bos, Kilian Evang, and Noortje J. Venhuizen (2012c). UGroningen: Negation detection with Discourse Representation Structures. In Agirre, Eneko, Johan Bos, and Mona Diab, editors, *Proceedings of the First Joint Conference on Lexical and Computational Semantics (*SEM 2012)*, volume 1, pages 301–309. Association for Computational Linguistics.

Beaver, David I. (1992). The kinematics of presupposition. In Dekker, Paul and Martin Stokhof, editors, *Proceedings of the Eighth Amsterdam Colloquium*, pages 17–36. University of Amsterdam, ILLC.

Beaver, David I. (1997). Presupposition. In van Benthem, Johan and Alice ter Meulen, editors, *The Handbook of Logic and Language*, pages 939–1008. Elsevier, Amsterdam.

Beaver, David I. (2001). *Presupposition and Assertion in Dynamic Semantics*. CSLI Publications, Stanford, CA.

Beaver, David I. (2002). Presupposition in DRT. In Beaver, David I., Luis D. Casillas Martínez, Brady Z. Clark, and Stefan Kaufmann, editors, *The Construction of Meaning*. CSLI Publications, Stanford, California.

Beaver, David I. and Bart Geurts (2011). Presupposition. In Zalta, Edward N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, CSLI, Stanford University, summer 2011 edition.

Bjerva, Johannes (2014). Multi-class animacy classification with semantic features. In Wintner, Shuly, Desmond Elliott, Konstantina Garoufi, Douwe Kiela, and Ivan Vulić, editors, *Proceedings of the Student Research Workshop at the 14th Conference of the European Chapter of the Association for Computational Linguistics*, pages 65–75, Gothenburg, Sweden. Association for Computational Linguistics.

Bos, Johan (2003). Implementing the binding and accommodation theory for anaphora resolution and presupposition projection. *Computational Linguistics*, 29(2):179–210.

Bos, Johan (2008). Wide-coverage semantic analysis with Boxer. In Bos, Johan and Rodolfo Delmonte, editors, *Semantics in Text Processing. STEP 2008 Conference Proceedings*, volume 1 of *Research in Computational Semantics*, pages 277–286.

Bos, Johan (2014). Semantic annotation issues in parallel meaning banking. In Bunt, Harry, editor, *Proceedings of the 10th Joint ISO-ACL SIGSEM Workshop on Interoperable Semantic Annotation*, pages 17–20.

Bos, Johan, Valerio Basile, Kilian Evang, Noortje J. Venhuizen, and Johannes Bjerva (2015). The Groningen Meaning Bank. In Ide, Nancy and James Pustejovsky, editors, *Handbook of Linguistic Annotation – Part Two: Case studies*, Text, Speech and Language Technology. Springer. (forthcoming).

Bos, Johan, Paul Buitelaar, and Anna-Marie Mineur (1995). Bridging as coercive accommodation. In *Proceedings to the Workshop on Computational Logic for Natural Language Processing (CLNLP)*, Edinburgh.

Bos, Johan, Kilian Evang, and Malvina Nissim (2012). Annotating semantic roles in a lexicalised grammar environment. In Bunt, Harry, editor, *Proceedings of the 8th Joint ISO-ACL SIGSEM Workshop on Interoperable Semantic Annotation*, pages 9–12, Pisa. ILC-CNR.

Bos, Johan and Malvina Nissim (2008). Combining discourse representation theory with framenet. *Frames, Corpora, and Knowledge Representation*, pages 169–183.

Bresnan, Joan, Anna Cueni, Tatiana Nikitina, and R. Harald Baayen (2007). Predicting the dative alternation. In Bouma, Gerlof, Irene Krämer, and Joost Zwarts, editors, *Cognitive Foundations of Interpretation*, pages 69–94. Royal Netherlands Academy of Arts and Sciences.

Burkhardt, Petra (2006). Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. *Brain and Language*, 98(2):159–168.

Chafe, Wallace L. (1976). Givenness, contrastiveness, definiteness, subjects, topics and points of view. In Li, Charles N., editor, *Subject and Topic*, pages 27–55. Academic Press.

Chierchia, Gennaro (1995). *Dynamics of meaning: Anaphora, presupposition, and the theory of grammar*. University of Chicago Press.

Chierchia, Gennaro and Sally McConnell-Ginet (1990). *Meaning and Grammar: an introduction to semantics*. MIT Press, Cambridge, Mass.

Clark, Herbert H. (1975). Bridging. In Schank, Roger C. and Bonnie L. Nash-Webber, editors, *Proceedings of the 1975 workshop on Theoretical Issues in Natural Language Processing*, pages 169–174, Stroudsburg, PA, USA. Association for Computational Linguistics.

Clark, Herbert H. and Susan E Haviland (1977). Comprehension and the given-new contract. *Discourse Production and Comprehension. Discourse Processes: Advances in Research and Theory*, 1:1–40.

Croissant, Yves (2013). *mlogit: multinomial logit model*. R package version 0.2-4.

Curran, James R., Stephen Clark, and Johan Bos (2007). Linguistically motivated large-scale NLP with C&C and Boxer. In Ananiadou, Sophia, editor, *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics Companion Volume Proceedings of the Demo and Poster Sessions*, pages 33–36. Association for Computational Linguistics.

Del Gobbo, Francesca (2003). *Appositives at the Interface*. PhD Thesis, University of California, Irvine.

Dell'Orletta, Felice, Alessandro Lenci, Simonetta Montemagni, and Vito Pirrelli (2005). Climbing the path to grammar: A maximum entropy model of subject/object learning. In *Proceedings of the Workshop on Psychocomputational Models of Human Language Acquisition*, pages 72–81. Association for Computational Linguistics.

van Eijck, Jan and Hans Kamp (1997). Representing discourse in context. In van Benthem, Johan and Alice ter Meule, editors, *Handbook of logic and language*, pages 179–237. Elsevier, Amsterdam.

van Eijck, Jan and Albert Visser (2012). Dynamic semantics. In Zalta, Edward N., editor, *The Stanford Encyclopedia of Philosophy*. Winter 2012 edition.

Evang, Kilian, Valerio Basile, Grzegorz Chrupala, and Johan Bos (2013). Elephant: Sequence labeling for word and sentence segmentation. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (EMNLP 2013)*, pages 1422–1426.

Evang, Kilian and Johan Bos (2013). Scope disambiguation as a tagging task. In *Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013)*, pages 314–320.

Evans, Gareth (1977). Pronouns, quantifiers, and relative clauses (i). *Canadian journal of philosophy*, 7(3):467–536.

Fellbaum, Christine, editor (1998). *WordNet: An Electronic Lexical Database*. MIT Press.

Fraurud, Kari (1990). Definiteness and the processing of noun phrases in natural discourse. *Journal of Semantics*, 7(4):395–433.

Frege, Gottlob (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100:25–50.

Gardent, Claire, Hélène Manuélian, and Eric Kow (2003). Which bridges for bridging definite descriptions. In *Proceedings of the EACL 2003 Workshop on Linguistically Interpreted Corpora*, pages 69–76.

Gazdar, Gerald (1979a). *Pragmatics: Implicature, presupposition, and logical form*. Academic Press, New York.

Gazdar, Gerald (1979b). A solution to the projection problem. In Oh, Choon-Kyu and David Dineen, editors, *Syntax and Semantics*, volume 11, pages 57–89. Academic Press, New York, presupposition edition.

Geach, Peter Thomas (1962). *Reference and generality: An examination of some medieval and modern theories*, volume 88. Cornell University Press.

van Geenhoven, Veerle (1998). *Semantic incorporation and indefinite descriptions: Semantic and syntactic aspects of noun incorporation in West Greenlandic*. CSLI Publications, Stanford, California.

Geurts, Bart (1997). Good news about the description theory of names. *Journal of Semantics*, 14(4):319–348.

Geurts, Bart (1999). *Presuppositions and pronouns*. Elsevier.

Geurts, Bart (2007). Really fucking brilliant. *Theoretical Linguistics*, 33:209–271.

Geurts, Bart (2010). Specific indefinites, presupposition and scope. In Bäuerle, Rainer, Uwe Reyle, and Thomas E. Zimmermann, editors, *Presuppositions and Discourse: Essays offered to Hans Kamp*, pages 125–158. Emerald, Bingley.

Geurts, Bart and David I. Beaver (2011). Discourse Representation Theory. In Zalta, Edward N., editor, *The Stanford Encyclopedia of Philosophy*. Fall 2011 edition.

Geurts, Bart and Emar Maier (2003). Layered DRT. Ms. University of Nijmegen.

Geurts, Bart and Emar Maier (2013). Layered Discourse Representation Theory. In Capone, Alessandro, Franco Lo Piparo, and Marco Carapezza, editors, *Perspectives on Linguistic Pragmatics*, pages 311–327. Springer International Publishing.

Geurts, Bart and Rob A. van der Sandt (2004). Interpreting focus. *Theoretical Linguistics*, 30:1–44.

Grice, H. Paul (1975). Logic and conversation. In Cole, Peter and Jerry L. Morgan, editors, *Syntax and semantics*, volume 3 of *Speech acts*, pages 41–58. Academic Press, San Diego, CA.

Groenendijk, Jeroen A. G. and Martin J. B. Stokhof (1991). Dynamic predicate logic. *Linguistics and Philosophy*, 14(1):39–100.

de Groote, Philippe (2006). Towards a Montagovian account of dynamics. In Gibson, Masayuki and Jonathan Howell, editors, *Proceedings of SALT*, volume 16, pages 1–16.

Grouin, Cyril, Sophie Rosset, Pierre Zweigenbaum, Karën Fort, Olivier Galibert, and Ludovic Quintard (2011). Proposal for an extension of traditional named entities: From guidelines to evaluation, an overview. In *Proceedings of the Fifth Linguistic Annotation Workshop*, pages 92–100. Association for Computational Linguistics.

Gundel, Jeanette K., Nancy Hedberg, and Ron Zacharski (1993). Cognitive status and the form of referring expressions in discourse. *Language*, pages 274–307.

Heim, Irene (1982). *The semantics of definite and indefinite noun phrases*. PhD thesis, UMass, Amherst.

Heim, Irene (1983). On the projection problem for presuppositions. In *Proceedings of the West Coast Conference on Formal Linguistics*, volume 2, pages 144–226.

Heim, Irene (1992). Presupposition projection and the semantics of attitude verbs. *Journal of Semantics*, 9(3):183–221.

Heringa, Herman (2012). *Appositional Constructions*. PhD Thesis, University of Groningen.

Hockenmaier, Julia and Mark Steedman (2007). CCGbank: a corpus of CCG derivations and dependency structures extracted from the Penn Treebank. *Computational Linguistics*, 33(3):355–396.

Hoeks, John C. J. and Harm Brouwer (2014). Electrophysiological research on conversation and discourse processing. In Holtgraves, Thomas M., editor, *The Oxford Handbook of Language and Social Psychology*, chapter 23, pages 365–386. Oxford University Press, New York.

Hovy, Eduard, Mitchell Marcus, Martha Palmer, Lance Ramshaw, and Ralph Weischedel (2006). OntoNotes: the 90% solution. In *Proceedings of the Human Language Technology Conference of the NAACL, Companion Volume: Short Papers*, pages 57–60. Association for Computational Linguistics.

Hunter, Julie (2010). *Presuppositional Indexicals*. PhD thesis, University of Texas.

Hunter, Julie (2013). Presuppositional indexicals. *Journal of Semantics*, 30(3):381–421.

Hunter, Julie and Nicholas Asher (2005). A presuppositional account of indexicals. In Dekker, Paul and Michael Franke, editors, *Proceedings of the Fifteenth Amsterdam Colloquium*, volume 201, pages 119–124.

Irmer, Matthias (2009). *Bridging Inferences in Discourse Interpretation*. PhD thesis, University of Leipzig.

Irmer, Matthias (2013). Inferring implicatures and discourse relations from frame information. *Lingua*, 132:29–50.

Kadmon, Nirit (2001). *Formal Pragmatics: Semantics, Pragmatics, Preposition, and Focus*. Wiley-Blackwell.

Kamp, Hans (1981). A theory of truth and semantic representation. In Groenendijk, Jeroen A. G., Theo M. V. Janssen, and Martin B. J. Stokhof, editors, *Formal Methods in the Study of Language*, number 135, pages 277–322. Mathematisch Centrum.

Kamp, Hans, Josef van Genabith, and Uwe Reyle (2011). Discourse Representation Theory. In Gabbay, Dov M. and Franz Guenthner, editors, *Handbook of Philosophical Logic*, volume 15, pages 125–394. Springer Netherlands.

Kamp, Hans and Uwe Reyle (1993). *From discourse to logic: Introduction to modeltheoretic semantics of natural language, formal logic and Discourse Representation Theory*. Kluwer, Dordrecht.

Karttunen, Lauri (1971). Some observations on factivity. *Research on Language & Social Interaction*, 4(1):55–69.

Karttunen, Lauri (1973). Presuppositions of compound sentences. *Linguistic inquiry*, 4(2):169–193.

Karttunen, Lauri (1974). Presupposition and linguistic context. *Theoretical linguistics*, 1(1-3):181–194.

Karttunen, Lauri and Annie Zaenen (2005). Veridicity. In Katz, Graham, James Pustejovsky, and Frank Schilder, editors, *Annotating, Extracting and Reasoning about Time and Events*, number 05151 in Dagstuhl Seminar Proceedings, pages 1–9, Dagstuhl, Germany. Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI).

Kasher, Asa and Dov M Gabbay (1976). On the semantics and pragmatics of specific and non-specific indefinite expressions. *Theoretical Linguistics*, 3(1-3):145–190.

Kilgarriff, Adam and Joseph Rosenzweig (2000). Framework and results for English SENSEVAL. *Computers and the Humanities*, 34(1-2):15–48.

King, Jeffrey C. (2013). Anaphora. In Zalta, Edward N., editor, *The Stanford Encyclopedia of Philosophy*. Summer 2013 edition.

Kipper, Karin, Anna Korhonen, Neville Ryant, and Martha Palmer (2008). A large-scale classification of English verbs. *Language Resources and Evaluation*, 42(1):21–40.

Kleene, Stephen Cole (1952). *Introduction to Metamathematics*. North-Holland, Amsterdam.

Koev, Todor (2014). Two puzzles about appositives: Projection and perspective shift. In Etxeberria, Urtzi, Anamaria Fălăuş, Aritz Irurtzun, and Bryan Leferman, editors, *Proceedings of Sinn und Bedeutung 18*, pages 217–234, Bayonne and Vitoria-Gasteiz.

Kracht, Marcus (1994). Logic and control: How they determine the behaviour of presuppositions. In van Eijck, Jan and Albert Visser, editors, *Logic and Information Flow*, pages 89–111. MIT press, Cambridge, MA.

Krahmer, Emiel (1998). *Presupposition and anaphora*. CSLI Publications, Stanford, CA.

Kratzer, Angelika (1999). Beyond 'Ouch' and 'Oops'. How descriptive and expressive meaning interact. Cornell Conference on Theories of Context Dependency. Comments on a paper by Kaplan.

Kripke, Saul A. (2009). Presupposition and anaphora: Remarks on the formulation of the projection problem. *Linguistic Inquiry*, 40(3):367–386.

Langendoen, D. Terence and Harris Savin (1971). The projection problem for presuppositions. In Fillmore, C. and D. Terence Langendoen, editors, *Studies in Linguistic Semantics*, pages 373–388. Holt, Reinhardt and Winston, New York.

Lee, Heeyoung, Angel Chang, Yves Peirsman, Nathanael Chambers, Mihai Surdeanu, and Dan Jurafsky (2013). Deterministic coreference resolution based on entity-centric, precision-ranked rules. *Computational Linguistics*, 39(4):885–916.

Levinson, Stephen C. (1983). *Pragmatics*. Cambridge textbooks in linguistics. Cambridge University Press.

Lewis, David (1979). Scorekeeping in a language game. *Journal of philosophical logic*, 8(1):339–359.

Maier, Emar (2009a). Presupposing acquaintance: a unified semantics for de dicto, de re and de se belief reports. *Linguistics and Philosophy*, 32(5):429–474.

Maier, Emar (2009b). Proper names and indexicals trigger rigid presuppositions. *Journal of Semantics*, 26(3):253–315.

Marcus, Mitchell P., Beatrice Santorini, and Mary Ann Marcinkiewicz (1993). Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics*, 19(2).

McCawley, James D. (1998). *The syntactic phenomena of English*. University of Chicago Press.

McFadden, Daniel (1977). *Quantitative methods for analyzing travel behavior of individuals: some recent developments*. Institute of Transportation Studies, University of California.

Michel, Jean-Baptiste, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K Gray, Joseph P Pickett, Dale Hoiberg, Dan Clancy, Peter Norvig, Jon Orwant, et al. (2011). Quantitative analysis of culture using millions of digitized books. *Science*, 331(6014):176–182.

Miltsakaki, Eleni, Rashmi Prasad, Aravind Joshi, and Bonnie Webber (2004). The penn discourse treebank. In *Proceedings of the 4th International Conference on Language Resources and Evaluation*.

Minnen, Guido, John Carroll, and Darren Pearce (2001). Applied morphological processing of English. *Journal of Natural Language Engineering*, 7(3):207–223.

Mitkov, Ruslan (2002). *Anaphora Resolution*. Longman.

Muskens, Reinhard (1996). Combining Montague semantics and discourse representation. *Linguistics and Philosophy*, 19(2):143–186.

Nissim, Malvina (2006). Learning information status of discourse entities. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, pages 94–102. Association for Computational Linguistics.

Nissim, Malvina, Shipra Dingare, Jean Carletta, and Mark Steedman (2004). An annotation scheme for information status in dialogue. In *Proceedings of the 4th International Conference on Language Resources and Evaluation*, pages 1023–1026. Citeseer.

Nouwen, Rick (2007). On appositives and dynamic binding. *Research on Language & Computation*, 5:87–102.

Nouwen, Rick (2014). A note on the projection of appositives. In McCready, Eric, Katsuhiko Yabushita, and Kei Yoshimoto, editors, *Formal approaches to semantics and pragmatics: Japanese and beyond*, volume 95 of *Studies in Linguistics and Philosophy*. Springer.

Orasan, Constantin and Richard Evans (2007). NP animacy identification for anaphora resolution. *Journal of Artificial Intelligence Research*, 29:79–103.

Palmer, Martha, Daniel Gildea, and Paul Kingsbury (2005). The Proposition Bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31(1):71–106.

Piwek, Paul and Emiel Krahmer (2000). Presuppositions in context: Constructing bridges. In *Formal aspects of context*, pages 85–106. Springer.

Poesio, Massimo and Renata Vieira (1998). A corpus-based investigation of definite description use. *Computational Linguistics*, 24(2):183–216.

Postolache, Oana, Ivana Kruijff-Korbayová, and Geert-Jan M Kruijff (2005). Data-driven approaches for information structure identification. In *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pages 9–16. Association for Computational Linguistics.

Potts, Christopher (2003). *The Logic of Conventional Implicatures*. PhD Thesis, University of California.

Potts, Christopher (2005). *The Logic of Conventional Implicatures*. Oxford University Press, USA.

Potts, Christopher (2013). Presupposition and implicature. In Lappin, Shalom and Chris Fox, editors, *The Handbook of Contemporary Semantic Theory*, volume 2. Wiley-Blackwell, Oxford. To appear.

Prasad, Rashmi, Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Livio Robaldo, Aravind Joshi, and Bonnie Webber (2008). The Penn Discourse Treebank 2.0. In Calzolari, Nicoletta, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, and Daniel Tapias, editors, *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC)*, pages 2961–2968. European Language Resources Association (ELRA).

Prince, Ellen F. (1981). Toward a taxonomy of given-new information. *Radical pragmatics*.

Pulman, Stephen G. (2007). Formal and computational semantics: a case study. In *Proceedings of the Seventh International Workshop on Computational Semantics: IWCS-7, Tilburg, The Netherlands*, pages 181–196.

Quine, Willard Van Orman (1956). Quantifiers and propositional attitudes. *Journal of Philosophy*, 53(5):177–187.

Quine, Willard Van Orman (1960). *Word and Object*. The MIT press, Cambridge, MA.

R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Reyle, Uwe (1993). Dealing with ambiguities by underspecification. *Journal of Semantics*, 10(2):123–179.

Reyle, Uwe (1995). On reasoning with ambiguities. In Abney, Steven P. and Erhard W. Hinrichs, editors, *Proceedings of the seventh European chapter of the Association for Computational Linguistics*, pages 1–8, Belfield, Dublin, Ireland. University College Dublin, The Association for Computational Linguistics.

Roberts, Craige (1989). Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy*, 12(6):683–721.

Roberts, Craige (1996). Information structure in discourse: Towards an integrated formal theory of pragmatics. In Yoon, Jae Hak and Andreas Kathol, editors, *Working Papers in Linguistics*, volume 49, pages 91–136. The Ohio State University, Department of Linguistics.

Rosenbach, Anette (2014). English genitive variation–the state of the art. *English Language and Linguistics*, 18(02):215–262.

Rosset, Sophie, Cyril Grouin, Karën Fort, Olivier Galibert, Juliette Kahn, and Pierre Zweigenbaum (2012). Structured named entities in two distinct press corpora: Contemporary broadcast news and old newspapers. In *Proceedings of the 6th Linguistic Annotation Workshop*, pages 40–48. Association for Computational Linguistics.

Russell, Bertrand (1905). On Denoting. *Mind*, 14(4):479–493.

Russell, Bertrand (1957). Mr. Strawson on referring. *Mind*, 66(263):385–389.

van der Sandt, Rob A. (1989). Presupposition and discourse structure. In Bartsch, Renate, Johan van Benthem, and Peter van Emde Boas, editors, *Semantics and Contextual Expression*, volume 11 of *Groningen-Amsterdam Studies in Semantics*, pages 278–294. Foris, Dordrecht.

van der Sandt, Rob A. (1992). Presupposition projection as anaphora resolution. *Journal of Semantics*, 9:333–377.

Schlenker, Philippe (2007). Anti-dynamics: Presupposition projection without dynamic semantics. *Journal of Logic, Language and Information*, 16(3):325–356.

Schlenker, Philippe (2008). Be articulate: A pragmatic theory of presupposition projection. *Theoretical Linguistics*, 34(3):157–212.

Schlenker, Philippe (2009). Local contexts. *Semantics and Pragmatics*, 2(3):1–78.

Schlenker, Philippe (2011). DRT with local contexts. *Natural language semantics*, 19(4):373–392. to appear.

Schlenker, Philippe (2013). Supplements without bidimensionalism. Ms.

Schwarz, Florian, editor (2015). *Experimental Perspectives on Presuppositions*. Studies in Theoretical Psycholinguistics. Springer International Publishing.

Sekine, Satoshi, Kiyoshi Sudo, and Chikashi Nobata (2002). Extended named entity hierarchy. In *Proceedings of LREC*, volume 2.

Simons, Mandy, Judith Tonhauser, David Beaver, and Craige Roberts (2010). What projects and why. In Li, Nan and David Lutz, editors, *Proceedings of SALT*, volume 20, pages 309–327. CLC Publications.

Singh, Raj (2008). *Modularity and locality in interpretation*. PhD thesis, MIT.

Soames, Scott (1979). A projection problem for speaker presuppositions. *Linguistic Inquiry*, 10(4):623–666.

Soames, Scott (1982). How presuppositions are inherited: A solution to the projection problem. *Linguistic Inquiry*, 13(3):483–545.

Spenader, Jennifer (2002). Presupposed propositions in a corpus of dialogue. In van Deemter, Kees and Rodger Kibble, editors, *Information Sharing: Reference and Presupposition in Language Generation and Interpretation*, pages 35–60. CSLI Publications, Stanford.

Stalnaker, Robert (1970). Pragmatics. *Synthese*, 22.

Stalnaker, Robert (1973). Presuppositions. *Journal of philosophical logic*, 2(4):447–457.

Stalnaker, Robert (1974). Pragmatic presuppositions. *Semantics and Philosophy*, pages 197–213.

Steedman, Mark (2001). *The syntactic process*. The MIT press.

Strawson, Peter Frederick (1950). On referring. *Mind*, 59(235):320–344.

Strawson, Peter Frederick (1964). Identifying reference and truth-values. *Theoria*, 30(2):96–118.

Tonhauser, Judith, David Beaver, Craige Roberts, and Mandy Simons (2013). Toward a taxonomy of projective content. *Language*, 89(1):66–109.

Veltman, Frank (1991). Defaults in update semantics. In Kamp, Hans, editor, *Conditionals, defaults and belief revision*, pages 28–64. Dyana Deliverable R2.5A, Edinburgh.

Venhuizen, Noortje and Harm Brouwer (2014). PDRT-SANDBOX: An implementation of Projective Discourse Representation Theory. In Rieser, Verena and Philippe Muller, editors, *Proceedings of the 18th Workshop on the Semantics and Pragmatics of Dialogue (DialWatt - SemDial 2014)*, pages 249–251, Edinburgh.

Venhuizen, Noortje J., Valerio Basile, Kilian Evang, and Johan Bos (2013a). Gamification for word sense labeling. In *Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013) – Short Papers*, pages 397–403, Potsdam, Germany. Association for Computational Linguistics.

Venhuizen, Noortje J., Johan Bos, and Harm Brouwer (2013b). Parsimonious semantic representations with projection pointers. In Erk, Katrin and Alexander Koller, editors, *Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013) – Long Papers*, pages 252–263, Potsdam, Germany. Association for Computational Linguistics.

Venhuizen, Noortje J., Johan Bos, Petra Hendriks, and Harm Brouwer (2014a). Harnessing projection: implementing Projective Discourse Representation Theory. in prep.

Venhuizen, Noortje J., Johan Bos, Petra Hendriks, and Harm Brouwer (2014b). How and why conventional implicatures project. In Snider, Todd, Sarah D'Antonio, and Mia Wiegand, editors, *Semantics and Linguistic Theory (SALT)*, volume 24, pages 63–83, New York, USA. LSA and CLC Publications.

Venhuizen, Noortje J., Johan Bos, Petra Hendriks, and Harm Brouwer (2015). Aligning discourse semantics with surface structure. Ms.

Von Fintel, Kai and Lisa Matthewson (2008). Universals in semantics. *The Linguistic Review*, 25(1-2):139–201.

Zaenen, Annie, Jean Carletta, Gregory Garretson, Joan Bresnan, Andrew Koontz-Garboden, Tatiana Nikitina, M. Catherine O'Connor, and Tom Wasow (2004). Animacy encoding in english: why and how. In *Proceedings of the 2004 ACL Workshop on Discourse Annotation*, pages 118–125. Association for Computational Linguistics.

Zeevat, Henk W. (1992). Presupposition and accommodation in update semantics. *Journal of Semantics*, 9(4):379–412.

Zeevat, Henk W. (1996). A neoclassical analysis of belief sentences. In Dekker, Paul J. E. and Martin J. B. Stokhof, editors, *Proceedings of the 10th Amsterdam colloquium*, pages 723–742. ILLC, Universiteit van Amsterdam.

Zeevat, Henk W. (2002). Explaining presupposition triggers. In van Deemter, Kees and Rodger Kibble, editors, *Information Sharing: Reference and Presupposition in Language Generation and Interpretation*, pages 61–87. CSLI Publications, Stanford.

Zou, Lei, Kun Xu, Jeffrey Xu Yu, Lei Chen, Yanghua Xiao, and Dongyan Zhao (2014). Efficient processing of label-constraint reachability queries in large graphs. *Information Systems*, 40:47–66.

List of Publications

1. Venhuizen, Noortje J., Johan Bos, Petra Hendriks, and Harm Brouwer (2015). Aligning discourse semantics with surface structure. Ms

2. Bos, Johan, Valerio Basile, Kilian Evang, Noortje J. Venhuizen, and Johannes Bjerva (2015). The Groningen Meaning Bank. In Ide, Nancy and James Pustejovsky, editors, *Handbook of Linguistic Annotation – Part Two: Case studies*, Text, Speech and Language Technology. Springer. (forthcoming)

3. Venhuizen, Noortje J., Johan Bos, Petra Hendriks, and Harm Brouwer (2014b). How and why conventional implicatures project. In Snider, Todd, Sarah D'Antonio, and Mia Wiegand, editors, *Semantics and Linguistic Theory (SALT)*, volume 24, pages 63–83, New York, USA. LSA and CLC Publications

4. Venhuizen, Noortje and Harm Brouwer (2014). PDRT-SANDBOX: An implementation of Projective Discourse Representation Theory. In Rieser, Verena and Philippe Muller, editors, *Proceedings of the 18th Workshop on the Semantics and Pragmatics of Dialogue (DialWatt - SemDial 2014)*, pages 249–251, Edinburgh

5. Venhuizen, Noortje J., Johan Bos, and Harm Brouwer (2013b). Parsimonious semantic representations with projection pointers. In Erk, Katrin and Alexander Koller, editors, *Proceedings of the 10th International Conference on Computational Semantics*

(IWCS 2013) – Long Papers, pages 252–263, Potsdam, Germany. Association for Computational Linguistics

6. Venhuizen, Noortje J., Valerio Basile, Kilian Evang, and Johan Bos (2013a). Gamification for word sense labeling. In *Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013) – Short Papers*, pages 397–403, Potsdam, Germany. Association for Computational Linguistics

7. Basile, Valerio, Johan Bos, Kilian Evang, and Noortje J. Venhuizen (2012a). Developing a large semantically annotated corpus. In Calzolari, Nicoletta, Khalid Choukri, Thierry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk, and Stelios Piperidis, editors, *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, pages 3196–3200, Istanbul, Turkey. European Language Resources Association (ELRA)

8. Basile, Valerio, Johan Bos, Kilian Evang, and Noortje J. Venhuizen (2012c). UGroningen: Negation detection with Discourse Representation Structures. In Agirre, Eneko, Johan Bos, and Mona Diab, editors, *Proceedings of the First Joint Conference on Lexical and Computational Semantics (*SEM 2012)*, volume 1, pages 301–309. Association for Computational Linguistics

9. Basile, Valerio, Johan Bos, Kilian Evang, and Noortje J. Venhuizen (2012b). A platform for collaborative semantic annotation. In Daelemans, Walter, Mirella Lapata, and Lluís Màrquez, editors, *Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 92–96, Avignon, France. Association for Computational Linguistics

Nederlandse samenvatting

De betekenis van een zin kan worden opgedeeld in verschillende onderdelen. Wanneer we bijvoorbeeld de zin “Petra’s artikel heeft een prijs gewonnen” horen, leiden we hier niet enkel uit af dat er een artikel is dat een prijs heeft gewonnen, maar ook dat er een persoon bestaat die “Petra” heet, en dat zij dit artikel heeft geschreven. We kunnen dus zeggen dat deze ene zin uit ten minste drie verschillende bijdragen bestaat: er is een persoon die “Petra” heet, Petra heeft een artikel geschreven, en dit artikel heeft een prijs gewonnen. Deze bijdragen zijn niet onafhankelijk van elkaar. Sterker nog, het bestaan van Petra is een voorwaarde voor het gegeven dat zij een artikel heeft geschreven, en het bestaan van dit artikel is op zijn beurt weer een voorwaarde voor het winnen van een prijs. In de taalkunde noemt men dit soort voorwaarden *presupposities*, oftewel ‘vooronderstellingen’.

Presupposities hebben een aantal interessante taalkundige eigenschappen. De belangrijkste hiervan is dat ze zich anders gedragen dan reguliere bijdragen ten opzichte van verschillende talige kenmerken, zoals ontkenning en de vragende vorm. Als we bovenstaande zin ontkennen (“Petra’s artikel heeft geen prijs gewonnen”) of in vragende vorm zetten (“Heeft Petra’s artikel een prijs gewonnen?”), verandert de betekenis: het volgt nu duidelijk niet meer dat het artikel een prijs gewonnen heeft. Wat echter opvalt is dat de presupposities niet veranderen. Uit zowel de ontkenning als de vragende vorm volgt nog steeds dat er iemand bestaat die “Petra” heet en dat deze persoon een artikel heeft geschreven. Deze eigenschap van presupposities—dat ze niet beïnvloed worden door de aanwezigheid van ontkenning en vragende vorm—wordt ook wel *projectie* genoemd. Om deze term te begrijpen, zouden we een zin kunnen vergelijken met een pop in een schimmenspel: wanneer deze met een lichtstraal wordt beschenen, laten bepaalde onderdelen het licht door, terwijl andere het juist tegenhouden. In deze analogie zijn de presupposities van een zin de stralen die op de muur geprojecteerd worden; deze worden niet beïnvloed door kenmerken zoals de kleur van de pop, en maken enkel de contouren zichtbaar.

Lang heeft men aangenomen dat projectie een eigenschap is die uniek is voor presupposities. In de loop der tijd zijn er echter andere taalkundige fenomenen geïdentificeerd die net als presupposities kunnen projecteren, maar die op een andere manier bijdragen aan de betekenis van een zin. Hiervan zijn de zogenaamde *conventionele implicaturen* een goed voorbeeld. In de zin "Johan, een vogelaar, heeft een boek geschreven" hebben we wederom te maken met meerdere bijdragen: er is een persoon die "Johan" heet, Johan is een vogelaar, en hij heeft een boek geschreven. Net als in de vorige zin functioneert de eerste bijdrage als een presuppositie (het bestaan van Johan wordt aangenomen). De tweede bijdrage dient daarentegen niet als aanname, maar verschaft nieuwe achtergrondinformatie over de persoon Johan, namelijk dat hij een vogelaar is. Dit noemt men een 'conventionele implicatuur', omdat het een afleiding is die volgt uit de manier waarop de betrekkelijke bijzinconstructie conventioneel wordt gebruikt. Presupposities en conventionele implicaturen verschillen dus met betrekking tot de status van de overgebrachte informatie; waar de bijdrage van presupposities als *gegeven* wordt beschouwd, worden conventionele implicaturen gebruikt om *nieuwe* achtergrondinformatie te verschaffen. Ondanks dit onderscheid hebben presupposities en conventionele implicaturen een belangrijke eigenschap met elkaar gemeen: ze projecteren—probeer bovenstaande zin maar eens in vragende vorm te zetten of te ontkennen! Deze observatie stelt ons voor een aantal belangrijke vragen: Hoe verhouden presupposities en conventionele implicaturen zich tot elkaar? Wat zijn de taalkundige eigenschappen die bepalen of een bijdrage projecteert of niet? En hoe kunnen we de relatie tussen de verschillende bijdragen van een zin in een theoretisch model representeren?

Om een antwoord te vinden op deze vragen, beschrijf ik in dit proefschrift de eigenschappen van verschillende projecterende fenomenen, waarvan presupposities en conventionele implicaturen (maar bijvoorbeeld ook anaforen) een belangrijk deel vormen. Ik beargumenteer dat projectie een inherent onderdeel moet zijn van de manier waarop betekenis wordt afgeleid uit de verschillende onderdelen van een zin of tekst. Om dit proces beter te begrijpen, maak ik gebruik van de abstracte betekenisrepresentaties van een veelgebruikt taalkundig formalisme genaamd *Discourse Representation Theory* (DRT). In deze representaties wordt de interactie tussen de verschillende onderdelen van betekenis expliciet gemaakt. Ik laat zien dat de bestaande analyse van presupposities in DRT een aantal tekortkomingen heeft, met name met betrekking tot de generalisatie naar andere projecterende fenomenen, zoals conventionele implicaturen. Vervolgens stel ik een uitbreiding van het DRT-formalisme voor waarin projectie centraal staat: dit nieuwe formalisme noem ik *Projective Discourse Representation Theory* (PDRT). Het idee achter PDRT is dat speciale *projectie-variabelen* aangeven wat de status is van de informatie bijgedragen door de verschillende onderdelen van een zin; zo worden presupposities expliciet gerepresenteerd als een verwijzing naar gegeven informatie in de context, en conventionele implicaturen als toevoeging van nieuwe informatie aan de beschrijving van een specifieke persoon of entiteit. De formele (wiskundige) eigenschappen van het PDRT formalisme heb ik uitgewerkt in de computerimplementatie PDRT-SANDBOX, waarmee ik laat zien dat deze analyse betrouwbare betekenisrepresentaties oplevert.

In het laatste deel van mijn proefschrift gebruik de betekenisrepresentaties van PDRT om meer te leren over projectiegedrag. Dit doe ik door middel van een *computationele* analyse op basis van de data uit de *Groningen Meaning Bank* (GMB): een semantisch geannoteerd corpus dat is ontwikkeld als onderdeel van het grotere onderzoeksproject waarvan ik deel heb uitgemaakt. Dit corpus bestaat uit een grote digitale collectie teksten (meer dan tienduizend in totaal), waarvoor automatisch een betekenisrepresentatie in termen van PDRT is afgeleid. Deze automatisch verkregen representaties moeten zoveel mogelijk overeen komen met de taalkundige intuïties van mensen. Om deze intuïties te verkrijgen, hebben wij een verzameling online spelletjes ontwikkeld, genaamd *Wordrobe*. Elk van de spelletjes richt zich op een klein onderdeel van de taalkundige analyse, zoals bijvoorbeeld het bepalen van het type woord (zelfstandig naamwoord of werkwoord), en het bepalen van diegene of datgene waarnaar een persoonlijk voornaamwoord verwijst. Door het spelen van Wordrobe kan iedereen dus bijdragen aan betere taalkundige representaties in de Groningen Meaning Bank.

De betekenisrepresentaties kunnen vervolgens worden gebruikt om meer te leren over verschillende aspecten van taal, zoals bijvoorbeeld projectiegedrag. In mijn computationele analyse richt ik mij op referentiële expressies, zoals namen, zelfstandige naamwoorden en persoonlijke voornaamwoorden. Ik onderzoek welke taalkundige eigenschappen van invloed zijn op de bijdrage van een bepaalde referentiële expressie. Deze eigenschappen kunnen worden opgedeeld in drie categorieën: eigenschappen van de expressie zelf, de context waarin hij wordt gebruikt, en kenmerken van de formele PDRT-representatie. De verkregen resultaten ondersteunen de PDRT-analyse op twee belangrijke punten. Ten eerste wordt de hypothese bevestigd dat de bijdragen van verschillende projecterende fenomenen in termen van 'gegevenheid' in de context kunnen worden geformuleerd (denk aan de beschrijving van presupposities als *gegeven* informatie en conventionele implicaturen als *nieuwe* informatie). Ten tweede laten de resultaten zien dat de betekenisrepresentaties van PDRT informatie bevatten die van invloed is op de bijdrage van referentiële expressies, en dat de PDRT-representaties dus gebruikt kunnen worden voor het maken van betere taalkundige voorspellingen.

De bevindingen in dit proefschrift laten zien hoe een combinatie van theoretische en computationele analyses toegepast kan worden voor het ontwikkelen van taalkundige analyses. Door gebruik te maken van computationele toepassingen zoals de Groningen Meaning Bank en PDRT-SANDBOX in combinatie met statistische methoden, kunnen bestaande taalkundige analyses onder de loep worden genomen en verfijnd. Dit draagt niet alleen bij aan een beter theoretisch begrip van menselijke taal, maar helpt ook bij de verbetering van taalgeoriënteerde computationele toepassingen, zoals automatisch vertalen (bijvoorbeeld *Google Translate*), spraakherkenning (bijvoorbeeld *Apple*'s *Siri*), en vraag-antwoord systemen (bijvoorbeeld *IBM*'s *Watson*).

English summary

The meaning of a sentence can be divided into different parts. For instance, when hearing the sentence "Petra's article won a prize", we not only infer that some article won a prize, but also that there is a person named "Petra", and that she is the author of the article. In other words, this single sentence contains at least three different contributions: there is a person named "Petra", Petra wrote an article, and this article won a prize. These different contributions are not independent. On the contrary, the existence of Petra is a precondition for the fact that she wrote an article, and the existence of this article, in turn, is a precondition for it to win a prize. In linguistics, these kinds of preconditions are called *presuppositions*, since they are pre-supposed.

Presuppositions have several interesting linguistic properties. The most important of these is that they behave differently than regular contributions with respect to linguistic constructions, such as negation and interrogatives. This becomes clear if the aforementioned sentence is negated ("It's not the case that Petra's article won a prize") or turned into an interrogative sentence ("Did Petra's article win a prize?"): its meaning changes, that is, it clearly does not follow anymore that the article won a prize. Interestingly, however, the presuppositions of the sentence survive. Both the negated and interrogative versions of the sentence imply that there is a person named "Petra" and that she wrote an article. This property of presuppositions—to remain unaffected by negation or interrogation—is called *projection*. This term can be explained by comparing a sentence to a puppet in a shadow play: when light shines on the shadow puppet, some of its parts block the light beams, while others let the light shine through. In this analogy, the presuppositions of a sentence are the beams that are projected on the wall; remaining unaffected by properties such as the color of the puppet, they merely unveil its outlines.

It has long been assumed that presuppositions are the only phenomena that possess the property of projection. Recently, however, several linguistic phenomena have been identified that project just like presuppositions, but that make a different contribution to the meaning of a sentence. This is for example the case for a class of contributions called *conventional implicatures*. In the sentence “Johan, a bird watcher, wrote a book” we can again distinguish between different contributions: there is a person named “Johan”, Johan is a bird watcher, and he wrote a book. Just like in the previous example, the first contribution is a presupposition (the existence of Johan is assumed). The second contribution, on the other hand, does not function as an assumption, but provides novel background information about the person named “Johan”, namely that he is a bird watcher. This is called a ‘conventional implicature’, because the implication follows from the conventional way in which the appositional construction is used. Presuppositions and conventional implicatures thus differ with respect to the status of the information they convey; while the contribution of a presupposition is considered to be *given*, conventional implicatures are used to convey *new* information. Despite this distinction, presuppositions and conventional implicatures share an important linguistic property: they project—as becomes evident if we try to negate the sentence or turn it into the interrogative form. This observation raises some important questions: What is the relation between presuppositions and conventional implicatures? What are the linguistic properties that determine whether contributions project or not? And how can we represent the different contributions of a sentence in a theoretical model?

In this thesis, I aim to answer these questions by investigating the behaviour of different projection phenomena, which include presuppositions and conventional implicatures (but also, for instance, anaphora). I argue that projection should be an inherent part of the way in which the meaning of a sentence or text is deduced from its individual parts.

To obtain a better understanding of this process, I employ the abstract meaning representations from a widely used linguistic formalism called *Discourse Representation Theory* (DRT). In these meaning representations, the interaction between the different contributions that make up the meaning of a text (discourse) are made explicit. I show that the existing analyses of presuppositions in DRT have its limitations, in particular with respect to generalization to other projection phenomena, such as conventional implicatures. I then propose an extension of the DRT formalism, in which projection has a central role. This formalism is called *Projective Discourse Representation Theory* (PDRT). The idea behind PDRT is that special *projection variables* indicate the information status of the different parts of a sentence; presuppositions are explicitly represented as referring to given information in the context, and conventional implicatures as contributing novel information to the description of a specific person or entity. The formal (mathematical) properties of the PDRT formalism are worked out in a computer implementation called PDRT-SANDBOX, which shows that the analysis results in reliable and robust meaning representations.

In the last part of my thesis, I use the representations from PDRT to learn more about projection behaviour. For this, I use a *data-driven* analysis based on data from the *Groningen Meaning Bank* (GMB): a semantically annotated corpus that was developed as part of the larger research project of which I was part. This corpus consist of a large digital collection of texts (more than ten thousand in total) that are automatically annotated with a meaning representation based on PDRT. These automatically derived representations should resemble human linguistic intuitions as much as possible. In order to collect these intuitions, we developed a collection of online crowd-sourcing games, called *Wordrobe*. Each of these games focuses on a small part of the linguistic analysis, such as determining the word type (noun versus verb), or determining the referent of a personal pronoun. By playing Wordrobe, everyone can contribute to the development of better meaning representations in the Groningen Meaning Bank.

The resulting meaning representations can be used to learn more about different aspects of linguistic meaning, for example projection behaviour. In my computational analysis, I focus on referential expressions, such as proper names, common nouns, and pronouns. I investigate which linguistic characteristics determine the contribution of a given referential expression. These characteristics can be classified into three categories: properties of the referential expression itself, the context in which the expression is used, and characteristics of the formal meaning representation from PDRT. The obtained results support the PDRT analysis in two different respects. Firstly, they confirm the hypothesis that the contributions of different projection phenomena can be explained in terms of the notion of 'givenness' (remember the description of presuppositions as *given* information and conventional implicatures as *new* information). Secondly, this study shows that the meaning representations from PDRT contain information that critically affects the contribution of referential expressions. This means that the PDRT representations contribute to making better predictions about linguistic behaviour.

This thesis shows how a combination of theoretical and computational analyses can be used to obtain better linguistic analyses. By making use of computational resources such as the Groningen Meaning Bank and PDRT-SANDBOX, in combination with statistical methods, existing linguistic analyses can be tested and refined. This not only contributes to a better theoretical understanding of natural language, but also aids the improvement of linguistically oriented computer applications, such as machine translation (for example, *Google Translate*), speech recognition (for example, *Apple*'s *Siri*), and question answering systems (for example, *IBM*'s *Watson*).

Groningen Dissertations in Linguistics (GRODIL)

1. Henriëtte de Swart (1991). *Adverbs of Quantification: A Generalized Quantifier Approach.*
2. Eric Hoekstra (1991). *Licensing Conditions on Phrase Structure.*
3. Dicky Gilbers (1992). *Phonological Networks. A Theory of Segment Representation.*
4. Helen de Hoop (1992). *Case Configuration and Noun Phrase Interpretation.*
5. Gosse Bouma (1993). *Nonmonotonicity and Categorial Unification Grammar.*
6. Peter Blok (1993). *The Interpretation of Focus: an epistemic approach to pragmatics.*
7. Roelien Bastiaanse (1993). *Studies in Aphasia.*
8. Bert Bos (1993). *Rapid User Interface Development with the Script Language Gist.*
9. Wim Kosmeijer (1993). *Barriers and Licensing.*
10. Jan-Wouter Zwart (1993). *Dutch Syntax: A Minimalist Approach.*
11. Mark Kas (1993). *Essays on Boolean Functions and Negative Polarity.*
12. Ton van der Wouden (1994). *Negative Contexts.*
13. Joop Houtman (1994). *Coordination and Constituency: A Study in Categorial Grammar.*

14. Petra Hendriks (1995). *Comparatives and Categorial Grammar.*
15. Maarten de Wind (1995). *Inversion in French.*
16. Jelly Julia de Jong (1996). *The Case of Bound Pronouns in Peripheral Romance.*
17. Sjoukje van der Wal (1996). *Negative Polarity Items and Negation: Tandem Acquisition.*
18. Anastasia Giannakidou (1997). *The Landscape of Polarity Items.*
19. Karen Lattewitz (1997). *Adjacency in Dutch and German.*
20. Edith Kaan (1997). *Processing Subject-Object Ambiguities in Dutch.*
21. Henny Klein (1997). *Adverbs of Degree in Dutch.*
22. Leonie Bosveld-de Smet (1998). *On Mass and Plural Quantification: The Case of French 'des'/'du'-NPs.*
23. Rita Landeweerd (1998). *Discourse Semantics of Perspective and Temporal Structure.*
24. Mettina Veenstra (1998). *Formalizing the Minimalist Program.*
25. Roel Jonkers (1998). *Comprehension and Production of Verbs in Aphasic Speakers.*
26. Erik F. Tjong Kim Sang (1998). *Machine Learning of Phonotactics.*
27. Paulien Rijkhoek (1998). *On Degree Phrases and Result Clauses.*
28. Jan de Jong (1999). *Specific Language Impairment in Dutch: Inflectional Morphology and Argument Structure.*
29. Hae-Kyung Wee (1999). *Definite Focus.*
30. Eun-Hee Lee (2000). *Dynamic and Stative Information in Temporal Reasoning: Korean Tense and Aspect in Discourse.*
31. Ivilin Stoianov (2001). *Connectionist Lexical Processing.*
32. Klarien van der Linde (2001). *Sonority Substitutions.*
33. Monique Lamers (2001). *Sentence Processing: Using Syntactic, Semantic, and Thematic Information.*
34. Shalom Zuckerman (2001). *The Acquisition of "Optional" Movement.*
35. Rob Koeling (2001). *Dialogue-Based Disambiguation: Using Dialogue Status to Improve Speech Understanding.*
36. Esther Ruigendijk (2002). *Case Assignment in Agrammatism: a Cross-linguistic Study.*
37. Tony Mullen (2002). *An Investigation into Compositional Features and Feature Merging for Maximum Entropy-Based Parse Selection.*

38. Nanette Bienfait (2002). *Grammatica-onderwijs aan allochtone jongeren.*
39. Dirk-Bart den Ouden (2002). *Phonology in Aphasia: Syllables and Segments in Level-specific Deficits.*
40. Rienk Withaar (2002). *The Role of the Phonological Loop in Sentence Comprehension.*
41. Kim Sauter (2002). *Transfer and Access to Universal Grammar in Adult Second Language Acquisition.*
42. Laura Sabourin (2003). *Grammatical Gender and Second Language Processing: An ERP Study.*
43. Hein van Schie (2003). *Visual Semantics.*
44. Lilia Schürcks-Grozeva (2003). *Binding and Bulgarian.*
45. Stasinos Konstantopoulos (2003). *Using ILP to Learn Local Linguistic Structures.*
46. Wilbert Heeringa (2004). *Measuring Dialect Pronunciation Differences using Levenshtein Distance.*
47. Wouter Jansen (2004). *Laryngeal Contrast and Phonetic Voicing: A Laboratory Phonology Approach to English, Hungarian and Dutch.*
48. Judith Rispens (2004). *Syntactic and Phonological Processing in Developmental Dyslexia.*
49. Danielle Bougaïré (2004). *L'approche communicative des campagnes de sensibilisation en santé publique au Burkina Faso: les cas de la planification familiale, du sida et de l'excision.*
50. Tanja Gaustad (2004). *Linguistic Knowledge and Word Sense Disambiguation.*
51. Susanne Schoof (2004). *An HPSG Account of Nonfinite Verbal Complements in Latin.*
52. M. Begoña Villada Moirón (2005). *Data-driven identification of fixed expressions and their modifiability.*
53. Robbert Prins (2005). *Finite-State Pre-Processing for Natural Language Analysis.*
54. Leonoor van der Beek (2005). *Topics in Corpus-Based Dutch Syntax.*
55. Keiko Yoshioka (2005). *Linguistic and gestural introduction and tracking of referents in L1 and L2 discourse.*
56. Sible Andringa (2005). *Form-focused instruction and the development of second language proficiency.*

57. Joanneke Prenger (2005). *Taal telt! Een onderzoek naar de rol van taalvaardigheid en tekstbegrip in het realistisch wiskundeonderwijs.*
58. Neslihan Kansu-Yetkiner (2006). *Blood, Shame and Fear: Self-Presentation Strategies of Turkish Women's Talk about their Health and Sexuality.*
59. Mónika Z. Zempléni (2006). *Functional imaging of the hemispheric contribution to language processing.*
60. Maartje Schreuder (2006). *Prosodic Processes in Language and Music.*
61. Hidetoshi Shiraishi (2006). *Topics in Nivkh Phonology.*
62. Tamás Biró (2006). *Finding the Right Words: Implementing Optimality Theory with Simulated Annealing.*
63. Dieuwke de Goede (2006). *Verbs in Spoken Sentence Processing: Unraveling the Activation Pattern of the Matrix Verb.*
64. Eleonora Rossi (2007). *Clitic production in Italian agrammatism.*
65. Holger Hopp (2007). *Ultimate Attainment at the Interfaces in Second Language Acquisition: Grammar and Processing.*
66. Gerlof Bouma (2008). *Starting a Sentence in Dutch: A corpus study of subject- and object-fronting.*
67. Julia Klitsch (2008). *Open your eyes and listen carefully. Auditory and audiovisual speech perception and the McGurk effect in Dutch speakers with and without aphasia.*
68. Janneke ter Beek (2008). *Restructuring and Infinitival Complements in Dutch.*
69. Jori Mur (2008). *Off-line Answer Extraction for Question Answering.*
70. Lonneke van der Plas (2008). *Automatic Lexico-Semantic Acquisition for Question Answering.*
71. Arjen Versloot (2008). *Mechanisms of Language Change: Vowel reduction in 15th century West Frisian.*
72. Ismail Fahmi (2009). *Automatic term and Relation Extraction for Medical Question Answering System.*
73. Tuba Yarbay Duman (2009). *Turkish Agrammatic Aphasia: Word Order, Time Reference and Case.*
74. Maria Trofimova (2009). *Case Assignment by Prepositions in Russian Aphasia.*

75. Rasmus Steinkrauss (2009). *Frequency and Function in WH Question Acquisition. A Usage-Based Case Study of German L1 Acquisition.*
76. Marjolein Deunk (2009). *Discourse Practices in Preschool. Young Children's Participation in Everyday Classroom Activities.*
77. Sake Jager (2009). *Towards ICT-Integrated Language Learning: Developing an Implementation Framework in terms of Pedagogy, Technology and Environment.*
78. Francisco Dellatorre Borges (2010). *Parse Selection with Support Vector Machines.*
79. Geoffrey Andogah (2010). *Geographically Constrained Information Retrieval.*
80. Jacqueline van Kruiningen (2010). *Onderwijsontwerp als conversatie. Probleemoplossing in interprofessioneel overleg.*
81. Robert G. Shackleton (2010). *Quantitative Assessment of English-American Speech Relationships.*
82. Tim Van de Cruys (2010). *Mining for Meaning: The Extraction of Lexico-semantic Knowledge from Text.*
83. Therese Leinonen (2010). *An Acoustic Analysis of Vowel Pronunciation in Swedish Dialects.*
84. Erik-Jan Smits (2010). *Acquiring Quantification. How Children Use Semantics and Pragmatics to Constrain Meaning.*
85. Tal Caspi (2010). *A Dynamic Perspective on Second Language Development.*
86. Teodora Mehotcheva (2010). *After the fiesta is over. Foreign language attrition of Spanish in Dutch and German Erasmus Students.*
87. Xiaoyan Xu (2010). *English language attrition and retention in Chinese and Dutch university students.*
88. Jelena Prokić (2010). *Families and Resemblances.*
89. Radek Šimík (2011). *Modal existential wh-constructions.*
90. Katrien Colman (2011). *Behavioral and neuroimaging studies on language processing in Dutch speakers with Parkinson's disease.*
91. Siti Mina Tamah (2011). *A Study on Student Interaction in the Implementation of the Jigsaw Technique in Language Teaching.*

92. Aletta Kwant (2011). *Geraakt door prentenboeken. Effecten van het gebruik van prentenboeken op de sociaal-emotionele ontwikkeling van kleuters.*
93. Marlies Kluck (2011). *Sentence amalgamation.*
94. Anja Schüppert (2011). *Origin of asymmetry: Mutual intelligibility of spoken Danish and Swedish.*
95. Peter Nabende (2011). *Applying Dynamic Bayesian Networks in Transliteration Detection and Generation.*
96. Barbara Plank (2011). *Domain Adaptation for Parsing.*
97. Çağrı Çöltekin (2011). *Catching Words in a Stream of Speech: Computational simulations of segmenting transcribed child-directed speech.*
98. Dörte Hessler (2011). *Audiovisual Processing in Aphasic and Non-Brain-Damaged Listeners: The Whole is More than the Sum of its Parts.*
99. Herman Heringa (2012). *Appositional constructions.*
100. Diana Dimitrova (2012). *Neural Correlates of Prosody and Information Structure.*
101. Harwintha Anjarningsih (2012). *Time Reference in Standard Indonesian Agrammatic Aphasia.*
102. Myrte Gosen (2012). *Tracing learning in interaction. An analysis of shared reading of picture books at kindergarten.*
103. Martijn Wieling (2012). *A Quantitative Approach to Social and Geographical Dialect Variation.*
104. Gisi Cannizzaro (2012). *Early word order and animacy.*
105. Kostadin Cholakov (2012). *Lexical Acquisition for Computational Grammars. A Unified Model.*
106. Karin Beijering (2012). *Expressions of epistemic modality in Mainland Scandinavian. A study into the lexicalization-grammaticalization-pragmaticalization interface.*
107. Veerle Baaijen (2012). *The development of understanding through writing.*
108. Jacolien van Rij (2012). *Pronoun processing: Computational, behavioral, and psychophysiological studies in children and adults.*
109. Ankelien Schippers (2012). *Variation and change in Germanic long-distance dependencies.*

110. Hanneke Loerts (2012). *Uncommon gender: Eyes and brains, native and second language learners, & grammatical gender.*
111. Marjoleine Sloos (2013). *Frequency and phonological grammar: An integrated approach. Evidence from German, Indonesian, and Japanese.*
112. Aysa Arylova (2013). *Possession in the Russian clause. Towards dynamicity in syntax.*
113. Daniël de Kok (2013). *Reversible Stochastic Attribute-Value Grammars.*
114. Gideon Kotzé (2013). *Complementary approaches to tree alignment: Combining statistical and rule-based methods.*
115. Fridah Katushemererwe (2013). *Computational Morphology and Bantu Language Learning: an Implementation for Runyakitara.*
116. Ryan C. Taylor (2013). *Tracking Referents: Markedness, World Knowledge and Pronoun Resolution.*
117. Hana Smiskova-Gustafsson (2013). *Chunks in L2 Development: A Usage-Based Perspective.*
118. Milada Walková (2013). *The aspectual function of particles in phrasal verbs.*
119. Tom O. Abuom (2013). *Verb and Word Order Deficits in Swahili-English bilingual agrammatic speakers.*
120. Gülsen Yılmaz (2013). *Bilingual Language Development among the First Generation Turkish Immigrants in the Netherlands.*
121. Trevor Benjamin (2013). *Signaling Trouble: On the linguistic design of other-initiation of repair in English Conversation.*
122. Nguyen Hong Thi Phuong (2013). *A Dynamic Usage-based Approach to Second Language Teaching.*
123. Harm Brouwer (2014). *The Electrophysiology of Language Comprehension: A Neurocomputational Model.*
124. Kendall Decker (2014). *Orthography Development for Creole Languages.*
125. Laura S. Bos (2015). *The Brain, Verbs, and the Past: Neurolinguistic Studies on Time Reference.*
126. Rimke Groenewold (2015). *Direct and indirect speech in aphasia: Studies of spoken discourse production and comprehension.*

127. Huiping Chan (2015). *A Dynamic Approach to the Development of Lexicon and Syntax in a Second Language.*
128. James Griffiths (2015). *On appositives.*
129. Pavel Rudnev (2015). *Dependency and discourse-configurationality: A study of Avar.*
130. Kirsten Kolstrup (2015). *Opportunities to speak. A qualitative study of a second language in use.*
131. Güliz Güneş (2015). *Deriving Prosodic structures.*
132. Cornelia Lahmann (2015). *Beyond barriers. Complexity, accuracy, and fluency in long-term L2 speakers' speech.*
133. Sri Wachyunni (2015). *Scaffolding and Cooperative Learning: Effects on Reading Comprehension and Vocabulary Knowledge in English as a Foreign Language.*
134. Albert Walsweer (2015). *Ruimte voor leren. Een etnogafisch onderzoek naar het verloop van een interventie gericht op versterking van het taalgebruik in een knowledge building environment op kleine Friese basisscholen.*
135. Aleyda Lizeth Linares Calix (2015). *Raising Metacognitive Genre Awareness in L2 Academic Readers and Writers.*
136. Fathima Mufeeda Irshad (2015). *Second Language Development through the Lens of a Dynamic Usage-Based Approach.*
137. Oscar Strik (2015). *Modelling analogical change. A history of Swedish and Frisian verb inflection.*
138. He Sun (2015). *Predictors and stages of very young child EFL learners' English development in China.*
139. Marieke Haan (2015). *Mode Matters. Effects of survey modes on participation and answering behavior.*
140. Nienke Houtzagers (2015). *Bilingual advantages in middle-aged and elderly populations.*
141. Noortje Joost Venhuizen (2015). *Projection in Discourse: A data-driven formal semantic analysis.*

GRODIL
Center for Language and Cognition Groningen (CLCG)
P.O. Box 716
9700 AS Groningen
The Netherlands