# A Coherence-Based Approach to the Interpretation of Non-Sentential Utterances in Dialogue

*David Schlangen*

Doctor of Philosophy

Institute for Communicating and Collaborative Systems
School of Informatics
University of Edinburgh

2003

# Abstract

This thesis is concerned with the syntax, compositional semantics and contextually-situated interpretation of a certain kind of non-sentential utterance occuring in dialogue, namely one where the utterance, despite its 'incomplete' syntactic form, is intended to convey a proposition, a question or a request. Perhaps the most prominent type of such utterances is the *short answer*, as in "A: Who came to the party? — B: Peter.", but there are many other types as well. Following (Morgan 1973) and (Ginzburg 1999*b*) and others, we will call such utterances *fragments*. Clearly, the interpretation of fragments is highly context dependent. We will provide evidence that there are complex syntactic, semantic and pragmatic constraints governing the use of fragments. In particular, following (Ginzburg 1999*b*), we will present evidence that while the main resolution must be semantic, some limited syntactic information nevertheless has to persist beyond the boundaries of sentences to allow for the formulation of certain constraints on fragments. We will argue that consequently only a theory that has at its disposal a wide array of information sources —from syntax through compositional and lexical semantics to domain and world knowledge, and reasoning about cognitive states— can do justice to the complexity of their interpretation. As we will show, however, it is desirable to encapsulate these knowledge sources as much as possible, in order to maintain computability. Our main thesis then is that the resolution of the intended content of fragments can be modelled as a by-product of the establishment of *coherence* in dialogue, which (following much of the work on discourse) we define as the establishment of certain connections of the content of the current utterance to the content of its discourse context. We will show that all constraints on the form and content of fragments follow from how they are connected to the context.

The central role of discourse coherence in our account of fragments, together with having access to different kinds of information, distinguishes our theory from prior attempts. The work of Jonathan Ginzburg and colleagues ((Ginzburg 1999*b*, Ginzburg & Sag 2001) *inter alia*), for example, provides an approach to some types of fragments which is based on unification-operations on HPSG-signs. This approach, as we will show, fails to offer a convincing model of the interpretation of fragments where missing content is linguistically implicit and has to be inferred. Carberry (1990), on the other hand, employs computationally expensive plan-recognition techniques for the interpretion of fragments. This fails to predict certain empirical facts and we will furthermore show that the complex reasoning with cognitive states that she employs can often be replaced with much simpler inferences based on linguistic information.

In this thesis, we offer an analysis of the syntax and compositional semantics of fragments, and we provide a computational and formally precise theory of how the compositional semantics is supplemented with further content via reasoning about the context—both linguistic and non-linguistic. We also describe an implementation of our approach, based on an extension of a wide-coverage grammar and an accompanying discourse reasoning component for a simple domain.

# Acknowledgements

First, and foremost, I'd like to thank my supervisor, Alex Lascarides. During the last three and a half years, she has been the best supervisor one could have: always accessible, always quick to respond even to half- or quarter-baked ideas, always encouraging. If the work presented here is more than a fragment that is in no small degree due to her.

While being less involved in the day-to-day affairs of the research, Claire Grover, my second supervisor, has provided invaluable input at crucial points, patiently reading ever more voluminous drafts.

I have greatly benefited from the discussion of the work presented here at my *viva voce* with my examiners, Dan Flickinger and Johan Bos.

I am also grateful to the administrative staff at the department, who helped me sort out the bureaucratic stuff, kept the computers running, and suffered me coming in almost daily asking whether that important parcel had finally arrived.

I had the privilege of being able to present the work at a number of conferences, and I benefited greatly first from the reviews of the papers and then from the feedback to the presentations. Especially helpful was a Post-Graduate workshop of the *Europäisches Graduierten Kolleg* (a collaboration between Saarbrücken and Edinburgh), where I profited from discussions with Alexander Koller and Ivana Kruijff-Korbayova.

I gratefully acknowledge the financial support of the Graduate School of Informatics of the University of Edinburgh, ICCS, the Edinburgh University Development Fund, and the organizers of SIGDIAL '03 that enabled me to travel to those conferences. The whole thesis project was supported by EPSRC through a grant to Alex Lascarides, which I also gratefully acknowledge.

I am thankful to Jonathan Ginzburg and his group at King's College in London (especially Matt Purver and Raquel Fernández), who, even though we might disagree on the Proper Treatment of Fragments, have nevertheless been very open to discussion, and always helpful in exchanging information. Thanks are also due to Ann Copestake for the early collaboration on RUDI, and some interesting discussions. I am also grateful to Malte Gabsdil, and to Nikiforos Karamanis, for discussions not only about fragments. Bernd Eickmann and Jason Baldridge read some chapters of the thesis and prevented a number of mistakes from making it into the final version. Lastly, I want to thank Stephan Oepen for help with `[incr tsdb()]`.

I will not list here the friends who helped me get through the long, dark Scottish winters and the long, rainy Scottish "summers"—they know who they are (and they probably have better things to do than read theses). Maximilian Negwer, however, deserves a special mention, for without his brilliant invention this work might not have been possible.

Finally, I am deeply thankful to my parents for their support, first through one of those interminable

German degrees, then through yet another degree.

Parts of the work for this thesis have been done sitting at a desk in Hamburg, and later at the same desk in Berlin. For providing the desk, and for providing a refuge when Scotland became too cold, dark, or rainy, and for just about everything else, I thank Sophie.

# Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*David Schlangen*)

# Contents

# Chapter 1

# Introduction

Q: “What is this thesis about?”
A: “Non-sentential utterances, like this one.”

In this introductory chapter we will expand on this answer: we state what the problem is—namely interpreting fragments like the answer above—and we sketch our solution to it. We will also briefly define the problem *ex negativo*, by describing related phenomena we will *not* deal with in this thesis.

## 1.1 What this Thesis is about

### 1.1.1 The Problem: Interpreting *Fragments*

Imagine you overhear the following dialogue between people sitting at a neighbouring table in a Japanese restaurant:[1]

(1)
a. Paul: Who wants some Wakame?
b. Mary: Peter.
c. Peter: What is it?
d. Some kind of vegetable?
e. Paul: Yes,
f. seaweed.
g. Mary: Try it. It's good for you.
h. Peter: Why?
i. Mary: Lots of vitamins.
j. Peter: How do you eat this stuff?
k. Paul: I think with a fork.
l. Mary: No,
m. with chopsticks.

Even though many of the contributions in this dialogue would not be classified as sentences by most grammars, people do not have any problems in understanding the 'sentential-type' messages (i.e., the propositions and questions) conveyed by these non-sentential utterances. For example, as a competent speaker of English, you would probably understand (1-b) as meaning something that can be paraphrased as 'Peter wants some Wakame'. That this NP is intended to convey this proposition has to be *inferred*; it is information that goes beyond what is derivable from the compositional semantics (i.e., the meaning of the words and the syntax) of the fragment "Peter". This inference must take into account *contextual information*, in particular the fact that this fragment was meant to provide an answer to the question (1-a).

The challenge to a theory of dialogue semantics is then to model this competence, and this is what we want to do in this thesis. Our goal is to provide a competence model of the interpretation of *fragments* (as we will call this kind of non-sentential utterance). Apart from being of theoretical interest, such a theory can also potentially improve human–computer dialogue systems, since it provides a theoretical basis to enable them to understand human input of this type. Typically, about 10% of utterances in natural dialogue are fragmental (see the numbers given in Chapter 2), and so handling this kind of input

[1] (1-j) and (1-k) are taken from (Morgan 1973).

is a pressing task. We will demonstrate the theory's utility by deriving a prototype implementation that can resolve fragments in dialogues about a simple domain.

The key idea of our approach is that the resolution of the intended meaning of fragments can be modelled as a by-product of establishing the coherence of the dialogue, which we define (in line with much present research on discourse) as the establishment of certain connections of the content of the current utterance to the content of its discourse context.[2] We will show that establishing coherence depends on and interacts with both linguistic and extra-linguistic information sources, which thus have to be combined in a principled way.

In the remainder of this section, we give an idea of what the components of such a competence model must be, i.e. what information about fragment and context is needed to interpret the utterance and detect infelicitous uses. This will provide a preview of the sort of data we will be concerned with in this thesis. We will then briefly sketch how our approach to interpreting fragments makes use of the various component sources of information that influence their meaning. In Section 1.3, we summarise this overview in a concise list of the aims of the thesis, and finally in Section 1.4 we give a roadmap through the remaining chapters.

### 1.1.2 Relevant Information Sources

Despite the fact that fragments are not grammatical *sentences*, there are nevertheless some grammatical constraints on fragments. For example, the non-sentential utterances in the following examples seem to be *ungrammatical*.

(2) a. Peter: What is this?
*Picture?
b. Paul: I saw a man. *Policeman.

This is a first indication that probably fragments must be syntactically at least *phrases* (a claim already made by (Morgan 1973)).[3] Further, these phrases can potentially be modified by adverbs, as shown in (3).

(3) Paul: Who wants some Wakame?
Peter: Not me. Maybe Mary?

[2] Such an approach has been successfully used to account for a variety of phenomena, among them for example lexical ambiguity (Hobbs, Stickel, Appelt & Martin 1993) and presupposition (Asher & Lascarides 1998*b*). Kehler's (2002) approach to VP-ellipsis and gapping also falls under this rubric.

[3] We will later see exceptions to this general rule.

In certain cases, combinations of phrases seem not to be well-formed, as (4-a) shows, whereas in others they are (4-b).

(4) a. Paul: Who kicked whom?
Mary: *Peter me.
b. Paul: Who relies on whom?
Mary: Peter on Sandy.

Now, fragments like those in (2) and (4-a) seem to be ill-formed regardless of the context they are in, and despite the fact that pragmatics would suggest an interpretation in the given contexts. Hence we can conclude that there are conventional constraints on the well-formedness of fragments, and not just semantic or pragmatic ones. This claim is further supported by the observation that there are cross-linguistic differences: the German translation of (4-a), shown in (5), *is* grammatical.

(5) Paul: Wer trat wen?
Maria: Peter mich.

Determining which fragments are well-formed or ill-formed in this sense is the task of the *grammar of fragments*. This grammar must be part of the overall grammar of the language, i.e., it must be specified in a way that is compatible with current theories of the grammar of full sentences (and syntactic constituents).

The grammar must also exploit syntax to produce a logical form (LF) that fully reflects the content these fragments have independent of their context. Clearly, this LF will be highly underspecified. It must, however, contain sufficient information that a principled theory of dynamic semantics and pragmatics can infer how the underspecified arguments are resolved to specific values.

To resolve the underspecification that is generated by the grammar, additional information has to be taken into account. One of these is *discourse structure*, and we will explore the interplay between this structure and the interpretation of fragments in detail in this thesis. Discourse structure consists of rhetorical relations between utterances. In our approach, the rhetorical connection between a fragment and its context is the main clue for how the underspecification in the fragment should be resolved. For example, the information that a fragment and a previous utterance stand in a question-answer-pair relation provides strong semantic constraints on the fragment, thereby resolving its underspecified semantics as generated by the grammar. The following example illustrates the importance of the rhetorical connection:

(6) a. Peter: What healthy substance does seaweed contain?
Mary: Lots of vitamins.
*(= Seaweed contains* lots of vitamins.*)*

b. Paul: Seaweed contains lots of minerals.
Mary: No, lots of vitamins.
*(*= No, *seaweed contains* lots of vitamins.*)*

c. Mary: Seaweed is good for you.
Lots of vitamins.
*(= Seaweed is good [. . . ] because it contains* lots of vitamins.*)*

d. Mary: Seaweed contains good stuff.
For example, lots of vitamins.
*(*= For example, *seaweed contains* lots of vitamins.*)*

Informally, the rhetorical relation connecting the two utterances is *Question-Answer-Pair* in (6-a), *Correction* in (6-b), *Explanation* in (6-c), and *Elaboration* in (6-d). Note that (6-c) differs from the other examples in (6) because here 'contains' isn't explicitly in the context; we will come back to this difference below.

Certain properties of the relations between segments of the context can account for the fact that in some cases intervening material need not restrict the choice of where to connect the fragment, while in others it does. In (7) for example, (7-d) can (and, for the dialogue to be coherent, indeed must) be resolved to being an answer to (7-a), even though there are other utterances more adjacent to the fragment. On the other hand, Paul's utterance in (8) seems to disallow a resolution of Sandy's fragmental question to something like "Was it a letter from Joe ?". Here, the intervening material apparently restricts the choice of antecedents.[4]

(7) a. Paul: Who wants some Wakame?
b. Peter: What is it?
c. Mary: Seagrass.
d. Peter: Euw. Not me, then.

(8) Paul: Peter gave Mary a letter. Then she gave him a present.
Sandy: From Joe?

[4]Note that the situation seems to be a bit more complex. A more informative fragment like "a letter from Joe?" seems to be able to force a connection to the penultimate utterance in (8), at the cost of seemingly ignoring the previous utterance. We will further discuss this dialogue dynamics below in Chapter 8.

As (Morgan 1973), and expanding on this (Ginzburg 1999*b*) observed, there is also in certain cases what seems to be a *syntactic* 'influence' from the utterance to which the fragment is a reply on the fragment itself. For example, the preposition in the fragment-answer (9-a), being a verb particle, is normally seen as being semantically empty and serving only a syntactic function. However, only this and no other functional preposition can be used in answers to this question, as (9-b) shows.[5] This suggests that an approach that only relies on semantic information would not be able to capture this constraint; and so it seems that indeed at least some syntactic information needs to persist beyond the boundaries of the grammar.

(9) Peter: Who can we rely on?

   a. Paul: On Mary.
   b. Paul: #At Mary.

In our approach, this 'syntactic influence' will be a consequence of a general parallelism preference in discourse, which will be required when connecting the fragment to the context with certain rhetorical relations. Not all rhetorical connections trigger this syntactic parallelism; for example answers to *who*-questions do, but answers to *why*-questions don't:

(10) Peter: Why is it good for me?
     Mary: Lots of vitamins.

The fragment in (10) illustrates how sometimes world knowledge is needed to compute the intended meaning. To interpret Mary's utterance as an answer to the question (and hence as an explanation of the presupposed proposition "it [seaweed] is good for Peter"), we have to use knowledge about how reference to "lots of vitamins" can explain why something is good for someone. This will presumably lead to an interpretation that can be paraphrased as "it *contains* lots of vitamins".

Information about the goals of speakers that are normally connected with certain speech acts and the plans they follow (i.e. reasoning about cognitive states of dialogue participants), can also help to resolve underspecification, as example (11) (after (Asher & Lascarides 1999)) demonstrates.

(11) a. Peter: Let's meet the weekend after next.
     b. Mary: OK, but not Saturday.
     c. Peter: #So, 2pm then?

[5] We use '#' in this thesis to mark utterances that are infelicitous in their context or pragmatically ill-formed, but not ungrammatical.

According to (Clark 1977), definite descriptions in discourses have to be “bridged” (i.e., connected semantically) to entities that have already been introduced in the context. This bridging relation and the related entitity are however typically not specified by the grammar and have to be inferred. For resolving the definite “Saturday” in (11-b), there are at least two accessible antecedents to which it can be bridged: “now” (which we assume to be always accessible to temporal expressions) and “the weekend after next”. The resolution of “Saturday” to its intended meaning in this context (Saturday of the weekend after next) is guided by the underlying intention of agreeing on a time to meet, and the assumption that the discourse participants are cooperative. So reasoning about the *intentions* behind the contributions of the dialogue participants can help the resolution.

The same example, however, shows that there are also cases where even if the underlying intentions strongly suggest a reading, this reading might for other reasons not be available. This is demonstrated by (11-c). It is odd, because discourse structure enforces a resolution of “2pm” to Saturday 2pm, which would then conflict with the plan of finding a time, since Saturday has already explicitly been ruled out. A reading of 2pm as referring to Sunday 2pm, which would be compatible with the plan, doesn’t seem to be available for linguistic reasons (accessibility of antecedents; cf. (Kamp & Reyle 1993)).

This section has shown that there is a wide range of information sources that must be respected when constructing a model of the interpretation of fragments. In the following subsection, we briefly sketch how our theory makes use of these sources to arrive at an interpretation of fragments in context.

### 1.1.3 Questions of Design

Although we have already at least implicitly taken a stance in the previous section, for the sake of generality we list here some questions regarding the design of a theory of fragment interpretation.

First, there is the question of how these information sources are best brought together. One radical position would be to represent *all* information, from syntactic to domain-knowledge, in one format and to do all reasoning within a logic working on these representations (e.g. (Hobbs et al. 1993)). Various more conservative modifications of this position are also conceivable, where modularisation is gradually increased. We will later review an approach which tends towards the radical non-modular position, which we will show has certain disadvantages. Our position is situated more towards the modular end of that scale, as we will see.

Given our aim for modularity, the next question is how the interfaces between the modules should be designed. For example, if we separate grammar and discourse interpretation, what then should the output be of the grammar? One can think of two alternatives here. First, one could let the discourse interpretation work on the ‘standard’ meanings of the phrases of which fragments consist. Or, one could assign fragments special kinds of representations that distinguish the meaning of phrases used as fragments from that of the same phrases used differently. (We will later argue for that second position.)

### 1.1.4 Sketch of the Proposed Solution

As the first component, our approach specifies a grammar of fragments, in the formalism of HPSG (Pollard & Sag 1994). Our grammar basically directly realizes the constraint mentioned above, namely that fragments are phrases, possibly modified by adverbs. In a phrase structure-like notation, the rule would be something like “S-frag → (ADV) XP”. To mention one technical detail, the grammar makes use of *constructions* (i.e. couplings of form and meaning particular to certain grammatical constructions; introduced to HPSG by (Sag 1997)), and so does not have to postulate phonologically empty elements.

The grammar builds underspecified logical forms which represent the compositional semantics of the fragments, i.e. their meaning independent from any context. To capture this semantics, we extend an extant representation formalism with a new constraint. Roughly, the underspecified representation for the fragment “Peter.” will express that it will resolve to a proposition in which the entity denoted by the NP will play some, to be specified, role.

The final component then is a theory of how these underspecified values are resolved during discourse processing. Here we use a theory of discourse semantics called SDRT (Asher 1993, Asher & Lascarides 2003) as our backbone. This theory implements the idea that discourses are coherent only if all bits of information in them are connected by rhetorical relations or, equivalently in this theory, by speech act types. The theory already offers a method for managing most of the information sources we have listed above. One important extension, though, is that we will need access to a limited amount of syntactic information as well, to handle examples like (9). We will explain the differences in whether syntactic constraints on fragments are present or not by extending a general constraint on discourse updates, namely one that, all else being equal, always prefers the resolution that results in the most parallel pair of representations.

There is a connecting theme in the three components, namely the use of descriptions or constraints to specify information. All the components mentioned above do in some form use the distinction between description and entity that is being described: the grammar is a collection of constraints specifying linguistic *signs*; the compositional semantics of fragments is given in the form of a description as well; the discourse interpretation module finally defines how constraints describing discourse structure are accumulated during the discourse. Hence, a technically appropriate name for our approach would be ‘constraint-based’; however, given the primary importance of the notion of coherence we will dub our approach ‘coherence-based’.

## 1.2 What this Thesis isn’t about

There is a number of phenomena related to fragments we will *not* discuss in detail in this thesis. First of all, we should stress that we see utterances exhibiting VP-ellipsis (as for example B’s utterance in (12) below) *not* as fragments. VP-ellipsis is defined by the replacement of a particular syntactic constituent (the VP) with an element that must be contextually interpreted, whereas, as we will show, fragments cannot simply be defined syntactically, and also the 'missing' information is never marked syntactically. Nevertheless, the phenomenon of VP-ellipsis is certainly related to that of fragments; this relation will be discussed in Chapter 3.

(12) A: Who wants to come with us?
B: I do.

We have been careful so far to define fragments as "non-sentential utterances that intuitively seem to be intended *as they are* to convey propositions, questions and requests". Looking at transcriptions of dialogues, one can find many utterances that do not consist of full sentences (McKelvie 1998, Heeman & Allen 1999). Many of those, however, will be 'accidental', either dysfluencies during speech production, or recognition or transcription errors. The difference here is one of speaker intention: fragments in our sense are utterances which, to use the time-honoured Gricean definition, are intended in that form to induce a belief in an audience, where it is intended that the utterance is recognised as so intended (Grice 1957). Utterances that are non-sentential because of production errors (repetitions, self-repairs, etc.) or even only reach the addressee in a non-sentential form (because of transmission or recognition errors) are directly excluded by this definition, since they don't meet the criterion that they are intended to have that form. This distinction is important for our purposes, because we crucially assume that all the necessary information to resolve the message is there in the content of the signal that is available to us. If we can't resolve the intended meaning on that basis, then we label the discourse as incoherent, rather than incomprehensible.[6,7]

Secondly, there are two kinds of non-sentential utterance that will not be discussed here even though they fit our definition of "fragments". The first is what has been called antecedent-free or 'situationally controlled' fragments (Klein 1993, Schwabe 1994). (13) gives an example of a fragment of this kind.

(13) a. Joe [walking up to the counter at a coffee shop]: One frappuccino, please.
b. Johann: Einen$_{ACC}$ Eiskaffee, bitte.

These discourse-initial fragments will, given the situational context, resolve to the request "bring me a (cup of) coffee". This resolution, however, must rely to an even greater degree on knowledge about situational scripts, expectations, etc. Moreover, it is not clear to what extent these are conventionalised

[6]This is not to mean that some of the techniques developed here cannot be useful for this extended problem of dealing with non-sentential input. There seems to be a certain overlaps between our approach and for example Milward's (2000) approach to robust processing, but we leave a deeper comparison to further work.

[7]Of course we idealise a bit here, in that we assume that it always is possible to recognise whether an utterance is a fragment or an 'accidentally' non-sentential utterance. An examination in Section 2.2.3 of examples from corpora will show that sometimes classifying an utterance as one or the other is difficult.

forms; (13-b) for example requires a certain case marking that suggest a certain linguistic form for a question. We believe that similar principles regarding coherence govern the use of fragments of this type; after all, establishing linguistic coherence is only one part of interpreting behaviour as rational. However, trying to formalise this would lead us too far away from linguistic matters, and so we restrict ourselves to fragments that are in some form replies to other utterances.[8]

The other kind of non-accidentally non-sentential utterance we will not deal with here are those fulfilling mostly ‘discourse management’ functions (Bunt 1994), like “bye” or “hello”. We will justify this decision in Section 2.2.2.

Third, an interesting question related to fragments is why people actually use them. Why use something that has to be resolved to be understood, rather than using full sentences? One could speculate that fragments belong to the class of constructions that (Halliday & Hasan 1976) call “cohesive forms” (which includes for example various types of ellipsis, pronouns, etc.), because they contribute to the cohesion of a discourse by establishing links to previous utterances that contain the information needed to resolve them. In a similar vein, (Traum 1999*a*, p.125) comments on the effect of some kinds of fragments on the process of grounding information: “a good example [for evidence that a previous utterance has been understood] is a short answer to a question, which would be incoherent if the question had not been asked and understood.” These are questions that need to be answered by any theory of language *production*, but we ignore them here, instead focussing on working out what was said or meant by an utterance of a fragment. We believe that linking fragments to their semantics is a pre-requisite to working out when and how to produce them anyway.

Lastly, there are some issues that also concern interpretation, or rather recognition of well-formed and non-well-formed fragments. The following example is adapted from (Carberry 1990, p.168), and it shows that after a certain amount of intervening material has been processed, fragments become harder to interpret. We will put aside questions of how this can be modelled, though, since like the dysfluencies mentioned above this also seems to be more of a performance issue.

[8] Cf. (Schwabe 1994) for an approach to such situationally controlled fragments.

(14) (a) A: Do you want to take CS320?
(b) B: Who is teaching it?
(c) A: Dr. Raff and Dr. Owen.
(d) B: When does Dr. Raff's class meet?
(e) A: At 7:00pm on Wednesday night.
(f) B: Where?
(g) A: At Wilcastle.
(h) B: Does Dr. Owen's class meet on campus?
(i) A: Yes.
(j) B: When?
(k) A: It meets Tuesday and Thursday at 8:00 am.
(l) B: ??Yes, with Dr. Raff. [as answer to (a)]

## 1.3 Aims of the Thesis

To summarise, here are the aims of the thesis:

1. To model the grammatical well-formedness constraints on fragments;
2. To analyse their compositional semantics;
3. To provide a computationally tractable and formally precise theory of how the compositional semantics is supplemented via reasoning about the (linguistic and non-linguistic) context;
4. To implement the approach by extending a wide-coverage grammar and providing an accompanying discourse reasoning component for a simple domain.

## 1.4 Thesis Outline

The thesis can be separated into three main parts. Chapter 2 constitutes the main empirical part. In it, we present a taxonomy of fragment types, which we validate with data from several dialogue corpora. Chapters 3, 4 and 5 review the relevant literature, first looking at a related phenomenon—verb-phrase-ellipsis—and then at the existing literature on fragments themselves. On the basis of this review, we will draw some general conclusions about the criteria that a model of fragments must meet. Our approach–which adopts these criteria—will be developed in Chapter 6, 7 and 8, describing (in order) the compositional semantics of fragments, their syntax, and the way these components can be used in a theory of discourse interpretation to compute their intended meanings. These elements are brought together in Chapter 9, where a computational implementation of a restricted portion of the theory is described.

# Chapter 2

# A Taxonomy of Fragments

In the previous chapter we said that one very important clue to how a fragment is resolved in its context is the rhetorical relation in which it stands to some element of that context (or equivalently: the speech act type it instantiates). To illustrate this, we gave an example of the same fragment in different contexts in which it receives different interpretations. We will use this example here to motivate a classification of fragments along two dimensions. The first dimension is the speech act that has been performed with the fragment (we will make a further classification here); the second dimension concerns the question of where the information needed to resolve the fragment comes from. These two dimensions will be further discussed in sections 2.2 and 2.3, respectively; the taxonomy of speech act types will also be validated with data from a small corpus study. Section 2.4 will then compare our taxonomy with others in the literature, showing that we can offer a more principled classification which also captures novel generalisations. We close with a summary of the observations made in this chapter, and with desiderata for a theory of the interpretation of fragments that can be derived from them, which will guide the discussion of the related approaches and the development of our own in the subsequent chapters.

## 2.1 Two Dimensions for Classification

We gave an example in the previous chapter of a fragment—a bare NP—in different contexts in which it received different interpretations. Below in (15) we repeat this example, with an additional context (15-c).

(15) a. Peter: What healthy substance does seaweed contain?
Mary: Lots of vitamins.
*(= Seaweed contains* lots of vitamins.*)*

b. Paul: Seaweed contains lots of minerals.
Mary: No, lots of vitamins.
*(=* No, *seaweed contains* lots of vitamins.*)*

c. Peter: What does seaweed contain?
Lots of vitamins?
*(= Does seaweed contain* lots of vitamins*?)*

d. Mary: Seaweed is good for you.
Lots of vitamins.
*(= Seaweed is good [. . . ] because it contains* lots of vitamins.*)*

e. Mary: Seaweed contains good stuff.
For example, lots of vitamins.
*(=* For example, *seaweed contains* lots of vitamins.*)*

Our thesis is, put simply, that the fragment receives different interpretations in these contexts because of the different *functions* it plays with respect to the utterance to which it is a reply. For example, in (15-a) the fragment NP is intended to 'replace' the *wh*-phrase in the resolved meaning of B's utterance, and in (15-b) the object (and not the subject), because these fragments are intended to provide an answer and a correction, respectively. This functional difference or equivalently, the difference in what speech act type the fragments instantiate, is the first criterion for classification we will use. In the following section, we will investigate what kinds of speech act types can be realised by non-sentential utterances. We will introduce in that section a sub-classification which can be illustrated with (15-c) above. The fragment in that exchange has some similarity to the fragmental answer in (15-a), apart from it being uttered with an interrogative intonation. Its intended meaning is that of a polar question, which has the property that all (positive) answers to it will be answers to the question to which the fragment itself was a reply. Hence we will call this kind of rhetorical relation (we will see others like this) the 'question' or $_q$-versions of the speech act they determine for their answers.

The speech act a fragment realises, however, is not the only criterion for classifying it. If we look closer at the examples in (15), we see that in some of them all the material that is needed to interpret the fragment comes from the utterance to which it is a reply, while in others additional material is needed.[1] This is illustrated graphically in (16) below.

(16) a. Peter: What [...] does seaweed contain?

Mary: *Seaweed contains* lots of vitamins.

b. Mary: Seaweed is good for you.

?

*it contains* lots of vitamins.

In the interpretation of the fragmental answer in (16-a) (which is example (15-a) from (15) above), the fragment becomes the complement of a verb/predicate which is the same as the one in the question, and whose other argument also comes from the question. No additional material apart from that is needed. In (16-b) ((15-d) from (15)), on the other hand, something like 'seaweed contains ...' must be *inferred* to turn the fragment into an explanation of the assertion made in the previous utterance. This inference presumably requires general world knowledge about properties that are good in food. For these reasons, we will call fragments of the first type *resolution-via-identity* (or *res-via-id*) fragments, and those of the second type *resolution-via-inference* (*res-via-inf*) fragments. In the next section we will show that this difference is not a property of the speech act performed with the fragment but rather an independent factor for classification. For example, we will see question-answer-pairs where the fragmental answer is *res-via-inf* (unlike in (16-a) above), and elaborations where the fragment is *res-via-inf* (unlike (16-b)).

With these introductory distinctions in place, we can begin our survey of the different speech acts that can be performed with non-sentential utterances.

## 2.2 A Taxonomy of Non-Sentential Speech Acts

In this section we present our taxonomy of the types of speech acts that can be performed with non-sentential utterances (NSUs). We begin with a description of how the taxonomy was compiled, and introduce a further meta-classification of types of NSUs. The actual taxonomy is given in Subsection 2.2.3; the semi-formal definitions of the semantics of the speech act types presented in this section

[1]We are being deliberately vague about what kind of material this is in this chapter. Indeed, the question as to what material—syntactic or semantic—is needed to resolve ellipsis in general, will be what connects our problem to the survey of approaches to VP ellipsis in Chapter 3, and will also be the question along which we organise the description of extant approaches to fragments in Chapters 4 and 5. We will give a brief preview of the argumentation below in section 2.3, but apart from that we won't make any commitments here.

will form the basis for the formalisation of their semantics given in later chapters. The section closes with a small study of the coverage of fragments occurring in a corpus achieved with this taxonomy, i.e. of the descriptive adequacy of the classification.

### 2.2.1 Methodology

All classification attempts have to start from a theoretical position that tells one what there is to classify in the first place. For our approach to fragments this is, as already mentioned, speech act theory; more specifically, speech act theory as formalised in SDRT (Asher & Lascarides 2003).[2] Here is not the place to even give a brief overview of the development of speech act theory since its inception in (Austin 1962),[3] and so we just list a few points that make the particular theory we chose a promising starting point for our enterprise. First of all, speech acts in SDRT are *relations*; as the discussion in the previous chapter and in particular in the previous section has made clear, the relation of a fragment to its context is an important clue to its intended meaning, and so any taxonomy we use must incorporate this relationality. Secondly, the speech act types in this theory are provided with a well-defined formal semantics, unlike in many other more traditional ones; this degree of formalisation allows one to adopt a strict criterion for when a new relation is needed, namely if the truth-conditional consequences it has cannot be captured otherwise. Thirdly, unlike some other theories of discourse cohesion, SDRT allows a plurality of types to be instantiated by one utterance; a property we will make use of when we formalise the resolution of fragments in Chapter 8. Lastly, the types are always *linguistically* motivated, and their interaction with many linguistic phenomena is studied in great detail, which allows us to use these analyses if these phenomena occur in conjunction with fragments. A general advantage of using an extant speech-act typology of course is that it makes sure that the types are independently motivated, and hence what we do is compatible with theories of the semantics and pragmatics of full sentences.

As a starting point, we took the taxonomy of speech acts as described by (Asher & Lascarides 2003) for full sentences and constructed and also identified in corpora examples of fragments that instantiate these types.[4] To extend the taxonomy, we then systematically read through portions of the corpora, looking for fragments that weren't covered by any of the types. To not miss potential types of fragments and also to not prejudice the analysis, we cast our net as wide as possible in that search by extending our definition of 'fragment' from the one of the previous chapter ("non-sentential utterances that are

[2]Calling SDRT a formalisation of speech act theory is a bit simplifying—the theory brings together ideas from various fields, among them speech act theory—, but for the purposes of this chapter we will treat it as such.

[3]For a very readable attempt at providing such an overview, see (Traum 1999*b*).

[4]In detail, the corpora we looked through were:

- The British National Corpus (BNC, (Aston & Burnard 1998)); examples from this corpus are marked with a turn identifier like this: [`BNC FMM 1234`], where the three-letter code identifies the file and the number the sentence in that file.
- The London-Lund Corpus of Spoken English, (Svartvik 1990). Turns are marked like this: `1.4.610 A:`.
- The HCRC MapTask corpus, (Anderson, Bader, Bard, Boyle, Doherty, Garrod, Isard, Kowtko, McAllister, Miller, Sotillo, Thompson & Weinert 1991). Turns are marked like this: `q1ec7.f.30:`.

intended to convey propositions, questions and requests") to one that was easier to operationalise in a search, namely "utterances that are non-accidentally non-sentential". This wider definition meant that in a first step we included in our classification utterances like "Bye" or "Sorry?", although we will in the remainder of the thesis not focus on these types of non-sentential utterances. In the next subsection we discuss this distinction between fragments in the narrower sense given by the old definition and fragments in the wider sense of the latter definition in the next subsection.

### 2.2.2 A Further Distinction: 'Message-Type' NSUs and 'Non-Message-Type' NSUs

Our informal definition of the phenomenon, as mentioned above, requires that fragments convey messages. This criterion turned out to be not easily applicable in some cases, however. Undoubtedly, there are types of NSUs where it is clear that a message has been conveyed, for example the short-answers we saw in the previous section. (We will call NSU-speech acts of this type *message-type*-NSUs.) There are also types of NSUs where it is clear that *no* message was conveyed; for example it is probably not necessary to assign "Bye" a propositional content with truth conditions.[5] (We will call this type consequently *non-message-type* NSUs.) However, there are also types where it is less clear how they should be classified. Examples of more problematic classes of fragments are what we will call *Comment* and *Comment*$_q$ (cf. Sections 2.2.3.20 below), e.g. as in "cool!" and "really?". Do such non-sentential utterances resolve to messages like "<*the content of the previous utterance*> is cool" or "is <*the content of the previous utterance*> really true?", or are they better described as conventional acts that keep the dialogue running and do not convey any other message? We classify them as the latter, but without arguing much for this decision; in any case this indicates that this meta-classification is probably better understood as a continuum from acts like short-answers on the one side to greetings and 'backchannels' (utterances like "mmhm", see Section 2.2.3.20) on the other.

We should stress that we do not put much theoretical weight on this distinction—unlike the distinction between *resolution-via-identity* and *resolution-via-inference* introduced above and discussed below in Section 2.3— it's mainly intended to motivate why we single out certain kinds of NSUs for further analysis,[6] namely because only those types classified as *message-type* are context-dependent in an interesting sense. Only in this type of NSUs does the compositional semantics of the fragment make an interesting contribution to the resolved meaning, and this is what we want to study in this thesis.

We have to avert a possible misunderstanding concerning these meta-classes. Many theories of speech acts (although not SDRT) make a distinction between "task-oriented acts" and "dialogue control acts" (Bunt 1994) or "core speech acts" and "synchronisation acts" (Traum & Hinkelman 1992). Our distinc-

[5]One could of course argue that such utterances are not really non-sentential, and hence outside our domain here. We will not take such a step and consider them as non-sentences. (Although, as will become clear shortly, not as fragments—they do not have a plausible full-sentence counterpart.)

[6]And, as we will see in Section 2.4, including this kind of NSU will make our corpus results comparable with a previous study.

tion might seem similar at first blush, and to a large extent is indeed congruent with this distinction—for instance, the examples we gave above for *non-message-type* NSUs are also paradigm 'dialogue control acts'—but our criterion for making the distinction is a different one, and the classification is not fully congruent. For example, some acts which in the classification of (Bunt 1994) would be classified as 'dialogue control acts', namely clarifications as in (17-a), have a core that must be resolved contextually and hence are 'message-type NSUs' in our taxonomy; on the other hand there are acts that we classify as 'non-message type', like the *Acknowledgement* in (17-b), which are clearly task-related.

(17) a. A: I talked to Peter.
B: Peter Miller?
b. A: Let's meet next week.
B: OK.

We can now present the taxonomy.

### 2.2.3 The Classes

#### 2.2.3.1 Overview

The speech act types in this section are presented grouped into "families" according to certain shared properties. Before we delve into the details, a short overview of what can be expected is presented in Tables 2.1–2.3. In these tables the names of the speech act type, the informal definition, a short example, and finally a reference to the appropriate section is given. In the semi-formal definitions the utterance (or discourse segment) to which the fragment is being connected is called $\alpha$ and the fragment itself $\beta$. The subsequent utterance to the fragment $\beta$ is called $\gamma$. (As mentioned above, we deliberately do not specify here what labels like $\alpha$ actually label, be that syntactic or semantic representations of the utterances.) The speaker of an utterance is denoted by $Agent(\pi)$ (where $\pi$ is the label of the utterance). The main eventuality of a clause is denoted by $e$ with the label as index, e.g. $e_\alpha$ for that of $\alpha$.

The structure of the subsequent subsections is as follows: after giving the definition of the speech act, we show a few examples that illustrate the range of utterance pairs that fall under the respective definition. We first use made-up examples to make the points more clearly, but we have also strived to find attested examples in corpora, to show that these types do occur in naturally occurring dialogues. For some of the less frequent and more intricate speech acts, perhaps unsurprisingly, we didn't find attested examples, and so we took special care to get judgements of their 'naturalness' from native speakers.

We should point out here that there is one systematic omission in the taxonomy: we have focussed on relations with propositions and questions as arguments, at the price of neglecting requests. In principle, the taxonomy is easily extendable to include those as well, due to time and space constraints, however,

| *"Family"* / *Relation* | *Definition, Example* | *Section* |
|---|---|---|
| Question-Answer-Pair | | |
| *QAP* | β provides a direct answer to α. *"A: Who came to the party? — B: Peter."* | 2.2.3.2 |
| $QAP_q$ | Positive answers to y/n-question β provide a direct answer to α, negative answers a partial answer. *"A: Who was this? Peter?"* | 2.2.3.3 |
| Elaboration | | |
| $Elab_{pp}$ | β elaborates on some aspect of the indicative α. *"A: I talked to Peter. Peter Miller."* *"A: I talked to Peter. Yesterday."* | 2.2.3.4 |
| $Elab_{qp}$ | β elaborates on the intended meaning of α. *"A: Who did you talk to? Yesterday."* *( = "Who did you talk to yesterday?")* | 2.2.3.4 |
| $Elab_{pq}$ | Any answer to β elaborates on some aspect of the indicative α. *"A: I talked to Peter. — B: When?"* *"A: I talked to Peter. — B: Peter Miller?"* | 2.2.3.5 |
| $Elab_{qq}$ | Any answer to β elaborates on intended meaning of α. *"A: Did you talk to Peter? — B: Peter Miller?"* | 2.2.3.5 |
| Contrast | | |
| *Contr* | α and β have a *contrasting theme*. *"(A: Are they in the cupboard?) — B: No, in the fridge."* | 2.2.3.7 |

*(continued on next page)*

Table 2.1: Speech act types that can be realized with NSUs: Question-Answer-Pair, Elaboration, and Contrast

| *"Family"* *Relation* | *Definition, Example* | *Section* |
|---|---|---|
| Continuation, Alternation | | |
| *Cont* | β continues a topic of α. | 2.2.3.8 |
| | *"A: I am free on Monday. And on Wednesday."* | |
| *Q-Cont* | The question β continues a topic of the question α. | 2.2.3.8 |
| | *"A: What's his name? — B: ... — A: His address?"* | |
| *Q-Alt* | Answers to β answer an alternative question combined out of α and the fragment-phrase β′ α. | 2.2.3.9 |
| | *"A: Can you come on Tuesday? Or Wednesday?"* *(= "When can you come, Tuesday or Wednesday?")* | |
| Explanation, Result | | |
| *Expl* | β explains $e_\alpha$. | 2.2.3.10 |
| | *"A: Peter left early. Exams."* | |
| $Expl_q$ | All answers to β explain $e_\alpha$. | 2.2.3.11 |
| | *"A: Peter left early. — B: Exams?"* | |
| $Expl_q^*$ | All answers to β explain why α has been uttered. | 2.2.3.12 |
| | *"A: Are you married? — B: Why?"* | |
| *Res* | α explains $e_\beta$. | 2.2.3.13 |
| | *"A: He had a stroke. And died."* | |
| $Res_q$ | Answers to β are explained by α. | 2.2.3.13 |
| | *"A: He had a stroke. — B: And died?"* | |

*(continued on next page)*

Table 2.2: Speech act types that can be realized with NSUs: Continuation and Explanation

| *"Family"* *Relation* | *Definition, Example* | *Section* |
|---|---|---|
| Plan-Related Relations | | |
| *Plan-Elab* | β details a step in a plan to reach a goal behind α. *"A: Let's meet on Monday. At two o'clock."* | 2.2.3.14 |
| *Q-Elab* | Answers to β detail a step in a plan to reach a goal behind α. *"A: Let's meet on Monday. — B: At two o'clock?"* | 2.2.3.15 |
| *Ack* | β entails that *Agent*(β) has accepted or achieved *Agent*(α)'s goal behind uttering α. *"A: Let's meet on Monday. — B: OK."* | 2.2.3.17 |
| *Plan-Corr* | β indicates that *Agent*(β) doesn't accept or is unable to help achieve *Agent*(α)'s goal behind α. *"A: Let's meet on Monday. — B: No."* | 2.2.3.16 |
| $Ack_q$ | positive answers γ to β entails *Ack*(α, γ), negative *Plan-Corr*(α, γ). *"A: Let's meet on Monday. OK?"* | 2.2.3.18 |
| Comment | | |
| *Comnt* | β indicates a propositional attitude of *Agent*(β) towards the content of α. *"A: I talked to Peter. — B: Awesome!"* | 2.2.3.20 |
| $Comnt_q$ | Answers to β indicate a propositional attitude of *Agent*(α) towards the content of α. *"A: I talked to Peter. — B: Really?"* | 2.2.3.20 |
| Narration | | |
| *Narr* | $e_\beta$ occurs after $e_\alpha$, ... *"A: He went to Italy. And (then) to Spain."* | 2.2.3.19 |
| $Narr_q$ | Answers γ to β entail *Narr*(α, γ). *"A: He went to Italy. — B: And then?* | 2.2.3.19 |

Table 2.3: Speech act types that can be realized with NSUs: Plan-Related Relations to non-message-type NSUs (*Comment* and *Narration*)

we leave this to future work. Finally, note that these types do not impose mutually exclusive classes on the data, i.e., one fragment can stand in more than one relation to more than one α or, put differently, can fulfill more than one function in a dialogue.

#### 2.2.3.2 Question-Answer-Pair

Here is the informal definition of the speech act / rhetorical relation that connects answers to their questions:

| | |
|---|---|
| *QAP* | β provides a direct answer to α. |

To explain the terminology used in this definition, following (Groenendijk & Stokhof 1984) we understand under 'direct answer to a question' those that provide exactly the information that was asked for (e.g., "Who came to the party?" "Peter.").[7] We will talk later about 'partial answers' as well, which are those that at least exclude possible answers (e.g., "Not Peter."), and also 'indirect answers', which provide information from which a direct answer can be inferred ("I saw Peter.").

We have of course already seen many examples of this speech act type, and indeed it is one of the most frequent ones found in corpora (see Section 2.2.4). The following two examples show two randomly picked instances from the BNC:

(18) `1.4.610 A:` Who is it?
`1.4.611 B:` Pope Innocent the Fourth.

(19) `1.3.434 C:` Who was doing the interviewing?
`1.3.435 A:` *seven ladies*

Short-answers need not consist of noun-phrases, as the following examples show.

(20) a. A: What is John doing?
B: Suppressing dissent.
(from (Morgan 1973))

b. A: What are you doing?
B: Timesing it by X.
`[BNC FMM 404]`

c. A: How do you feel?
B: Hungry.

[7] We leave aside here issues of exhaustivity, i.e. of whether a direct answer has to be one which provides the full answer ("Peter came, and nobody else.") or whether it is enough that it provides a positive answer, as above.

d. A: When did Peter arrive at the party?
B: At midnight.

These examples, as well as most of those we have seen before, are *resolution-via-identity* fragments. However, we can also find examples of *res-via-inf* short answers, depending on the type of questions (this point will be discussed below). For instance, in (21) (a variation of a turn from the restaurant dialogue (1) from Chapter 1), we see a short answer that has to be resolved using inference.

(21) Peter: Why is it good for me?
Mary: Lots of vitamins.

Just to reinforce this point, here is another such example, followed by an instance of this speech act found in a corpus.

(22) a. A: Why can't you come tonight?
b. B: Too much to do.

(23) A: Now you're on the run from the Army at the moment?
B: Mhm.
A: Why did you run away?
B: Mental torture. I just couldn't handle it anymore.
[BNC HVD 279]

Note that despite these differences in the potential for being followed up by different kinds of fragments, we have chosen not to distinguish these pairs with different rhetorical relations. The semantic relation between the utterances, which as we said is the main criterion for individuating rhetorical relations for us, is the same in all cases, namely one of answerhood. The difference lies in the compositional semantics of the *wh*-element. The examples above all feature *why* as the *wh*-element; an indication, as we will see below, that there is another relation besides *QAP* present as well, namely *Explanation*. However, we can also find *res-via-inf* examples of short answers to *how*-questions, where *QAP* is the only relation, as in the examples below in (24).

(24) a. A: How do you make common salt?
B: Sodium and chlorine.
[BNC FMR 0387]
b. A: How would you copper plate things?
B: Electrolysis.
[BNC FMR 1351]

We will discuss the relation of this distinction between *res-via-id* and *res-via-inf* short answers to the distinction between complement-questions and adjunct-questions and to further constraints on the fragment below in 2.3 and at appropriate places further on. To anticipate, it seems that complement-questions like *who*- and *what*-questions always require *res-via-id*-fragments, whereas adjunct-questions are less restricted, as it were, and can trigger inferences to recover missing material. There seems to be in fact a continuum of 'strictness' in the relation of kinds of questions to their short-answers. At the one end of this continuum stand complement-questions like the ones in (18) and (19) above, which always resolve via *identity*. Somewhere in the middle we find *where*- and *when*-questions like those in (25) below, where the question word is syntactically a PP, but the answer can be an NP (and so a preposition has to be inferred). In (25-b), even more inference is required, since the NP is one denoting an activity and a deduction of a location is required. Towards the end of the spectrum we find successively more 'clause-like' answers, like those above in (24) and (23).

(25) a. A: When will he arrive?
B: 5 o'clock.

b. A: Where is everybody?
B: Playing football.

As we said, this observation will be further discussed below in Section 2.3 and in Chapter 4; we now return to our taxonomy of fragmental speech acts. Before we move on, we should also mention that we group direct answers to polar questions under this heading as well. They simply resolve to (the negation of) the propositional core of the polar-question.

(26) A: Did Peter show up?
B: Yes. / No.

Finally, fragments can also realise partial answers, as shown in (27).

(27) A: Who came to the party?
B: Not Peter.

The appropriate speech act type is defined as follows.

| | |
|---|---|
| *QPAP* | $\beta$ provides a partial answer to $\alpha$. |

#### 2.2.3.3 Question-Answer-Pair, $_q$-Version

As already shown above with (15-c), there is a speech act type that is closely related to *QAP*, but differs in that the fragment resolves to a question: $QAP_q$. Here is our informal definition of its semantics:

| | |
|---|---|
| $QAP_q$ | β is a y/n-question, and positive answers to it provide a direct answer to α, negative answers a partial answer. |

The following example shows in what sense this speech act is similar to *QAP* from the previous section: the same NP (modulo differences in intonation) can serve as an argument to $QAP_q$ (28-a) or as one to *QAP* (28-b).

(28) A: Who was this?

a. A: Peter?

b. B: Peter.

This simply follows from the definition above: to be resolvable to a polar-question of the required type, the fragment β must be such that it could be resolved to be a direct answer to α, were it uttered with declarative intonation. Or, to phrase it differently, the propositional core of the polar question β is a direct answer to α.

Fragments of this type again are fairly frequent, as we will see below in Section 2.2.4. (29) presents an instance of this speech act type from a corpus. Both the last utterance in `f.30` and utterance `f.32` stand in this relation to the question "where's the u-shape?".

(29) `q1ec7.f.30:` Hang on, where's the u-shape, *just underneath the mill?*

`q1ec7.g.31:` No, no.

`q1ec7.f.32:` *Just underneath the caravan park?*

`q1ec7.g.33:` Yeah.

A further apparent constraint on this speech act, namely that the $QAP_q$-fragment has to be uttered by the same speaker as α (or at least from someone collaboratively involved in wanting to know an answer), can be explained with more general constraints on cooperative use of questions. Normally (i.e., in default 'language-games', unlike for example exam-type situations), someone who asks a question is supposed to not already know its answer. So if an interlocutor replies to a question with such a fragmental counter-question, she implies that $Agent(\alpha)$ knows an answer to her own question, contrary to the cooperativity assumption mentioned above. This explains why the following exchange forces a construal of the dialogue-context as an exam- or quiz-situation.

(30) A: Who discovered America?
B: Isaac Newton?

Since this speech act type is our addition to the taxonomy of SDRT, we provide in (31) an attested example of a full-sentence instance as well, to show its general use.

(31) A: Why did President Clinton agree to give him [Gerry Adams] a visa?
Was it to keep the Irish population happy?
[BNC JSL 048]

#### 2.2.3.4 Elaboration

The definitions of the previous two speech act types left implicit what the sentence mood of α has to be, since it follows from the constraint that β (or utterances connected to β) have to be *answers* of α—hence α has to be a question. However, elaborations can elaborate propositions and questions—indeed, the elaboration itself can be a proposition or a question. So to distinguish these different versions, we will index the name of the relation with the semantic types of its arguments, so that for example $Elab_{qp}$ stands for the version where α denotes a question and β a proposition.

Here now is the definition of the most basic version, where both α and β are propositions.

| | |
|---|---|
| $Elab_{pp}$ | β elaborates on some aspect of the indicative α, e.g. by giving details about a sub-event of the event described in α, or by providing more information about participants involved in the event. |

Elaborations of this type are shown in the following example, where (32-a) is a specification of a participant in the event which α describes, and (32-b) provides more information about the whole event, by 'adding' an adjunct. (32-c) shows a special construction with which fragmental elaborations can be realised, which in effect 'adds' a conjunction to an element of α.[8]

(32) a. A: I went to the cinema.
The Odeon.
b. A: I went to the cinema.
With Peter.
c. A: I went to the cinema with Peter.
And Sandy.

Instances from corpora corresponding to these examples are given in (33), (34), and (35). In turn

[8] Note that all usual ambiguities connected with conjoined NPs (collective vs. distributive readings, etc.) are preserved in the resolved fragment.

r098c_NKH004 in (33), the NP "Thursday, June the third" elaborates the "Thursday afternoon" of the previous fragment, while the fragment 1.1.761 in (34) resolves in this context to "we haven't seen each other since that peculiar meeting with the language lecturers at the examiner's meeting.", and so elaborates on the event denoted by the utterance it attaches to. B's utterance in (35) finally adds a conjoined NP and thus elaborates on the object of the predicate in A's previous utterance.

(33) r098c_SMA003: [...] we will have to go to the following week, erm, either Thursday afternoon, or Friday morning.
r098c_NKH004: Thursday afternoon, for sure. Thursday, June the third.

(34) 1.1.757 A: We haven't seen each other
1.1.758 A: since that peculiar meeting with the
1.1.758 A: language lecturers.
1.1.759 A: Remember?
1.1.760 B: aha
1.1.761 A: *At the examiners' meeting.*
1.1.762 A: Right.

(35) A: And we had erm *[pause]* chicken noodles with bamboo shoots.
B: And mushrooms.
[BNC KC4 0272]

Elaborations can also concern optional arguments of elements of $\alpha$ that aren't realised in $\alpha$, as in the following examples. We will come back to this observation when we discuss the additional *syntactic* constraints on this kind of fragment below in 2.3 and in more detail later in the thesis.

(36) a. A: JFK was shot.
By the CIA(, to be precise).
b. A: I gave her a solemn promise.
(Namely,) that I'll never sing in the shower ever again.
c. A: I took a picture.
Of Madonna, no less.
d. A: Peter was reading when I saw him.
A book about Montague semantics, I think.
e. A: Die Verleihung war spektakulär.
[Des Oskars]$_{gen}$, meine ich.
*The handing over of the award was spectacular. Of the Oscar, I mean.*

The elaborations we have seen so far were all examples of *resolution-via-identity* fragments. We can however also construct fragmental elaborations of the *resolution-via-inference* kind, as in the following mini-dialogues.

(37) a. Max had a lovely evening yesterday.
A nice film, lots of good food, wine, dancing.

b. I went to the cinema yesterday.
Spiderman.

Now, fragments like these seem not to sound as natural as those discussed before, which presumably is due to the high inferential effort involved in constructing an event which involves the entity denoted by the fragment that is appropriate for the demands of the rhetorical relation. Nevertheless, the elaborations in (37) are not judged as infelicitous; (37-a) for example is seen to resolve to something like "Max saw a nice film, ate lots of good food, drank wine, did some dancing". Only with a resolution like this does β satisfy the semantic constraint on $Elab_{pp}$ given above.

This was the version of elaboration where the elaborated utterance denotes a *proposition*. Now, what shall we say is the semantics of a fragment like that in (38),[9] which seems to elaborate on a *question*?

(38) A: Did Peter call?
Today*(, I mean)*.

Intuitively, (38) can be paraphrased as "I meant 'Did Peter call today?"'. This paraphrase gives an indication of how we should define the semantics of the rhetorical relation present in this case, namely via *intentions*. (This is the first example of a speech act type which is defined in terms of intentions of discourse agents; we will see more below. In particular, we will discuss in Subsection 2.2.3.15 the relation of this speech act type to a similar one, *Q-Elab*, which is also makes reference to intentions.)

| | |
|---|---|
| $Elab_{qp}$ | β elaborates on the intended meaning of α, e.g. by giving details about a sub-event of the main-event of α, or by providing more information about participants involved in the event. |

Example (39) shows a corpus instance of this speech act.[10]

(39) A: Why don't we get married?
Today.

[9] It is difficult to precisely specify what the intonation of this kind of fragment is, interrogative or declarative. In the corpora we looked at, they mostly have been transcribed using full stops, i.e. as having been uttered with declarative intonation. This matches our intuitions, although a certain degree of uncertainty remains.

[10] But note that this is not from the conversation part of the BNC, and so does not actually constitute an example of this speech act in dialogue.

[BNC ACE 0126]

Note that α in this relation is not restricted to be a polar question; *wh*-questions can be elaborated in the same way, as (40) shows.

(40) A: Who was that on the phone?
This morning*(, I mean)*.

We will also subsume fragments like (41) under this class, where the additional information presumably is provided to help the hearer to identify the referent of an expression. Why we have decided not to model this with a separate speech act type will be discussed below in relation to clarification questions.[11]

(41) A: Did Peter call?
Peter Miller.

Finally, it seems that *resolution-via-inference* fragments are harder to get for this kind of elaboration, as (42) illustrates. The reason for this might be that this kind of elaboration restricts the question in such a detailed way that it is odd not to have phrased the question differently in the first place. We will return to this point later when we formalise the semantics of this speech act.

(42) a. A: Did you go to the cinema?
#Spiderman.
*(= Did you see (the film) Spiderman?)*
b. A: Did he have a good time?
?#A nice meal, etc..
*(= ??Did he have a good time, for example by eating a nice meal, etc..)*

#### 2.2.3.5 Elaboration, $_q$-Version

This speech act type is the 'dual' to $Elab_{pp}$ in the same way as $QAP_q$ is the dual to *QAP*: all answers to this speech act stand in that relation to α, and also β would count as an $Elab_{pp}$ were it not uttered with interrogative intonation. Here is the informal definition of this relation:

| | |
|---|---|
| $Elab_{pq}$ | Any answer to β elaborates on some aspect of the indicative α, e.g. by giving details about a sub-event of the event described in α, or by providing more information about participants involved in the event. |

[11] In a more general sense, (38)–(40) can also be construed as cases where the speaker provides help in identifying the referent of an expression—in this case the identity of the event that is part of the compositional semantics of the VP.

The examples from the previous section can easily be transferred. (43) shows some of them adapted to this speech act type; (44) gives an example from the VM-corpus.

(43) a. A: I went to the cinema.
B: The Odeon?
b. A: I went to the cinema.
B: With Peter?
c. A: JFK was shot.
B: By the CIA?
d. A: I gave her a solemn promise.
B: That you'll never sing in the shower ever again?

(44) `r083c_JMO004:` I am available [in the] afternoon on Tuesday.
`r083c_SMA005:` *Tuesday the sixteenth?* That looks good to me.

Note that non-realized optional arguments of elements of $\alpha$ can also be elaborated with fragments (45-a), and further, that *res-via-inf* fragments seem to be relatively natural again here (45-b).

(45) a. A: I made a purchase.
B: The ugly hat you're wearing?
b. A: I went to the cinema yesterday.
B: The new James Bond?

As mentioned above, we will subsume what often is called 'clarification question' as in (46) below under this class as well. As we will argue when we formalise this relation, the intended effect of asking something like "is the person you refer to as 'Peter' the person I refer to as 'Peter Miller'?" will fall out of the general definition of $Elab_{pq}$ as a side-effect, and so does not have to be modelled with a relation specific to this use of elaboration-questions.

(46) A: Peter called.
B: Peter Miller?

In the examples we have seen so far, $\beta$ is always resolved to a polar question. We can however also find examples where the fragment is intended to convey a *wh*-question; these are fragments that consist of a bare *wh*-phrase, as in the following examples.[12] Our definition of $Elab_{pq}$ given above is general enough to cover these cases as well.

[12] Fragments consisting only of "why" will also fall under this class, but note that they trigger additional semantic consequences that are modelled by another relation, namely $Explanation_q$, which will be defined below.

(47) A: You were funny with that wheel-barrow out the front here this afternoon.
B: Who?
A: You.
[BNC KD6 1661]

(48) A: It's a microphone. Recording conversations.
B: With who?
A: Everybody.
[BNC KD5 1514]

Example (47) shows a corpus example of such a fragmental query asking for a specification of a complement of α, whereas (48) asks for more details about the event described in α. (The elaboration in (48) is between B's utterance and A's second utterance.) Again we make no distinction between 'clarification-questions' ((47) would presumably be one) and other kinds of elaborations like (48).

There is also a special form of fragment that expresses clarification queries, namely where an element of α is repeated *verbatim*:

(49) A: Did Shmul call?
B: Shmul? *(= Who is Shmul?)*

Finally, we come to the last combination of sentence modes / semantic types in elaborations: $Elab_{qq}$. Again, the definition can be adapted from $Elab_{pq}$ as before, and the examples given for the '$_p$-version' are readily transferred.

| | |
|---|---|
| $Elab_{qq}$ | Any answer to β elaborates on the intended meaning of α, e.g. by giving details about a sub-event of the main-event of α, or by providing more information about participants involved in the event. |

(50) a. A: Did you go to the cinema?
B: The Odeon?
b. A: Did you go to the cinema?
B: With Peter?
c. A: Did you go to the cinema with Peter?
B: Peter Miller?

Again, elaborations of optional arguments and *res-via-inf* cases seem to be much harder to get:

(51) a. A: Did you make a purchase?

B: ?That hat?

b. A: Did you go to the cinema?
B: The new James Bond movie?

Once more we postpone the discussion of the relation of this speech act type to *Q-Elab*, which similarly refers to intentions, until this latter speech act type has been introduced. What we can discuss here is how this speech act relates to $QAP_q$, to which it at least superficially is similar in that this relation also takes two questions as arguments. However, $Elab_q$s, elaborate questions in the sense that they further specify details about participants in the questioned event or about the event itself, whereas $QAP_q$-acts as defined above are intended to restrict the range of possible answers (and in that sense 'elaborate' on the *wh*-phrase in the question): answers to β in this relation simply *are* answers to α. This difference is best captured with a different speech act.

Finally, there is a context in which this speech act can express surprise, as in the following.

(52) A: Peter called.
B: Peter? But I thought he's away for the week?

We analyse this as a further implicature of the performance of an $Elab_{qq}$ and not as a separate speech act.

#### 2.2.3.6 Correction

Another frequent use of non-sentential utterances in dialogues is as corrections of elements of previous utterances. The following examples show first a variant of a turn from the restaurant dialogue (1) from the previous chapter, and then a few instances of this speech act type from a corpus. Note that the "no" is analysed here as discourse particle, signalling the presence of this relation; this cue word is optional, as (56) demonstrates.

(53) Paul: You eat this with a fork.
Mary: No, with chopsticks.

(54) A: You're a hamster you are. You store food in your cheeks.
B: No, pouches.
[BNC KCH 4143]

(55) A: Nearly a thousand already.
B: What number?

A: On number eight. No, seven.
[BNC KCU 05876]

(56) A: Speed limit's ninety around here.
B: Seventy.
[BNC KSV 1142]

Here is the informal definition of this relation:

| | |
|---|---|
| *Corr*$_1$ | β corrects (an element of) α. |

Looking closer at these examples, it seems that there is a constraint that the corrected element must be in focus in α. In the following example, for instance, the reading of the fragment indicated in brackets is rather odd, given the focus/background-structure of α.

(57) A: Peter loves [Sandy]$_F$.
B: No, Carl.
*(= #Carl loves Sandy)*

This contrasts with the full sentence case, in which a non-focussed element can be corrected:

(58) A: Peter loves [Sandy]$_F$.
B: No, [Carl]$_F$ loves Sandy.

We however think that there is an interfering issue here. Backgrounded material is normally assumed to be material whose truth has been settled, and that explains why it is odd to correct this material. A speaker is expected (given assumptions of cooperativity and rationality) to dispute only new material; otherwise, as shown in (58), a full, non-ambiguous expression has to be used to override this preference. We will elaborate on this in Chapter 8.

We also need a version of this speech act where α is a question, to describe examples like the following.

(59) A: Who did you see with Peter? Erm, no, Sandy.
*(= What I meant to ask was: "who did you see with Sandy?")*

As the paraphrase given in the example indicates, the semantics of this relation must be defined in terms of intentions. (Again, there is a related relation, *Plan-Correction*; we will discuss the distinctive features below in Subsection 2.2.3.16.)

| | |
|---|---|
| *Corr*$_2$ | β corrects (an element of) the intention behind uttering the interrogative α. |

Note that fragments like (60) will not be classified as *Correction*s in our taxonomy, but rather as $Elab_{qp}$ (with an additional *Contrast* relation (see below) connecting the two fragments).

(60) A: Who did you see with Peter?
Not Peter Miller, Peter Smith.
*(= Who did you see with Peter Smith?)*

Finally, it seems that *Correction*s always are *res-via-id* fragments, and that no $_q$-version of this relation is required.

#### 2.2.3.7 Contrast

Here is the definition of this speech act type:

| | |
|---|---|
| *Contr* | α and β have a *contrasting theme*. |

The following example illustrates what is meant with 'contrasting theme'. The theme of A's utterance is something like 'things that John loves', whereas that of B's utterance is 'things that John doesn't love', and so these two themes are maximally contrasting.

(61) A[to C]: John loves all kinds of sport.
B[to C]: But not football.

A common situation in which this speech act type occurs is as follow-up to a negative answer to polar questions, as in the following examples:

(62) a. A: Were they in tents?
B: No, caravans.
[BNC GYS 020]

b. A: Is your brother older than you?
B: No, younger.
[BNC HE1 116]

c. A: Were they nicer than usual?
B: No, different.
[BNC KBK 3344]

We will analyse negative answers to polar questions as resolving to the negated proposition from the polar question; for instance in example (62-a), the "no" resolves to "they weren't in tents". The fragment in these examples above then forms a contrast to this proposition; for (62-a), this is “they were in caravans”. Note that, as demonstrated in (63), these contrasts must be licensed by a certain focus/background structure of $\alpha$ (where we assume that that ‘no’ takes on the same partitioning as the question it answers).

(63) a. A: Did Peter $[\text{sing}]_F$?
B: No, shout. / #No, Sandy.
b. A: Did $[\text{Peter}]_F$ sing?
B: No, Sandy. / #No, shout.

The constraint of having a contrasting theme indicates that the argument of this relation will be relatively close structurally, and indeed there do not seem to be *res-via-inf* examples of this relation.

#### 2.2.3.8 Continuation

The following fragment is an instance of the class *Continuation*. This class is most often realised with ‘CONJ XP’-fragments.[13]

(64) A: I’m free on Monday.
And on Wednesday afternoon.

Here is the definition of this speech act.

| | |
|---|---|
| *Contn* | $\beta$ continues a topic of $\alpha$. |

There is also a related, and actually more frequent speech act which we call *Q-Contn*. It is defined as follows.

| | |
|---|---|
| *Q-Contn* | The question $\beta$ continues enquiry about a topic of the question $\alpha$. |

Example (65) shows an instance of this relation from a corpus. The continued topic here is something like “things the hearer has [on her map]”.

[13] There is a general problem with this kind of fragment form: since they, unlike most other kinds of fragments, form a grammatical sentence when adjoined to $\alpha$, it is difficult to decide whether constructions like this do indeed form independent fragments or just intonationally separated segments of previous utterances. However, since fragments like these can come from different speakers as well, as we will see in the next sections, we have decided to view them as fragments.

(65) `q1ec8.g.73:` Right, have you got the lake?

`q1ec8.f.74:` No.

`q1ec8.g.75:` No.

`q1ec8.f.76:` To the ... no.

`q1ec8.g.77:` *Roman baths?*

`q1ec8.f.78:` No.

`q1ec8.g.79:` Hmm. *Antelopes?*

`q1ec8.f.80:` Mm, they're over to the east of the mountain.

Here is another example, taken from (Carberry 1990, p.161). The topic here is "properties of CS360".

(66) A: What is the meeting time of CS360?
B: 7.00 P.M. on Monday night.
A: The meeting place?

Depending on how broadly we define topic-hood, it is possible to imagine *res-via-inf* cases of this speech act type, as for example in the last utterance in example (67) below (taken from (Carberry 1990, p.184)). However, as we will see when we formalise the semantics of this relation, in the interest of a manageable notion of topic we will have to exclude cases like this, and so conclude that *Continuation*s are always *res-via-id*.

(67) A: On what days does CS440 meet next semester?
B: Two sections meet on Monday and one section meets on Tuesday next semester.
A: What time do the sections on Monday meet?
B: One section meets on Monday at 4:00 pm and another section meets on Monday at 7:00 pm.
A: The texts?

#### 2.2.3.9 Q-Alt

(68) instantiates what we call *Q-Alt* (for 'question alternation'). The intended reading of the fragment here is something like "on which of the following days are you free: Monday, Tuesday?", i.e. an alternative-question, to which one has to reply with an NP, and not the question "are you free either on Monday or Tuesday?" which must be answered with "yes" or "no".

(68) A: Are you free on Monday?
Or on Tuesday?

We define the semantics of this speech act as follows.

| | |
|---|---|
| *Q-Alt* | $\beta$ resolves to an alternative-question involving an element of $\alpha$. |

#### 2.2.3.10 Explanation

This speech act type is defined as follows:

| | |
|---|---|
| *Expl* | $\beta$ explains $e_\alpha$. |

Again we first give a constructed example and then some instances from corpora.

(69) A: Peter left early.
B: Exams tomorrow.

(70) `1.1.143 B:` I've got about a week
`1.1.144 B:` of fairly hard work
`1.1.145 B:` after the fourth of July.
`1.1.146 B:` *This CSC _stuff*,
`1.1.147 B:` you see.

(71) `1.1.909 B:` and I'm determined
`1.1.910 B:` to get that sort of stuff
`1.1.911 B:` into the comprehension questions,
`1.1.912 B:` rather than all this high-faluting literature
`1.1.912 B:` stuff
`1.1.913 B:` where they can set imagery
`1.1.914 B:` and all that kind of thing.
`1.1.915 B:` Because the
`1.1.915 B:` scientists don't want that sort of stuff.

`1.1.917 A:` Mhm.
`1.1.918 A:` *Far further from the students'*
`1.1.918 A:` *experience*
`1.1.919 A:` *and so on*
`1.1.920 A:` yes

Note that these fragments all are *res-via-inf*. This is to be expected given the semantics of the relation, since the fragment has to provide an independent event that can explain $\alpha$, and so no great degree of structural similarity is likely.

#### 2.2.3.11 Explanation$_q$

There exists a $_q$-type speech act corresponding to *Expl*, as shown in the following example:

(72) A: We finish at one o'clock next term.
B: The heat?

We define the relation as follows.

| | |
|---|---|
| $Expl_q$ | All answers to β explain $e_\alpha$. |

Like the *Explanation*s we have seen in the previous section, this seems to allow *res-via-inf* fragments.

There is also a special realization of this speech act type, namely via the question word "why", as in the following example.

(73) A: We finish at one o'clock next term.
B: Why?
A: Because of the heat.
[BNC KCK 1183]

#### 2.2.3.12 Explanation$_q^*$

Upon further inspection, it seems that there is a systematic ambiguity in the intended meaning of such *why*-fragments. The relation $Expl_q$ captures the reading where an explanation of the event is requested, but we can also understand such fragments as asking for the intention behind uttering α. (74) shows an example where such a latter reading seems to be preferred.[14]

(74) A: Did dad tell you what happened this morning?
B: No, why?
[BNC KD5 3268]

In SDRT such speech acts are called 'meta-linguistic' because they are about speech acts and not their content, and are distinguished with a superscripted star. Here is the definition of the class to which we assign (74):

| | |
|---|---|
| $Expl_q^*$ | All answers to β explain why α has been uttered. |

As the next example shows, this speech act is not restricted to *why*-questions:

[14] In general, this reading seems to be the preferred one when "why" follows a question.

(75) A: Can we go now?
B: Tired?

Also, it is not restricted to an $\alpha$ that denotes a question, as (76) shows.[15]

(76) A: You look tired.
B: Why?

In general such $^*_q$-fragments seem odd if they come from the same speaker as $\alpha$; this can be explained with assumptions about cooperativity: the speaker simply is supposed to know his own speech act related goals (whereas the hearer can only infer them), and so the speaker does not have to (and can't even coherently) inquire about them.[16]

#### 2.2.3.13 Result, Result$_q$

We also have in our taxonomy a speech act type where the order of explanans and explanandum is reversed compared to *Explanation*, as illustrated by the attested instance (77) (the relation holds between the last utterance and the one before it).

(77) A: And it was the First World War that killed him, he had a brain damage, you know, a stroke. And died.
`[BNC FXW 033]`

The class is defined as follows.

| | |
|---|---|
| *Res* | $e_\beta$ is the result of $e_\alpha$. |

Again, we can easily construct a question version of this, as the following example shows.

(78) A: He had a stroke.
B: And died?

The definition accordingly is:

| | |
|---|---|
| *Res* | If $\gamma$ is a positive reply to $\beta$ then $e_\gamma$ is the result of $e_\alpha$. |

[15] However, the difference between these two readings is less striking for propositional $\alpha$s, because the event that caused another event will also be a cause for saying something. And so an answer like "you have shadows under your eyes" to (76) both satisfies an $Expl_q$ and an $Expl^*_q$ reading.

[16] This presupposition does not hold for rhetorical questions, as in "Did we ask you to accept that? No. Why not? Because . . . ".

Note that these examples, like all fragments that are introduced by a conjunct, are *res-via-id* cases. For example, the constructed variant (79) below of (78) sounds odd.

(79) A: He had a stroke.
B: ?#And death?

So far we have focussed on relations whose semantics are defined in terms of the contents of the things they connect—the so-called *content-level* speech acts. We come now to the family of speech acts whose semantics are defined in terms of *intentions* or *plans* of the dialogue participants, as well as perhaps in terms of the contents of the propositions (or questions) being related. These are called *plan-level* speech acts.

#### 2.2.3.14 Plan-Elab

*Plan-Elab* is another of SDRT's speech act types that can be realised with fragments. The informal definition goes as follows:

| | |
|---|---|
| *Plan-Elab* | β details a step in a plan to reach a goal behind α. |

The next two examples illustrate what falls under this definition.

(80) a. A: I want to cash this cheque, please.
Small bills only.

b. A: Let's meet sometime next week.
B: OK, but not on Tuesday.

Example (80-a) is taken from (Carberry 1990). The fragment resolves to something like 'I want small bills only', which specifies a step in a plan to achieve the goal of α, which is to get the cheque cashed. (80-b) is an example from the scheduling dialogue domain; the fragment indicates a plan to reach the goal to meet insofar as it narrows down the range of possible times in that interval. These two examples show that fragmental *Plan-Elab*s can be both *res-via-inf* and *res-via-id*.[17]

There is a general issue here of how these plan-related speech act types relate to others that mention intentions defined earlier. For example, what is the difference between $Elab_{qp}$ and *Plan-Elab*? Is the former a special case of the latter? To explain our position, we have to briefly go into what kinds of goals can be normally associated with a speech act. First, there is the goal of being heard or 'acoustically understood'; then there is the goal of being understood proper, i.e. of conveying the intended message;

[17] *Plan-Elab* is not the only rhetorical relation that can be inferred for these examples; in (80-a) there is possibly also a normal *Elab*, whereas in (80-b) there is also a *Contrast*. But, as noted above, plurality of rhetorical connections is one of the features of SDRT.

there is also a goal of being believed (in the case of assertions); finally, there are more 'higher level' goals, like for example knowing an answer to a question. We now deviate a bit from SDRT, which seems to include all these kinds of goals in their plan-related relations, and restrict this relation and the following to the kind of goals mentioned last above. This explains how *Plan-Elab* and $Elab_{qp}$ are distinguished: at least in some uses, the latter attends to the goal of being understood, whereas the fragments above in (80) attend to explicitly stated domain goals.

#### 2.2.3.15 Q-Elab

There is also a $_q$-version of the previous speech act type, which we however, following the SDRT-nomenclature, call *Q-Elab*. Here is the definition of this discourse relation:

| | |
|---|---|
| *Q-Elab* | Answers to $\beta$ detail a step in a plan to reach a goal behind $\alpha$. |

An example will make clearer what is covered by this definition. Answers to the questions in (81) (variations of the examples given above) will provide an indication about how the respective goal is to be realised.

(81) a. A: I want to cash this cheque, please.
B: Small bills?
b. A: Let's meet sometime next week.
B: (OK.) Tuesday?

The following two examples are taken from a corpus. In (82-a), the fragment "Monday [. . . ]" elaborates a plan to meet "sometime over the next two weeks", whereas in (82-b) the "how about [. . . ]" is resolved to "how about *we meet* [. . . ]", given the goal of the dialogue.

(82) a. `r037c_DEL000:` okay , so I guess we have to , get together again sometime over the next two weeks . {erm} *random starting point,* {*erm*} *Monday or Tuesday of the next two weeks , sometime after noon*?
b. `r037c_CAW007:` Saturday or Sunday afternoon would be pretty good for me. [. . . ]
`r037c_DEL008:` okay, how about Saturday, two to four pm?

These examples given in this section are all *res-via-inf*.

Again, we restrict the kind of goals figuring in the definition in the way described above for *Plan-Elab*. The speech act type *Pardon* defined below is related to *Q-Elab* in that it attends to the goal of being heard, whereas (some uses of) $Elab_{qq}$ attend to the goal of being understood.

#### 2.2.3.16 Plan-Correction

Another plan-related speech act type we need is called *Plan-Correction*, and is defined thus:

| | |
|---|---|
| *Plan-Corr* | β indicates that *Agent*(β) doesn't accept or is unable to help achieve *Agent*(α)'s goal behind α. |

A very simple example of this is the following.

(83) A: Close the window.
B: No.

The same restriction as above as to which goals are corrected holds (namely, only those not connected to understanding the message). Hence, *Plan-Corr* relates to $Corr_2$ in the same way as *Q-Elab* does to the $Elab_{.q}$ relations.

#### 2.2.3.17 Acknowledgement

The relation we call *Acknowledgement* is the dual to *Plan-Correction*; it is illustrated by the following example.

(84) A: Move down below the mill.
B: Down below the mill, OK.

The definition of this speech act goes as follows:

| | |
|---|---|
| *Ack* | β entails that *Agent*(β) has accepted or achieved *Agent*(α)'s goal behind uttering α. |

Again, there is a counterpart of this relation that has a similar function towards the goal of being heard / understood, namely *Backchannel* as defined below.

#### 2.2.3.18 Acknowledgement, $_q$-version

Finally, there is a $_q$-version of the previous speech act, defined as follows, and illustrated by (85).

| | |
|---|---|
| $Ack_q$ | β is a polar question, and positive answers to it are connected to α via *Ack*, negative answers via *Plan-Corr*. |

(85) A: Then you move down until you walk past the hut.
Okay?

(86) 2093 let us see how our schedules are, for the following weeks,
2094 okay?

We again want to reserve this speech act for non-understanding related goals, but note that it might in practice be difficult to do so, and utterances might be systematically ambiguous in this respect.

#### 2.2.3.19 Narration, Narration$_q$

We end this section with a look at some more restricted fragmental constructions. Fragments beginning with a conjunction can play a variety of roles, one of them being driving forwards a narration, as in the following example.[18]

(87) A: He drove to Italy. And then to Spain.

In this example the fragment ‘and then to Spain” forms a conjunction with the PP of the previous utterance, and, triggered by the “then” discourse particle, it introduces a new event. Note that the fragment without the “then” would presumably be interpreted as an elaboration.

In the following corpus example, the fragment is of the form “and VP”, and also forms a narration with the previous utterance (this is (77) from above; *Narration* is not the only relation present here).

(88) A: And it was the First World War that killed him, he had a brain damage, you know, a stroke. And died.
[BNC FXW 033]

We give a definition of this speech act below:

| | |
|---|---|
| *Narr* | $e_\beta$ occurs after $e_\alpha$, and entities referred to in $\alpha$ and $\beta$ are located at the beginning of $e_\beta$ where they are at the end of $e_\alpha$ (i.e., the two events overlap spatio-temporally). |

It is easy to construct examples for a $_q$-version of this speech act:

[18] Note our comments above in Footnote 13 about the problems with deciding whether constructions like this do indeed form independent fragments. We can add here in support of the decision to view them as such that in any case for sentences that are constructed like this (“He drove to Italy and then to Spain.”) rhetorical relations have to be inferred as well to capture the intended meaning that goes beyond compositional semantics (for this example, that the driving to Spain happened after the driving to Italy).

(89) A: He drove to Italy.
B: And then to Spain?

It also seems that there is a special version of this speech act where the fragmental query resolves to a *wh*-question (“and then what happened?”):

(90) A: He drove to Italy.
B: And then? / And?

The definition of this speech act $Narr_q$ is as expected:

| | |
|---|---|
| $Narr_q$ | All answers to $\beta$ describe events that occur after $e_\alpha$, and things referred to in $\alpha$ and $\gamma$ are at the beginning of $e_\gamma$ where they are at the end of $e_\alpha$. |

We close our survey with a list of remaining *non-message-type* classes.

#### 2.2.3.20 Some further *non-message-type* classes

Fragments of the type *Agree* signal that the speaker believes the other speaker, as in (91).

(91) A: Peter is an idiot.
B: Yes.

| | |
|---|---|
| *Agree* | $\beta$ indicates that $Agent(\beta)$ believes or agrees with the content of $\alpha$. |

Again, we do not classify this as an *Acknowledgement* because we have excluded the goal of belief-transfer from the definition of that relation.

Fairly frequent in dialogues are fragments that offer a commentary or an assessment of a previous utterance, as shown in (92).

(92) A: It has its own built in generator for the lighting.
B: Brilliant.
[BNC JPO 0609]

We call this speech act *Comment*, and define it in the following way:

| | |
|---|---|
| *Comnt* | $\beta$ indicates a propositional attitude of $Agent(\beta)$ towards $e_\alpha$. |

Backchannels are related to *Ack*, as mentioned above, and the class is defined as follows.

| | |
|---|---|
| *Backchn* | $\beta$ indicates that $Agent(\beta)$ understood $\alpha$, or simply is still attentive |

| | | | |
|---|---|---|---|
| (93) | PS1C0 | 0158 | But it means that she, she has a, a week's holiday and I think she's back for about three weeks and then breaks up again at school |
| | PS1BY | 0159 | Mm |
| | PS1C0 | 0160 | er Easter, so and then she'll have I think three weeks Easter |
| | PS1BY | 0161 | Mm |

The class labelled *Pardon* is the dual to the previous one, in that it signals the failure of the goal to be heard / understood.

| | |
|---|---|
| *Pardon* | $\beta$ indicates that *Agent*$(\beta)$ did *not* understand $\alpha$. |

(94) A: Did Peter call?
B: Pardon? / Sorry? / What?

Finally, we label as *Greet* all speech acts to do with beginning or ending conversations, i.e. "hello", "hi", "bye", but also certain uses of proper names, as in "Peter, (I want to tell you ...)". We forgo defining this speech act type, since it falls outside the range of what we later will be able to formalise.[19]

And this concludes our introduction of the taxonomy. In the following section we will study how this taxonomy fares with real-world dialogues; then in Section 2.3 we will look more closely at the other dimension for classification, that according to where the material in the resolved fragment comes from. In Section 2.4 we will compare our two-dimensional classification to others found in the literature; and finally in 2.5 we will summarise our findings and present some desiderata that can be derived from this survey.

### 2.2.4 Coverage

To get an idea of the frequency of fragments in typical dialogues (we have already cited these numbers in the introduction), and also to validate our taxonomy by finding out the percentage of "naturally occurring" fragments that can be classified with it, we ran a small corpus study, which we will describe in this section.[20] We describe in the next subsection our methodology for finding and annotating fragments, and then present, discuss and compare our results.

#### 2.2.4.1 Methodology

For this study we analysed material from two corpora, namely from the conversation part of the BNC, and from the verbmobil corpus.[21] More precisely, we analysed 5087 items of general free conversation from

[19] And also these acts are probably not best described as being relational; one main function indeed is to initiate a discourse.
[20] See for example (Walker & Moore 1997) for a discussion of the value of empirical studies in discourse.
[21] We held out this material during the stage of constructing the taxonomy, so as to avoid a manual "overfitting".

the BNC (dialogues KSU and KSV), and 4037 items of task-oriented dialogue from the VM/redwoods corpus (the 125 dialogues on the VM-CD-ROM 6).[22]

We proceeded in two stages. In a first pass, we semi-automatically marked up all NSUs, using the wide definition from the previous section: not a grammatical sentence, not the result of a disfluency. Semi-automatically, because we reduced the search space for both corpora: for the VM/redwoods data we used the wide-coverage grammar ERG (which will be described in some detail in Chapter 7) as a (rather sophisticated) filter, annotating only those items that the grammar did not parse as sentences.[23] For the BNC data we used a script that asked for a decision only for items shorter than a certain threshold of words, following the assumption that fragments in general are relatively short. We set the threshold to 5 words, which seemed a good cut-off point: judging by a manual check on a random sample it seemed to have covered all potential fragments, while still reducing the number of required decisions considerably.

In a second step we classified these NSUs according to the speech act types from the previous section, if possible, or with `other`, if not.[24] The results of this fine-grained annotation will be shown in the next section, but before we turn to that, a word about what constitutes the "items" in the corpora is in order, i.e. what the units are which we classified in the first step as potential fragment.

It turned out that the corpora differ slightly in how they segment the speech into transcribed items, and that had an influence on what had to be done during this first pass. First, the corpora differ in the treatment of dysfluencies. In the BNC, dysfluencies are marked up within an item, which continues until a sentence or at least a larger constituent is completed, if at all during the turn. A typical example is shown in (95-a). In the VM/redwoods corpus, on the other hand, dysfluencies like this lead to the utterance being split into several items, as the example (95-b) shows. Following our definition of fragmenthood ("not accidentally non-sentential"), we thus did not mark up in the VM/REDWOODS data items like ACM44 in (95-b) as NSUs in the relevant sense.

(95) a. PS1K5.1716 She wants to see who's interested in *#* po *#* in helicopters.

b. ACM 42 so if we can not, make it, on, Thursday afternoon, we
will have to, you know, look for something
43 on
44 between the fourth and sixth of August.

What we did mark as potential fragments are items resulting from interruptions, overlaps, and delivery in "installments" (Clark 1996). In the BNC, for example, overlapping or intervening material is transcribed

[22] More on what the "items" are in these corpora will follow presently; after the overview of the method.

[23] This might entail that we missed a few fragments, namely those that also have a sentential parse, as for example certain VP-fragments do (as imperatives); given the low number of VP-fragments we found in the BNC data, however, we expect the influence of this selection to be marginal.

[24] This annotation was done by the author. It would have been interesting to also use naive annotators and to compute inter-annotator reliability, however due to restricted time and resources we were not able to do this. We will however discuss below our confidence in the results.

as turns that break up the utterance at the closest constituent boundary. An example for this is shown in (96-a). In this example, we would mark item 0077 as a potential fragment, and not reconstruct a full utterance out of 0075 and 0077, since the decision of the annotators to split this might be based on information about pauses etc. that was not available to us anymore.

| | | | | |
|---|---|---|---|---|
| (96) | a. | PS1C0 | 0075 | but they want to go, bring all the trays of plants over |
| | | PS1BY | 0076 | A hundred, hundred odd |
| | | PS1C0 | 0077 | before before they go, so, so the greenhouse is empty |

In the VM/REDWOODS data, overlaps seem to not have occurred, or have been removed; here, however, we often find utterances that form grammatical sentences split into several items, as for example in (97).

| | | | |
|---|---|---|---|
| (97) | NBC | 3898 | so, how about from two to four. |
| | | 3899 | right after lunch. |
| | | 3900 | on Tuesday the ninth. |

Again, we decided not to revert this segmentation decision—which again might be based on information about pauses which was not available to us—, but we will have to keep in mind that this might have increased slightly the number of elaborations we find in the corpus.

Lastly, the VM/REDWOODS data seems to be 'cleaned' slightly in that no backchannels are annotated; they are quite frequent in the BNC data, as the following typical excerpt shows.

| | | | |
|---|---|---|---|
| (98) | PS1C0 | 0158 | But it means that she, she has a, a week's holiday and I think she's back for about three weeks and then breaks up again at school |
| | PS1BY | 0159 | Mm |
| | PS1C0 | 0160 | er Easter, so and then she'll have I think three weeks Easter |
| | PS1BY | 0161 | Mm |
| | PS1C0 | 0162 | and then I think she's back for, erm I think she gets back, [. . . ] well she'll be doing GCSEs anyway |
| | PS1BY | 0163 | Mm |

#### 2.2.4.2 Results

In the first pass, we marked 988 items out of the 5087-item selection from the BNC as NSU; that is 19.4%. At 13.6%, the number for the VM/redwoods corpus is lower; we will discuss the reason for this difference in a minute. Of the NSUs from the BNC, we were able to classify 96.1% into at least one of our classes from the previous section (recall that the classes are not disjoint, and so one fragment can fall in more than one), consequently 3.9% were classified as `other`.[25] The ratio is slightly worse for the VM/redwoods data: 95% falling into at least one category against 5% `other`. (These numbers are summarised in Table 2.4.)

| BNC | | | VM | | |
|---|---|---|---|---|---|
| items analysed | 5087 | | items analysed | 4037 | |
| NSUs | 988 | (= 19.4%) | NSUs | 550 | (=13.6%) |
| classfd. | 950 | (= 96.1% of NSUs) | classfd. | 523 | (= 95% of NSUs) |
| `other` | 38 | (= 3.9% of NSUs) | `other` | 28 | (= 5% of NSUs) |

Table 2.4: Results of annotation, all classes

Table 2.5 shows the distribution of types for the two corpora.[26] As this table shows, there is a marked difference between the corpora in the relative frequencies of the speech acts, presumably reflecting the difference in genre (free conversation vs. task-oriented dialogue). For example, explanations and requests for them do not occur at all in the VM setting—it seems that the dialogue participants always accept each others assertions—, while they are relatively frequent in the BNC data. The situation is reversed with acknowledgements: they are frequent in the VM data but don't occur much in the BNC data. Note that we ignored entailments between the classes for the purpose of this annotation task. For example, as mentioned above and discussed in more detail in Chapter 8, $QAP(\alpha,\beta)$ entails $IQAP(\alpha,\beta)$, but here we strengthened the definition of *IQAP* to only include 'genuine' indirect answers.

In view of the slight idealisations in the VM/REDWOODS corpus concerning things like backchannels which we discussed in the previous section, and considering our main focus on fragments proper, we have also compiled a statistic for just those types we labelled 'message-types NSUs' above. This statistic is shown in Table 2.6.

Interestingly, the percentages lie much closer together for the two corpora in this table, at around 10%. On the other hand, given the new basis for computation, the percentage of `other`-items is now increased, to around 7%. Does this mean our taxonomy is incomplete, or worse, inconsistent? A closer analysis of the unclassified items reveals that this is not the case. Many of those items in the BNC data were discourse initial,[27] and so could for principled reasons (our speech acts are *relations*) not be annotated. Other items were indexical, in the way demonstrated by the constructed example (99), and so could not be annotated, because a) gestures were not annotated, and b) they wouldn't constitute valid arguments for the relations.[28]

[25] The annotation was done by the author. Ideally, other, possibly even linguistically naive annotators should have been used as well; however, due to time and resource constraints, this was not done. We only want to show tendencies here, however, and so we think it is not a very big problem that we cannot give numbers about inter-annotator agreement.

[26] Note that the numbers don't need to (and in fact don't) add up to the number of NSUs found in the corpora, since some utterances might instantiate more than one type.

[27] Although we only annotated two files, these consisted of several conversations each.

[28] It should in principle be possible to extend SDRT with an account of non-linguistic acts, but we didn't want to make this step here.

| BNC | |
|---|---|
| *Backchn* | 218 |
| *QAP* | 204 |
| *Agree* | 103 |
| $Elab_{pq}$ | 101 |
| *Pardon* | 62 |
| $Elab_{pp}$ | 49 |
| *Comm* | 41 |
| `other` | 38 |
| *QCont* | 30 |
| $Expl_q$ | 30 |
| $Elab_{qq}$ | 21 |
| *Cont* | 16 |
| *Corr* | 16 |
| $Elab_{qp}$ | 12 |
| *P-Elab* | 11 |
| *IQAP* | 7 |
| $QAP_q$ | 7 |
| *Ack* | 7 |
| *P-Corr* | 6 |
| *Expl* | 4 |
| $Expl_q^*$ | 3 |
| $Comm_q$ | 2 |
| *Contr* | 2 |
| *Q-Alt* | 1 |
| $Expl^*$ | 1 |
| *Q-Elab* | 1 |

| VM | |
|---|---|
| *Greet* | 81 |
| *Ack* | 71 |
| *QAP* | 54 |
| *P-Elab* | 52 |
| $Elab_{pp}$ | 44 |
| *Cont* | 38 |
| $Elab_{qp}$ | 35 |
| `other` | 28 |
| *Q-Elab* | 28 |
| $Ack_q$ | 22 |
| *IQAP* | 16 |
| *Q-Alt* | 15 |
| *Comm* | 6 |
| *Corr* | 6 |
| *Contr* | 4 |
| $Elab_{pq}$ | 4 |
| *Res* | 4 |
| $Elab_{qq}$ | 2 |
| *Agree* | 1 |
| *P-Corr* | 1 |
| $QAP_q$ | 1 |

Table 2.5: Distribution of types of non-sentential speech acts

| BNC | | |
|---|---|---|
| items analysed | 5087 | |
| Fragments | 562 | (= 11.0%) |
| classfd. | 524 | (= 93.2% of fragments) |
| `other` | 38 | (= 6.8% of fragments) |
| VM | | |
| items analysed | 4037 | |
| Fragments | 369 | (= 9.1%) |
| classfd. | 341 | (= 92.4% of fragments) |
| `other` | 28 | (= 7.6% of fragments) |

Table 2.6: Results for Annotation, only fragments proper

(99) A: May I?
[B shakes her head]
A: Why not?

A third group of `other`-items were those where we weren't able to understand what was going on in the dialogue, due to a lack of background knowledge (we will discuss the relative strong requirement for such knowledge in a minute), or because dialogue participants were talking over each other, or even because the dialogue was indeed inconsistent briefly. Finally, both in the VM/REDWOODS and in the BNC data there are items marked `other` which seem to serve 'non-message-type' functions for which we didn't have a class; this indicates that our taxonomy might be incomplete in this direction, which however is not the direction we will be moving in in later chapters. As a first conclusion, it seems that this number of unclassified items is unavoidable and not a sign of a systematic deficiency of the taxonomy.

But does this mean our taxonomy is ideal? How easy was it to annotate the fragments that were classified? Are all classes equally good? Are the types for which we found only a few instances simply rare, or are they difficult to annotate? We can give at least a partial answer to these questions. During annotation, we also noted our confidence in the decision, and interestingly, this varies considerably with the type. For example, *QAP*s were always easy to annotate (but see below), whereas the distinction between *Comm(ent)* and *Agree* proved to be rather difficult to annotate, or between *Q-Elab* and the variants of *Elab*. Also, it turned out that the difference between *Expl* and $Expl_q$ on the one hand and the respective 'meta-talk' versions on the other was difficult to annotate. A further complicating factor are indirect speech acts. In the VM/REDWOODS corpus for example, exchanges like (100) are quite frequent.

(100) A: Can we meet next week?
B: OK.

In this exchange B replies with an *Acknowledgement* to a request which was expressed indirectly with a question. In this particular example, the indirect speech act is easy to detect, since the form 'can you VP' is strongly conventionalised (Searle 1975, Morgan 1978, Asher & Lascarides 2001), but other cases were more problematic, and so made the decision for *QAP*, *Ack*, or both less confident.

So, what do we conclude from this? It seems that a relatively high amount of information about the discourse is needed during annotation, information that might be very cost-intensive to get (e.g., to get at the difference between $Expl_q$ and $Expl^*_q$, one would need to model the cognitive state of the speaker in a fairly detailed way). Hence, for *descriptive* purposes it might make sense to collapse some of the classes. From an *analytic* standpoint, however, it can still be useful to insist on the differences. Moreover, a computer system that can produce this level of interaction will have to keep track of information at this level in any case, and so ideally should be able to deal with this (we will see a very simple attempt in this direction in Chapter 9). Finally, we should stress that we are not putting forward this taxonomy as an annotation schema to compete with for example (Core & Allen 1997) and similar attempts. Such schemes often use much broader classes, with only an informally given semantics. Our starting point is different, since we want to use a taxonomy of relations whose truth-conditional semantics are well-understood (see (Asher & Lascarides 2003) and later chapters of this thesis for more details). With this in mind, we are quite satisfied that the results lend some empirical credence to our taxonomy.

#### 2.2.4.3 Comparison to previous studies

To date, there have been only few empirical studies of non-sentential utterances in dialogues with which we could compare our results. A paper from the 1980s, (Thompson 1980), found that in a natural language interface to a database "approximately 10% of user utterances are elliptical." The classification used in this paper includes other kinds of ellipsis like VPE, but it is interesting to note that the result is roughly in the same order of magnitude as ours.

More relevant is a recent study by Fernández & Ginzburg (2002), where a taxonomy of non-sentential utterances is also presented (we will discuss the taxonomy separately below in Section 2.4 and concentrate on the empirical part of that paper here). The corpus used by the authors was the BNC, of which they selected a random excerpt of altogether 7542 items—fewer than our 9124 items, but from a wider range of files, spanning more genres than just free conversation and task-oriented dialogue (e.g. formal interviews and group meetings). They "identified and classified 841 NSUs, which make up 11.15% of the total number of sentences in the searched transcripts." (Fernández & Ginzburg 2002, p.19); a number which again seems to be very close to our findings for fragments proper. However, of the classified items they claim only 0.95% had to be classified as `other`, which is a significantly lower portion than with our taxonomy. We will now discuss possible reasons for this difference.

First of all, we think there are some problems with their approach which make a direct comparison difficult. For one, they are not very explicit on the criteria they used for annotation: they do not mention the problem of overlaps and how they influenced their data, i.e. whether they regarded constituents that were separated from the original utterance as shown above in (96-a) as NSUs or not. In general, they do not offer in the paper a clear definition of what they classified in the first place, beyond the statement that "dialogue is full of intuitively complete utterances that are not sentential"; and they do not make a distinction similar to ours between 'message-type NSUs' and 'non-message-type NSUs'. From the fact that they have a class *Acknowledgements* which include "utterances like 'okay', 'yes' and 'mm' that signal that the previous utterance was understood" (p.17) one can conclude that they used what we called the wide-definition of fragmenthood above.[29] If that is the case, however, then their numbers are significantly *lower* than ours: remember that if we include what we called 'non-message-type NSUs, we got 19.4% for the same corpus. This can presumably be partially explained by the fact that in some of the files they included (namely in the formal interviews) such kinds of utterances were not transcribed,[30] but it is difficult to say whether this accounts for the whole difference.

This vagueness about the data makes it difficult to interpret their 99.5% coverage. Perhaps more importantly, however, their classes are rather surface-oriented and hence not as much contextual information is needed for the classification task as with our taxonomy. This is an advantage for annotation, and could help to explain why they achieved greater coverage in the classification task than we did. But a surface-oriented approach to defining speech acts is a hindrance for formalising the resolution of fragments, as we will argue below in Section 2.4.

#### 2.2.4.4 Conclusions

We believe that this study has shown that our taxonomy offers a satisfactory coverage of the data, and thus forms a good starting point from which a formalisation can proceed without the danger of loosing sight of the actual data.

## 2.3 *resolution-via-identity* and *resolution-via-inference*

The discussion in the previous section has shown that the kind of resolution needed to get at the intended meaning of fragments—*res-via-id* or *res-via-inf*—is not dependent on the type of speech act that

[29] It is difficult to say whether those items we labelled *Pardon*, i.e. NSUs like "Sorry?", fall under one of their classes (which we will discuss below in Section 2.4); if not, then it seems rather unmotivated why they include backchannels and not other items of this kind.

[30] Which suggests that it would have been useful for them to make a similar classification and give separate numbers for those types of NSUs that were present in all files.

has been performed with the fragment, or at least not in the sense that the speech act *type* uniquely determines it; we will see later that the *semantics* of the speech act does indeed have an influence.

What we deliberately haven't said above is what kind of material from context and fragment has to be combined to arrive at the resolution of the fragment. For *resolution-via-inference* cases, this should be relatively clear: since the material coming from the fragment has to form the basis of an inferential process that might have to take into account extra-linguistic knowledge sources like world knowledge or plans, it seems that both the starting point and the result of this process will clearly be a semantic representation. Or to put it differently, it seems that the process is best defined on logical formulae, rather than on syntactic structures.[31] The question how *res-via-id* fragments are resolved, i.e. what is actually supposed to be 'identical' in these cases, however, is less easily answered, and will occupy us for much of the remainder of the thesis. In the rest of this section, we will give a preview of that discussion; we will go into more detail in Chapter 4, where we discuss Morgan's (1973) purely syntactic approach, and in Chapter 5, where we discuss a grammar-based approach that strives to mix a semantic with a syntactic approach.

The examples of *res-via-id* fragments we have seen so far seem to be compatible with both syntactic and semantic approaches. A syntactic approach for example could work by copying over bits of the question to the fragmental answer, to (re-)construct a syntactic structure for the resolved fragment which is then interpreted.[32] In such an approach, the material that is supposed to be identical in $\alpha$ and $\beta$ is syntactic structure. A semantic approach on the other hand could work by combining the appropriate semantic material from $\alpha$ and $\beta$ to result in a semantic representation of the intended meaning.

As we will see in our discussion on Morgan's (1973) approach, however, there are examples that seems to disfavour a syntactic resolution. In example (101), for instance, the syntactic structure corresponding to "he$_i$ himself likes John$_i$" is not well-formed, whereas the fragment (101-b) is.[33]

(101) a. A: Who likes Peter?
b. B: He himself.
*He$_i$ himself likes Peter$_i$.

Another example of this kind is the following, where similarly a binding constraint is violated:

(102) A: Who does Sandy think John likes.
B: Herself.
*Sandy$_i$ thinks John likes herself$_i$.

[31] Note that we talk of processes here just for convenience; the actual definition of resolution we give in Chapter 8 will be specified declaratively.

[32] We will discuss the distinction between syntactic and semantic approaches to ellipsis in general in Chapter 3 and to fragments in particular in Chapter 4.

[33] Similar examples are discussed in (Barton 1990) and (Ginzburg 1999*b*).

Nevertheless, this cannot be taken as evidence that a purely semantic approach would do better, as (Ginzburg 1999*b*) pointed out.[34] Why this is so can be shown in English with examples like (103) and (104). The preposition in (103), being a verb particle, is normally seen to be fulfilling only a syntactic role, and hence is not represented in the logical form (cf. e.g. (Pollard & Sag 1994, p.25)). This example, however, suggests that the questions license the use of the preposition in the answer: in (103-a), only the same preposition as in the question is allowed; one cannot use the (similarly semantically null) verb particle “of”. In (103-b) no semantically null preposition (i.e. preposition as a verb particle) is felicitous at all. If the presence of these verb particles, which are semantically empty, is not recorded in the logical form of either question or answer, then we can conclude from this that at least this kind of *syntactic* information has to be preserved in the representation of the context.

(103) a. A: On whom can Mary always rely?
B: On Peter. /#Of Peter.
b. A: Who does Mary like?
B: # On Peter.

Similarly, the grammatical idiosyncrasy of the verbs “make” and “do” in example (104) (they require an infinite verb or a base-form verb in their argument VP, respectively) seems to be ‘transmitted’ to the short-answer.[35] A semantic explanation of this fact would have to claim that the VPs in these short-answers are of different semantic type; a claim that presumably would lead to a fine-grainedness in semantic types that could cause problems elsewhere in the grammar.

(104) a. A: What did he make you do? B: Sing. / #To sing.
b. A: What did he force you to do? B: To sing.

Examples from case-rich languages like German can reinforce the observation even more. In the following example, it seems to be the case that the (quite idiosyncratic) case-assignment by the verb to its arguments is also effective across the sentence boundary.

(105) A: Wem$_{dat}$ hast Du geschmeichelt? B: [Dem Mann]$_{dat}$.
*A: Who did you flatter?* *B: The man.*
A: Wen$_{acc}$ hast Du gelobt? B: [Den Mann]$_{acc}$.
*A: Who did you praise?* *B: The man.*

(Ginzburg 1999*b*) and subsequent papers only investigate this syntactic influence for short answers to

[34](Ginzburg 1999*b*) does not discuss ‘why’-questions, and does not make a distinction comparable to ours between *resolution-via-inference* and *resolution-via-identity*. Indeed, as we will explain in 5, his approach seems for principled reasons to be only applicable to the latter.

[35]We will discuss later the fact that the complementizer in (104-b) does actually seem to be optional.

complement questions and certain types of fragmental questions, but as the following examples show, the situation seems to be similar for our whole class of *res-via-id* fragments.

(106) a. He made him sing.
Sing a whole aria(, to be precise.) / #to sing a whole aria.
b. He forced him to sing.
To sing a whole aria(, to be precise.)

(107) A: [Der Lehrer]$_{nom}$ gab [dem Schüler]$_{dat}$ [den Hammer]$_{acc}$.
*The teacher gave the student the hammer.*
a. B: [Der Mathelehrer]$_{nom}$(, um genau zu sein).
*The maths teacher(, to be precise).*
b. B: [Dem Klassensprecher]$_{dat}$(, um genau zu sein).
*The head of class(, to be precise).*
c. B: [Den neuen $\emptyset$]$_{acc}$(, um genau zu sein).
*The new (one)(, to be precise).*

Ginzburg & Sag (2001) explain this data for short-answers by stipulating a requirement of an identity of category between *wh*-phrase and the phrase constituting the fragment (the exact technical details of how they do this are discussed in detail in Chapter 5.).[36] In their approach, the pattern observed in (103-a) for example would be explained by the requirement of categorial identity between "on whom" / "of whom" (a PP[on] / PP[of], respectively) and "on Peter" / "of Peter" (similarly a PP[on] / PP[of]).[37] (103-b) would be ruled out, since the *wh*-phrase is an NP, and the short-answer a PP. Note that their approach does not require full syntactic reconstruction, but is restricted to exactly this amount of syntactic information.

We highlight some problems with this approach here. First, it seems that words from certain classes can be 'dropped' in short-answers. For example, questions where the *wh*-phrase is a PP, like those in (103) above and in (108) below, can nevertheless be answered felicitously with NP-fragments, as shown in (108-a).

(108) a. A: On whom did Mary rely?
B: Sandy. / On Sandy.
b. A: On what did you put the book?
B: The table. / On the table.

Similarly, the complementizer "to" from the VP-answer in (104-b) above seems to be dispensable, as

[36] They only offer an approach for some of the speech-acts we have classified above, but presumably their approach could be extended to deal with the observations for other *res-via-id* cases of fragments in a similar way by stipulating categorial identity.
[37] They do not explicitly deal with prepositional phrases, but this is what follows from the examples given in their paper.

(109) shows.[38]

(109) A: What did he force you to do?
B: Sing. / To sing.

Moreover, as we have seen above when we looked at elaborations, *res-via-id* fragments do not always *replace* an element in α (where this element would be the antecedent for Ginzburg's parallelism constraint). The examples in (36) above, two of which are repeated here as (110), have shown that also non-realised optional arguments of elements of α can be elaborated with fragmental utterances.

(110) a. A: Peter was reading when I saw him.
A book about Montague semantics*(, to be precise)*.
b. A: Die Verleihung war spektakulär.
[Des Oskars]$_{gen}$, meine ich.
*(The handing over of the award was spectacular. Of the Oscar, I mean.)*

Examples like this suggest that it is not directly the syntactic category of an element of α that is important, but rather the syntactic constraint of an element of α on its complements, be that a verb, as in (110-a) above, or a noun, as in (110). Note also that arguably we even need the requirements for a related verb, namely the transitive rather than the intransitive version of 'read', in (110-a).

As we said, we will return to the question of how this data that apparently pulls in different directions—contra syntactic, but also contra semantic reconstruction—should best be analysed. For now, this closes our introduction of our taxonomy, and we conclude by comparing it to ones that can be found in the literature.

## 2.4 Comparison with other Taxonomies

A number of taxonomies or classifications of non-sentential utterances already exist. In this section, we will present three of them, roughly in historical order. The first one was proposed by Ellen Barton (1990), the second one comes from the plan-based approach of (Carberry 1990), and the final taxonomy is based on Ginzburg's approach we have already mentioned, and is published in a paper by Fernández & Ginzburg (2002).

**Barton (1990):** Of these three classification attempts, Ellen Barton's is the only one that makes a distinction that is similar to our *res-via-id*/*res-via-inf*-dimension. The distinction she makes is based

[38] Note that all fragments in these examples are complements, and so this more relaxed 'parallelism' can't be due to differences between complements and adjuncts. (The latter are exempted from parallelism in Ginzburg & Sag's (2001) approach.)

on a difference in the "type of inference which is based on a different kind of contextual information" (Barton 1990, p. xvii). She illustrates this difference with the following example, where the first fragment is interpreted in her model in the "submodule of linguistic context", whereas the second is interpreted in the "submodule of conversational context".[39]

(111) a. A: What stops the White House staff from visiting House Speaker Tip O'Neill in his congressional office?
B: [An] old grudge.

b. A: The White House staff doesn't visit Tip O'Neill in his congressional office
B: [An] old grudge.

This distinction is very similar to ours between *res-via-id*, and *res-via-inf* (as which we would classify (111-a) and (111-b), respectively). Moreover, she makes similar observations regarding syntactic influences from the context on the fragment in the former case. What her approach doesn't offer, however, is a classification that goes beyond this major distinction, or that would offer a basis to explain it. We will develop such an approach later in the thesis. Also, her approach is only very roughly formalised, and so it is difficult to test predictions made by it (see Section 4.4.1 for details).

**Carberry (1990):** The main aim of Carberry's (1990) approach and the individuating criterion of her taxonomy is described in the following quote: "Understanding the intent behind elliptical fragments requires that the speaker's discourse goals be recognised. [. . . ] I have identified fifteen discourse goals that occur during information-seeking dialogues and that may be accomplished by means of elliptical fragments." (Carberry 1990, p.193). Following SDRT our aim is more modest: we want to describe the intended meaning of fragments, not necessarily all intentions behind uttering it. And we believe that this results in a much simpler approach, both in theoretical and in computational terms. This difference in aim allows us to generalise over some distinctions Carberry makes in her taxonomy. Moreover, by analysing certain constructions in more detail, we can even further generalise her classes.

The first effect is shown with the examples in (112), all of which we analyse as different types of *Elaboration*.

(112) a. A: What courses would you like to take?
B: For credit?

b. A: I want to get a degree.
CS major.

[39] A similar distinction of ellipsis in general can be found in (Rath 1979), who distinguishes between "Konstruktionsübernahme" (construction-adoption) and "Eigenkonstruktion" (self-constructed).

c. A: What is Dr. Smith teaching this fall?
   B: CS360.
   A: CS360?
   A′: The course in architecture?

d. A: Who's teaching CS360? The course in architecture.

In Carberry's taxonomy on the other hand, the fragments are all classified as belonging to completely different classes, which are in turn: *seek-clarify-question* ("IS [= information seeker] requests information relevant to clarifying a question posed by IP [= information provider]."), *provide-for-assimilation* ("IS provides information pertinent to constructing his underlying task-related plan."), *seek-identify* ("IS is unable to satisfactorily identify the referent of an item in IP's utterance and requests help from IP in doing so.") and *identify* ("IS attempts to identify an entity in his own utterance.").

The examples in (113) show that our analysis has the advantage of breaking down turns into smaller units, so that more general classes can be formed. Carberry classifies the fragments in (113) as a *answer-question-suggest-alternative* ("IS answers a yes/no-question negatively, providing a description of a desirable alternative.") and *answer-question-with-explanation* ("IS answers a yes/no-question with an explanation of the answer.") respectively. Since we analyse the negative answer "no" as an anaphor for the negation of the proposition contained in the polar question to which it is an answer, we can analyse these examples with rhetorical relations that are independently motivated, namely *Correction* or *Contrast* for (113-a) and *Explanation* for (113-b).

(113) a. A: Would you like to take CS360?
   B: No, CS470.
b. A: Do you want to take CS865?
   B: No, too late at night.

We will further discuss Carberry's approach below in Section 4.4.2, and only note here that our taxonomy captures all her examples, albeit with fewer classes (and hence fewer distinctions).

**Fernández *et al.* (2002):** As mentioned in the previous section, (Fernández & Ginzburg 2002) offer a taxonomy roughly based on Ginzburg's approach to non-sentential utterances. Theirs is the only paper of those discussed here where data about coverage are given; they claim that they can classify 99.05% of the fragments in the corpora they annotated. As already alluded to above, however, we think this result is difficult to compare to ours. We mentioned above that it is not quite clear what the data-basis for their annotation was; another factor that makes a direct comparison difficult is that the definitions of the classes used in the study (at least as they are described in (Fernández & Ginzburg 2002)) do not seem to be entirely consistent. Some of their classes appear to be defined according to the discourse function of the fragment (like our classes are, although they do not use relations), while others are based just on the form of the fragment; they also have classes that seem to be individuated by a mixture of these criteria, together with reasoning about the intentions of the interlocutors. In the following we will show that our taxonomy is both more consistent, being based on the sole criterion of the *use* made of fragments, and also more fine-grained. We first list the classes in (Fernández & Ginzburg 2002) for which there is a direct counterpart in our taxonomy. These are mostly fragments that are classified according to use by the authors as well.

(114) Short-Answer ≈ *QAP*
Correction ≈ *Correction*
Acknowledgement ≈ *Acknowledgement*
Clarification Ellipsis ≈ (one use of) $Elaboration_{pq}$

The class they label '*sluicing*' is an example of one defined by a mixture of form, function and intentions. Part of the definition is "sluices are bare question denoting *wh*-phrases" (all quotes here are from (Fernández & Ginzburg 2002, p.16)). However, as the authors make clear later in the paper, only one type of such phrases is actually supposed to belong to this class, namely that of fragments which "involve a request for additional information beyond what the speaker of the previous utterance thought was required." Note that this definition is in terms of intentions and beliefs of the interlocutors about their wider discourse goals ("what the speaker thought was required"). The example they give (from the BNC), presented here as (115), suggests that the fragments in this class are instances of what we label $Elab_{mq}$ (where $m$ is either $p$ or $q$). Our class however is only defined in terms of speech act related goals, which is a more restricted and conventionalised kind of goal than the general discourse goal.

(115) A: Can I have some toast please?
B: Which sort?

Two more examples that show that our classes are both more general and more fine-grained: First, they have a class called "bare modifier phrase" (and the name already indicates that this class is individuated by the form of the fragment), for which they give (116) as an example. In our taxonomy, however, this is just one way of expressing elaborations, and so our taxonomy captures generalisations that theirs doesn't capture.[40]

(116) A: They got men and women in the same dormitories.
B: With the same showers!

[40] Note that they ultimately have the same goal as we have, namely to specify how fragments are resolved, and so a classification just of different possible *surface-forms* of fragments is of little value to them as well.

Second, they describe a class they call "fragments introduced by connectives". As our examples (32-c) and (87) have shown (repeated here in (117), such fragments can serve a number of functions, and so our classification is more fine-grained.

(117) a. A: I went to the cinema with Peter.
And Sandy.
b. A: He drove to Italy. And then to Spain.

Their conclusion that "*with the context as given*, the principles by means of which NSU content is resolved do not involve complex domain sensitive reasoning." is not very strong, because the hard bit is to determine what exactly the context is. Computing the *question under discussion* (QUD), the main instrument for resolution in their approach (see below Chapter 5), will in all but the most trivial cases (*QAP*s) be a matter of pragmatic reasoning. The actual way of combining material once that is done is irrelevant to the question of how complex the reasoning in general is; we will expand on this point in Chapter 5.

## 2.5 Summary

In this chapter we have introduced a classification of fragments according to two dimensions: their rhetorical relation to an element of the context, and the relation of their resolved representations to that element. We have provided informal definitions for these rhetorical relations (we refer back to Tables 2.1–2.3 for an overview of these speech act types and their definitions) and we have investigated to what extent this classification captures data from corpora. We have then discussed certain properties of *res-via-id* fragments, namely that it seems that neither a fully syntactic nor a fully semantic resolution procedure can deal with them. This question of the right level of representation for reconstruction is what connects this chapter to the next. There we will look at the literature about a related phenomenon, namely VP-ellipsis, and we will see that this question guides the development of approaches in that field as well. We will survey the approaches and tools developed in that field, to see how far they can be helpful for us as well. In the Chapters 4 and 5 we will return to our main phenomenon, fragments, and see that the previous approaches can be roughly organised along this line of investigation as well.

But before we move on to the next chapter, here is a very concise list of desiderata we can compile from the long discussion in the present chapter.

An adequate approach to the interpretation of fragments must

- reflect the importance of rhetorical relations;
- model the differences between *res-via-inf* and *res-via-id* fragments; and, ideally also

- *explain* those differences by reducing them to more general principles.

This is what we will attempt in the later chapters when we develop our own approach. Now, however, we will look at the related phenomenon VP-ellipsis, to see whether techniques developed to tackle this problem can help us with ours.

# Chapter 3

# Other cohesive forms: VPE

In this chapter we look at the relation of the phenomenon of NSUs to the general class of ellipsis-phenomena, and briefly survey the literature on one particularly well-studied type of ellipsis, namely VP-ellipsis (VPE). As we will see, in this literature the main criterion for classifying approaches is the level of linguistic structure at which they place the resolution process; this is a question that we will have to answer later in the thesis for fragments as well. We undertake this review of the literature on VPE with an eye on whether similar data (favouring one or the other level) can be found for fragments, and more importantly whether methods similar to the ones developed for this problem can be applied to our phenomenon. Hence we are less interested in the details of the individual approaches here than rather in the general line of argumentation for or against certain kinds of approaches; this will provide a blueprint for the discussion of approaches to fragments in the next chapter. As one important result, this review gives us independent motivation for keeping syntactic information available for inter-sentential processes.

## 3.1 Introduction

The term 'ellipsis' is typically defined something like this: "the omission of one or more elements from a construction, especially when they are supplied by the context" (Matthews 1997). Clearly, fragments as defined in the previous chapter fall under this definition. In this chapter we will look at the literature on other, better researched kinds of ellipsis—a list of examples together with (informal) definitions is given in (118)—[1] to see if the approaches can be made useful for our problem.

(118) a. **VP-Ellipsis** (the head VP in the target is elided and replaced with a placeholder AUX)
*Noam met Ivan. But Ferdinand* didn't
*Noam read every book that Ivan* did.
b. **gapping** (the head verb and possibly some of its complements are elided, where source and target are coordinated)
*Noam wears glasses, and* Ivan contact lenses.
c. **pseudo-gapping** (the head verb is replaced by an auxiliary)
*Noam wears glasses, and Ivan* does contact lenses.
d. **stripping** (only one complement left in the target, again in coordinated structures)
*Noam wears glasses,* but not Ivan.
e. **sluicing** (all but the questioned elements are elided in the target)
*Noam quarrelled with Ivan, but he doesn't know* why.

We will mainly look at the literature on VP-ellipsis (VPE) here, where there is a lively debate about the nature of the resolution process.[2] Or in the words of Hobbs & Kehler (1997, p.394):

> *"The area is a tangled thicket of examples in which readings are mysteriously missing and small changes reverse judgements. It is a prime example of a phenomenon at the boundary between syntax and pragmatics."*

Approaches to VPE can be broadly classified along two dimensions. While all approaches to VPE agree that material from the context is needed to interpret the clause containing the ellipsis, they differ in what they assume the nature of that material to be. This is the first dimension of classification:

- *Syntactic approaches* postulate the existence at some stage of a syntactic representation of the target in which the information that is 'missing' on the surface is present; this syntactic representation is interpreted in the standard fashion. A prediction of this kind of approach is that the syntactic well-formedness of the full-sentence version influences the well-formedness of the elided version.

[1] In the list, and henceforth in this chapter, the clause containing the antecedent of the elided element is called "source clause" or simply "source", and the ellipsis site "target (clause)".

[2] The other ellipsis phenomena are less controversial and usually regarded as syntactic. See for example for gapping (Ross 1970, Steedman 1990) *inter alia*. Sluicing is another interesting case where semantic /pragmatic and syntactic approaches exists, but for reasons of space we will focus on VPE here.

- In a *semantic approach*, it is semantic material that is derived in some way from the source and is used to construct a semantic representation for the target. In the case of VPE, this amounts to recovering a property from the source that is then applied to the target. Unlike a syntactic approach, such an approach allows the 'missing' material to be implicit in the context, and it also does not predict any further syntactic constraints on the form of the VPE.

- *Discourse-based approaches*, which have become more prominent lately, put additional focus on the discourse relation between target and source. Some of these approaches fall somewhere in the middle between the two previously listed classes in that they mix elements of both; these will be particularly interesting for us.

The interaction of ellipsis with other phenomena can introduce ambiguity; two of these phenomena, which will show up again for fragments, are quantifier scope and so-called strict and sloppy readings of pronouns. The source in example (119-a) has a reading where "a book" has widest scope, and if this reading is chosen, then the target must be interpreted in the same way; i.e., no 'mixed' readings are possible where the target is interpreted with a different scope order from the source. This also seems to be the case for fragments, as (119-b) indicates, where the reading of the question fixes that of the answer.

(119) a. Peter gave a book to every student.
John did, too.

b. A: Who gave a book to every student?
B: Peter.

Example (120) demonstrates the effect of the other ambiguity source mentioned above. The mini-discourse can be interpreted as a description of a happy family life (in which Joe loves his wife, and Peter also loves his own wife; this is the *sloppy* reading) or as that of a (possibly unhappy) love triangle in which Peter loves Joe's wife (the *strict* reading).

(120) a. $\text{Joe}_1$ loves $\text{his}_1$ wife.
b. $\text{Peter}_2$ does, too. *(= love* $\textit{his}_{1/2}$ *wife)*

Again, similar ambiguities can be constructed with fragments, as in the fragmental correction in (121):[3]

(121) A: Carl has talked to his personal trainer.
B: No, Carl's father.
(= Carl's father has talked to his own personal trainer,

[3]This example admittedly does not sound as natural as with a VPE, but is acceptable with a strong intonational focus on 'Carl' in the first utterance and 'father' in the second one, where A and B are arguing about who has talked to the personal trainer.

*or* Carl's father has talked to Carl's personal trainer.)

These ambiguities must be captured by the approaches to VPE. The question of where in the process of recovery the ambiguity is seen to arise marks the other dimension of categorisation:

- *Identity-of-relations-analyses* demand that the representation of the elided element is identical to the respective representation from the source. Any ambiguity must therefore be postulated to already be present in the source, even if the sentence on its own (eg. (120-a)) doesn't normally exhibit this ambiguity.[4]

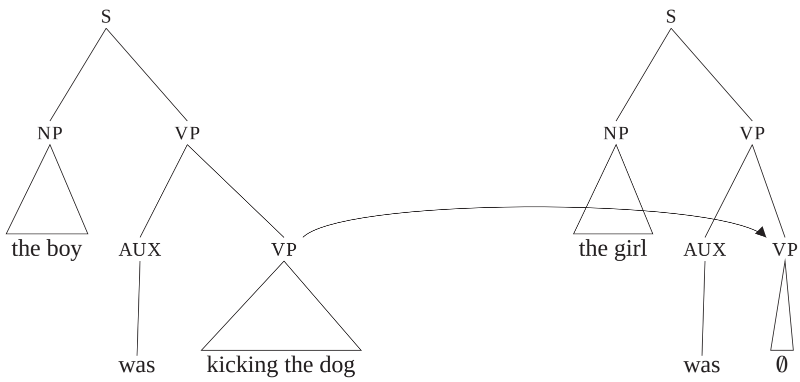
- *Non-identity-analyses* relax this condition. Here the operation can map one input representation to more than one output, if necessary. Therefore, the ambiguity is seen to arise during the recovery process and is not postulated to be in the source.

## 3.2 Syntactic Approaches

At least on first view, a syntactic approach to VPE seems promising. It seems that there is a strong *syntactic* parallelism between elements in the source and in the target, which allows the simple replacement of the VP that contains the "place-holder auxiliary" with *syntactic* material from the source, as shown schematically in (122-b).[5]

(122) a. The boy was kicking the dog. The girl was too.

b.




[4]Note that this assumption is only problematic for the strict/sloppy ambiguity explained above; in the case of scope ambiguities it is clear that the ambiguity is present in the source as well.

[5]The problem is presented here in the manner of speaking of a *reconstruction* approach, like for example (Williams 1977). A *deletion* approach sees the target as the transformational result of a full-sentence where elements were removed under certain conditions, cf. for example (Sag 1976). We abstract over these differences, since they meet in the claim that syntactic conditions on full-sentence paraphrases bear on the well-formedness of the clause containing the ellipsis.

There are a large number of approaches following this general line—a short selection of the most influential ones is (Sag 1976, Williams 1977, Fiengo & May 1994, Lappin & Shih 1996)—, differing for example in whether the process is construed as one of deletion or reconstruction or in the level of syntax they operate on, ie. syntactic logical form or surface semantics; the details are not important for our purposes here.

In any case, if interpretation of VPE does indeed proceed via syntactic trees, then the form of the target clause must be influenced by syntactic well-formedness conditions on the reconstructed form. Examples like the following are often used to support this claim:[6]

(123) a. The children asked to be squirted with the hose,
and so { they were *(squirted with the hose)*. / *we did *(squirt them with the hose)*. }
(Mismatch of voice.)

b. Hans schmeichelte jemandem, aber ich weiss nicht, { wem. / *wen. }
(Hans flattered someone-DAT, but I don't know { who-DAT. / *who-ACC. } )
(Violation of case agreement.)

c. John read everything which Bill believes he did.
*John read everything which Bill believes the claim he did.
(Subjacency violation.)

d. *Everyone else helped him$_i$, but he$_i$ didn't *(help* himself$_i$*)*
(Principle B violation.)

As we will see in the next chapter, a similar line of argument can be found in the literature on fragments, arguing that fragments can be shown to be subject to conditions on their well-formedness that can only be explained as syntactic. However, there as here, there are counterexamples to this claim, which we will discuss for VPE in the next section.

A consequence of syntactic approaches that is often overlooked is that even though VPE often occurs intra-sententially, it can equally happily occur inter-sententially, as in example (124). Since syntactic approaches place ellipsis resolution in the grammar, this entails that that grammar must be able to deal with units larger than sentences.

(124) A: No one visited me when I was ill.
B: That's not true. I did!

[6] After (Kehler 1993, Haïk 1987, Hardt 1997*c*). The (assumed) reason for non-well-formedness of the starred examples is given in parentheses.

## 3.3 Semantic Approaches

Examples like the following seem less readily explained with a syntactic approach, since here violations of certain syntactic conditions do seem to be allowed:[7]

(125) a. A: Do you think they'll like me?
B: Of course they will *(like* you*)*.
(Different pronoun required for same referent.)

b. I will hurt myself before he could *(hurt* me*)*.
(Again different pronoun required, here for syntactic reasons. Reconstruction with "myself" would lead to a Principle A violation.)

c. John got to Sue$_i$'s apartment before she$_i$ did *(get to* her$_i$ *apartment)*
(Reconstruction with proper name would lead to Principle C violation.)

d. A lot of this material can be presented in an informal fashion, and often I do *(present it that way)*.
(Mismatch of voice *is* felicitous here.)

e. [There is a problem with the security doors today.] If it's just a case of going out for a bun and coming back, I suggest you don't *(go out for a bun and come back)*.
(Antecedent is not head-VP.)

f. Ivan and Barbara want to go out together but Barbara can't *(go out with Ivan)*, because her father disapproves of Ivan.
(Antecedent not overtly expressed.)

g. John got sick. He went to the hospital.
Fred did { *(go to the hospital)* / *(get sick and go to the hospital)* / **(get sick) [if not followed by contrast]* } , too.
(Ambiguity in scope of antecedent, could be split antecedent.)

In semantic approaches, the target is not subject to syntactic constraints in the same way, and so these approaches do not have problems with examples of this kind. Possibly the most influential semantic approach—and also the clearest and simplest representative of this class—is that of (Dalrymple et al. 1991) (henceforth DSP); it is worthwhile here to take a closer look at it.

[7]Example (125-e) is from an email sent by a secretary of the University of Edinburgh, the others are taken from (Sag 1976, Hardt 1997*c*, Dalrymple, Shieber & Pereira 1991, Webber 1978, Asher 1993), respectively. The reasons why these examples pose a problem for syntactic approaches are given in parentheses.

In this account, ellipsis resolution amounts to recovering a property from the context. The property that is missing in the representation of the target clause is extracted from the context by solving, via higher order unification, an equation that reflects the parallelism between source and target. The approach is best explained with an example. (126) shows a typical instance of VPE; semantic representations for the source and target of this example are shown in (127)-a) and -b) respectively. To resolve the meaning of the target in this approach, the property $P$ must be specified. This is done by equation d), which states that the property must be such that it yields the source when applied to the parallel element *in the source*; ie., it is a property that is yielded by abstracting over the parallel element in the source. In our case, the equation has two solutions for $P$, which applied to the target result in the two desired readings resulting from the strict/sloppy ambiguity.[8] As this example shows, this ambiguity is modelled as an effect of the resolution mechanism and does not have to be postulated to exist at the source.

(126) a. Joe loves his wife, and ...
b. ... Peter does, too.

(127)

| | | | |
|---|---|---|---|
| (127) | a. | Representation source clause: | $love(joe, wife\_of(joe))$ |
| | b. | Representation target clause: | $P(peter)$ |
| | c. | Parallel elements: | $peter \leftrightarrow joe$ |
| | d. | Equation: | $P(joe) = love(joe, wife\_of(joe))$ |
| | e. | Solutions: | $P \mapsto \lambda x.love(x, wife\_of(joe))$ |
| | | | $P \mapsto \lambda x.love(x, wife\_of(x))$ |
| | f. | Substitution of e) in b), | $love(peter, wife\_of(joe))$ |
| | | β-reduced: | $love(peter, wife\_of(peter))$ |

This approach is quite flexible in what semantic material it can retrieve from the context, provided that the parallelism is set up correctly. If for instance in the representation of (128-a) the polarity of source and target is marked as parallel in the equation, as shown in (128-b), using the non-standard operators *pos* and *neg*, the approach provides the desired solution.[9]

(128) a. Dan didn't leave, but George did.
b. $dan \leftrightarrow george, \quad neg \leftrightarrow pos$
$P(dan, neg) = neg(left(dan))$
$P = \lambda x.\lambda S.S(left(x))$

Similarly, the phenomenon of 'sloppy tense' as exhibited by (129) can be handled; we forgo showing details here.

[8]These solutions can be computed automatically using Huet's (1975) algorithm for higher order unification. However, (Dalrymple et al. 1991) are at pains to stress that their approach is declarative and not procedural.

[9]The example is taken from (Dalrymple et al. 1991, p.22).

(129) You thought I was crazy. You probably still do. { (think I *was* crazy) / (think I *am* crazy) }

Clearly, correctly establishing the parallelism is crucial for this approach to work. Nevertheless, this task is factored out. The authors regard this as an advantage, because as they say a theory of parallelism is motivated independently, and secondly having this theory separated allows it to use syntactic as well as semantic / pragmatic methods to determine the parallelism. There are indeed several later approaches that try to formalise this process; some explicitly building on DSP, like (Gardent & Kohlhase 1997, Gardent 1999), others using slightly different formalisations as for example (Prüst, Scha & van den Berg 1994, Grover, Brew, Manandhar, Moens & Schöter 1995). These approaches all rely in some form on an ontology specifying sub-class relations between entities in the domain to compute what clauses have in common—the ‘common denominator’ in (Prüst et al. 1994), the ‘generalisation’ in (Grover et al. 1995). We will not describe these approaches in more detail, but note that we will return to them later when we describe the resolution of fragments.

Other semantic approaches that we have to at least mention are those that transfer ideas from DSP into a dynamic semantic setting (where meaning is defined as a relation between contexts; cf. the short introduction below in Section 8.1): (Klein 1987) is an early example; also (Bos 1994) falls under this rubric, which combines an approach to presupposition with the resolution of VPE; finally there is (Hardt 1997*a*).

A last semantic approach proper which we should mention here is (Egg, Niehren, Ruhrberg & Xu 1998). This approach offers an account of VPE on the level of *underspecified semantics* (where logical languages are used to *describe* formulae of the meaning language, cf. the introduction below in Section 6.3), a method which we will also use in our approach to fragments. The approach also takes the parallelism as given, and uses a special constraint to specify a parallelism-relation between two descriptions (viz. that of the source and the target) which constrains these descriptions in such a way that only the desired readings are described. We will briefly return to this approach when we describe our method for resolving *res-via-id*-fragments.

Modelling the establishment of parallelism is the first step towards what we will call ‘discourse-based approaches’, where the question of how the source can actually be found, and what influences this search, is tackled.

## 3.4 Discourse-Based Approaches

Example (130) demonstrates that there can be intervening material between ellipsis site and antecedent, and also, if there is multiple ellipsis, a rather complex resolution pattern.[10] The approaches described in

[10] The example is due to (Klein & Stainton-Ellis 1989).

the following introduce *discourse structure* to deal with such problems. Their common starting point is that VPE resolution is in some way informed or even determined by the process of establishing discourse cohesion.

(130) A: He's just [appointed Hawkeye chief surgeon over me]$_1$!
B: He can't $\emptyset_1$. Doesn't he [know that's [against regulations]$_3$]$_2$?
A: He did $\emptyset_1$, it is $\emptyset_3$, and he does $\emptyset_2$.

Asher (1993) analyses, in the context of SDRT, VPE as a subcase of concept anaphora, ie. as consisting of a relation between a concept discourse referent (introduced by the ellipsis) and a predicative DRS (i.e. a $\lambda$-abstracted DRS). This DRS can either be derived directly from a VP, or it can be constructed via reasoning. This approach tries to model not only the resolution of VPE but also to account for preferences for certain readings in certain contexts. It does this by driving the resolution of VPE by the need to maximise the quality of the discourse relation *Parallelism* (or *Contrast*, in some cases). This can explain the difference in (131), if we assume that there is a greater contrast between source and target in the first indicated resolution, and a better parallelism in the second.

(131) Kim said Joe hit her. { But Sam did *(hit her)*. / Sam did *(say that)*, too }

Moreover, the approach can explain the pattern observed above for (125-g), by allowing the relation *Parallel* to connect the target either to the previous clause or to a segment containing the two previous clauses: i.e., by making use of the structure of the preceding discourse.

The approach by (Prüst et al. 1994) also falls under the rubric 'discourse based'.[11] The authors similarly claim that "an adequate account of the structure of discourse and its semantics (taking into account parallelism effects) yields VPA [= VPE] resolution as a side effect" (p.266), and they also assume that generally parallelism is maximised in discourse. Additionally, the account is also not fully semantic in that the structure of the semantic representation is needed to establish parallelism and is, unlike in Montagovian approaches, not in principle dispensable. They call their representations "syntactic / semantic structures", but the 'syntax' in that name is that of the semantic representations.

(Kehler 1994, Kehler 2002) similarly assumes that it is not *constructions* as such that do or do not exhibit parallelism, but constructions *in a discourse context*. This approach is worth looking at in some more detail. Similar to SDRT, and hence to the approach to fragments we will develop in this thesis, it can be seen as coherence-based, where

[11]We have listed it also in the previous section because unlike (Asher 1993) they provide a detailed account of how parallelism beyond simple structural parallelism is computed.

> *"coherence establishment processes must be accounted for in analyses of [. . . ] linguistic phenomena [. . . ], and by extension, of other interclausal phenomena that have yet to be analysed in these terms."* *(Kehler 2002, p.8)*

The account builds on the rhetorical or coherence relations described by (Hobbs 1990), but further analyses these relations as belonging to one of three basic classes: resemblance relations, under which he groups for example *Parallel*, *Contrast*, but also *Elaboration*; cause-effect relations like *Result*; and finally contiguity relations like *Narration*.

The relations in these basic classes differ in which information they require to be established, and this is the most relevant point for us. While Cause-Effect and Contiguity relations only require the derivation of propositions, so Kehler, resemblance relations require the establishment of parallelism between individual elements of the clauses that are to be related. For this computation of parallelism, Kehler claims, syntactic information is required. This distinction can solve the apparent dilemma which the existence of examples like those in (123) that support syntactic approaches and like those in (125) which support semantic approaches poses, by claiming that there is an independent third factor that influences the resolution process. A minimal pair that illustrates this claim is shown in (132). The change in voice is not licensed in (132-a), where the clauses are connected by the resemblance relation *Parallel*; it is licensed in (132-b), which features *Explanation*.

(132) a. *John was shot by Bob, and Bill did too.
b. John was shot by Bob because Bill wouldn't.

We will not further validate these claims—they seem to be well argued for in (Kehler 2002)—but only note the approach offers welcome independent motivation for keeping syntactic information available during discourse processing. However, as an approach to discourse structure, it is a significantly less developed than SDRT; furthermore, it does not offer relations for dialogue at all.

## 3.5 Summary

We have seen that approaches to VPE can be classified according to the level of linguistic representation they work on, and we have seen the types of arguments typically brought forward in support of particular types of approaches, something that will recur in the next chapter. There seems to be a certain trend in the literature towards looking at VPE in the context of a broader theory of the relation of clauses in discourses, once more something we will take up again later. The notion of parallelism, to which we will come back, was introduced here, and finally we have briefly looked at an approach that argues for access to syntactic information for discourse processing.

# Chapter 4

# Previous Approaches to Fragments I: From Syntax to Pragmatics

In this chapter we review some of the literature on fragments. The approaches are organised in the same way as in the previous chapter, viz. according to which level of linguistic representation they are working on. We first review some arguments given by (Morgan 1973) in favour of a syntactic approach, and then look at two purely semantic approaches. The conclusion will be that both kinds of approach make wrong predictions. We then look at approaches that turn their attention to the influence of the wider context on fragment resolution. All approaches are shown to be unsatisfactory and lacking empirical coverage.

## 4.1 Introduction

A corpus study of VPE (Hardt 1997*b*) found that only 0.3% of clauses in the examined corpus contained this kind of ellipsis,[1] compared to the 11% for fragments we found, as described above in Section 2.2.4. And yet there is less literature on fragments than on other kinds of ellipsis. Fragments, however, occur almost exclusively in natural language *dialogue*, compared to VPE, which also frequently occurs in text. Hence, this discrepancy might just reflect a general bias of linguistics towards (monological) text.

Fragments are often handled in natural language interfaces to computer applications such as data bases.[2] However, approaches in this field are often restricted to a few kinds of fragments, normally short-answers, and rarely offer a principled account of fragments. The approaches we will discuss in this section here are different in that they mostly focus on linguistically principled analyses of fragments. The structure of this chapter is the same as that of the previous one: we classify the approaches according to the level of linguistic representation on which they place the resolution process, and we pick out paradigmatic representatives of syntactic and semantic approaches. We begin with an approach that is also historically the oldest, that of Morgan (1973).

## 4.2 Syntactic Approaches

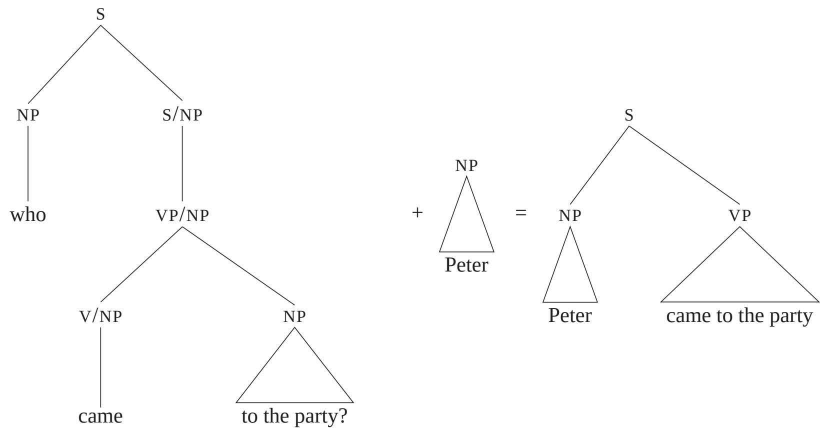
Consider the following simple example of a question–short-answer pair.

(133) A: Who came to the party?
B: Peter.

Examples like this make plausible an approach in which the short-answer is interpreted via reconstruction of syntactic structure. In this example, this reconstruction would be fairly trivial and involve only very simple operations on syntactic structures, as shown schematically in (134).

[1] This is a our own very rough estimation given the numbers cited in (Hardt 1997*b*). Hardt himself only states that he has identified 644 instances of VP-ellipsis in the Brown Corpus and Wall Street Corpus portions of the Penn Treebank (Marcus, Santorini & Marcinkiewicz 1993). These corpora taken together consist of 2 million words. Estimating the average length of a clause as 10 words, the combined corpus consists of 200,000 clauses, hence 664 clauses containing ellipsis make up only 0.3% of the overall number of clauses. Note again that this is our calculation and not Hardt's (1997*b*); and that we only use it to illustrate the difference in magnitude.

[2] Some earlier examples are (Thompson 1980) and (Carbonell 1983); we don't strive to offer a complete list here, for the reasons stated above.

(134) 


An early advocate of such an approach is Morgan (1973).[3,4] Following the tradition of generative linguistics, Morgan starts from the question how fragments are generated. He sees two basic possibilities:

- *Base generation*, where the grammar is extended so that certain other phrases than S are additionally allowed as start symbols, and interpretation is left to "some as yet unexplored principles of interpretation" (p.723). Morgan labels this theory DIT, for *direct interpretive theory*.
- *Ellipsis generation*, according to which fragments are generated by the grammar by deleting parts of full sentences. This is ET, the *ellipsis theory*, which Morgan argues for in that paper.[5]

From the perspective of linguistic theory nowadays, Morgan conflates two issues in the first alternative listed above: direct interpretation (i.e., interpretation of fragments without prior reconstruction of syntactic structure) does not imply base generation (i.e., allowing NPs or PPs to 'stand alone' syntactically). More modern techniques like semantic underspecification (cf. the detailed introduction in Section 6.3) allow a separation of these issues. In fact, we will argue later in the thesis for an approach that does not generate fragments 'in the base', but nevertheless could be called direct interpretation. Hence we will review in the following only Morgan's arguments specifically against direct interpretation, not those against base generation.

The following is a corollary to the claim made by ET that "fragments are generated by ellipsis transformations" (p.720):

> "ET claims that syntactic properties of full sentences will bear directly on the well-formedness and interpretation of fragments." (Morgan 1973, p.724)

Contrasting with this, DIT, like semantic approaches in general, predicts that there is no (purely) syntactic influence from full-sentences on fragments, since it sees fragments as being interpreted directly, without recourse to reconstruction of syntactic structure. The strategy for finding supporting evidence

[3] That paper is mostly concerned with presenting evidence in favour of such an approach; it doesn't actually work out the details of how it could be integrated in a syntactic theory.

[4] All page numbers in this subsection refer to (Morgan 1973).

[5] He also modestly calls this approach the 'common sense approach'.

for ET is clear then: one has to find syntactic idiosyncrasies that cannot be explained in any other way than through appeal to larger syntactic structures than are ‘visible / audible’. Or, to phrase it differently, one has to show that there are syntactic properties of the fragment that depend on syntactic properties of the antecedent clause, where the dependence is mediated through the condition of deletion under identity of structure. (Morgan 1973) offers several such examples, which we will first list and then discuss in the following.[6]

The first evidence Morgan presents relies on fairly theory specific syntax assumptions. Constructions like the one exhibited by the fragment in (135) were assumed in generative grammar of that time to result from movement in syntactic structure (the so-called ‘tough-movement’, where an object is moved out of its original position).

(135) A: If Hubert is hard to follow now, what will he be if he spoke more slowly?
B: Totally impossible for anyone to understand.

The details of this analysis are not really important here, the upshot is that in a grammar that follows a movement-analysis such things cannot be phrases, and hence they cannot be generated without ellipsis. We will soon see that other analyses are possible for such constructions that do not have this problem.

A second class of examples concerns what Morgan calls ‘complementizer choice’. In (136), the well-formedness of the fragment seems to depend on the sub-categorisation requirements of the verb. (One subcat frame for) ‘want’ requires a VP[*inf*] as complement, whereas ‘help’ is unusual in that it requires a VP[*bse*] complement. A semantic approach would have to show that these differences in syntactic type are reflected in different *semantic* types.

(136) a. A: What does John want?
B: To come over after dinner.
B′: *Come over after dinner.

b. A: What did John help you do?
B: *To wash my car.
B′: Wash my car.

Similarly, in (137) the type of complement-sentence licensed is determined by the syntactic requirements of the verb.[7]

[6] These arguments are repeated and neatly summarised in (Morgan 1989). Unless otherwise noted, all following examples in this section are either from (Morgan 1973) or from (Morgan 1989).

Note that we use the star and the hash-sign here simply to mark unacceptability in a wider sense. We do not make a decision here for whether the fragments marked thus are *ungrammatical* or just pragmatically infelicitous; i.e., we do not take a stance just yet on how fragment resolution works.

[7] This is Morgan’s (1973) example (59).

(137) A: What does John think?
B′: That Tricia has given birth to a 7-pound chin.
B″: *Tricia's having given birth to a 7-pound chin.
B‴: *For Tricia to have given birth to a 7-pound chin.

To this class of examples we might add (138), where the preposition is often seen as a purely functional verb particle without semantic impact. If that is the case, then it is difficult to explain in a semantic approach why the behaviour of such short-answers differs from that of NPs (see discussion above in Chapter 2 Section 2.3).

(138) a. A: On whom can we rely?
B: On Sandy.
b. A: Who did you see?
B: #On Sandy.

Examples like (139) below, so Morgan claims, show that binding theory, as defined for full sentences, determines the well-formedness of fragments. With reference resolved as indicated, the pronouns in (139-a)-B show the same pattern of licensing as in the possible full-sentence reconstruct B′; similarly the NP in (139-b) cannot be bound by another NP, just as in the full sentence (139-b).

(139) a. A: Who does $\text{Bill}_i$ like?
B: $\text{himself}_i$ / $\#\text{him}_i$
B′: Bill likes $\text{himself}_i$ / $\#\text{him}_i$

b. A: What does $\text{John}_i$ think?
B: #That the $\text{bastard}_i$ is spying on him.
B′: $\#\text{John}_i$ thinks that the $\text{bastard}_i$ is spying on him.

Next, as Morgan points out, certain adjectives, for example 'content', are restricted to occur in predicative position. As (140) shows, such adjectives can only be used to correct similarly predictive adjectives; in (140-b), where the correctum is used attributively, the fragmental correction is infelicitous.[8]

(140) a. A: Is the boy unhappy?
B: No, content.
B′: No, he's content.

b. A: Is he an unhappy boy?

[8]These examples are Morgan's (1973) (62)–(69). For some speakers, "content" seems to be allowed in attributive position; however, with "afraid" for instance these speakers do agree on the pattern observed by Morgan.

B: #No, content.
B′: *No, he's a content boy.
B″: No, he's content.

Fragments also are subject to subjacency constraints, according to Morgan. Just as in the full sentences in (141-a) and (141-b) below the operation of *wh*-movement—again speaking in terms of a syntactic theory that uses such operations—cannot move the indexed NP out of the 'island' complement sentence or relative clause, respectively, the fragmental clarification is not licensed. Similarly, conjunction builds an 'island', as shown in (141-c).

(141) a. A: That John shot someone$_i$ upset his father.
B: *Who$_i$? / *Whom$_i$?

b. A: A man who shot someone$_i$ has escaped.
B: *Who$_i$? / *Whom$_i$?

c. A: Judy and Peter$_i$ left.
B: No, Sandy$_i$. (= Judy and Sandy left.)

As a final piece of evidence, Morgan presents examples where the case marking of the fragment depends on its antecedent.[9]

(142) a. A: Wem$_{dat}$ hast Du geschmeichelt?
B: [Dem Mann]$_{dat}$.
B′: [Den Mann]$_{acc}$.
*(A: Who did you flatter? — B: The man.*$_{dat/*acc}$*)*

b. A: Wen$_{acc}$ hast Du gelobt?
B: [Den Mann]$_{acc}$.
B′: [Dem Mann]$_{dat}$.
*(A: Who did you praise? — B: The man.*$_{*dat/acc}$*)*

Unfortunately, the data presented so far is not as unequivocal as this presentation made it appear, as already (Morgan 1973) acknowledges and (Morgan 1989) reinforces. There are (at least) two classes of objections against this evidence and the approach it is meant to support. First, one can show exceptions to the assumed rules or present direct counter-evidence, or present different (non-syntactic) explanations; a second line of argumentation could be called the 'argument from uniformity'. It goes

[9]This class of examples is added in (Morgan 1989). There, Morgan gives examples from Korean, but since the author of this thesis happens to speak a language with strong case marking, we give an example from German here. This example was already used in Chapter 2.

back to (Yanofsky 1978) and (Barton 1990) and will be extended here; it is directed against this type of approach in general.[10]

Let us begin with the attack on the data. First of all, as already noted above, the problem with the examples containing ‘tough-movement’ simply doesn’t occur in mono-stratal, i.e. movement-free syntactic theories. Here constructions like ‘easy to follow’ are simply phrases that can be independently ‘generated’. Hence, this cannot be counted as evidence in support of ET.

Secondly, the examples presented above concerning complementizer choice must be augmented with examples where there are divergences between fragments and full sentence correlates. In (143-a), a modification of (136-a), the complementizer is optional. In (143-b), on the other hand, the complementizer is optional only in the full sentence case. Finally, in (143-c) the fragment is of a syntactic type that is not licensed in the corresponding full sentence, at least not in standard complement order.

(143) a. A: What does John want to do tonight?
B: To come over after dinner.
B′: Come over after dinner.

b. A: What does John believe?
B: That Optimality Theory is great.
B′: #Optimality Theory is great.
B″: John believes that Optimality Theory is great. /
John believes Optimality Theory is great

c. A: Concerning the weather, what can we rely on?
B: That it will rain.
B′: *We can rely on that it will rain.
B″: That it will rain we can rely on.

This difference will be handled by the fragment-grammar in our approach; for example in (143-b) B′ simply does not receive a parse as fragment, only as ‘normal’ sentence. That sentence (“Optimality Theory is great”) cannot be a *direct* answer to the question (“What does John believe?”), given the compositional semantics of the question.

One can also find examples that violate Morgan’s (1973) assumption that Binding Theory as defined on full sentences (always) influences the well-formedness of fragments. In (144), the fragments are licensed even though plausible full-sentence correlates aren’t.[11]

[10] There is also an ‘intra-theoretical’ objection we have to mention briefly, namely a problem with the changes to the scope of grammar implied by ET. Morgan’s ellipsis rule would have to work over *pairs* of sentences, and so the grammar must turn into a discourse grammar. This is a far-reaching conceptual change; one that hasn’t been fully appreciated by syntactic approaches to VPE, which would have to make the same step, too.

[11] These examples are due to (Ginzburg 1999*b*).

(144) a. A: Whose complaints annoyed Bill and Jill most intensely?
B: Each other's.
B′: #Each other$_i$'s complaints annoyed [Bill and Jill]$_i$ most intensely.

b. A: Who does Jill$_i$ think Bill desires?
B: Herself$_i$/ # Her$_i$.
Jill$_i$ thinks Bill desires *herself$_i$/ her$_i$.

Morgan (1973) himself points out that the island constraints are "mysteriously" weakened if the whole island is moved towards the end of the sentence, as in the following example.

(145) A: A man has escaped who shot someone$_i$.
B: Who$_i$? / Whom$_i$?

(Ginzburg 1999*b*) speculates that there is an independent factor that can explain the pattern above in (141), namely the influence of focus/background partition on the licensing of certain speech acts, and in turn the connection between that partitioning and subjacency. Where (Morgan 1973) explains the ill-formedness of (146-a) with a constraint that forbids the forming of fragments from islands, (Ginzburg 1999*b*) claims that such fragmental constructions have to be licensed by focus on the corrected/elaborated element, as in (146-b).[12] If focus is influenced by subjacency, then this analysis places the licensing factor on the antecedent utterance and not the fragment reconstruct.

(146) a. A: Did John and Bill$_i$ leave this morning?
B: #No, Harry$_i$ (= No, John and Harry left this morning.)

b. A: Can you help me with my [homework]$_F$?
B: Your homework, no. Your carpentry, yes.

Moreover, pronouns in English seem to violate the case constraint illustrated above with example (142), as again Morgan (1989) himself notes. In (147) the pattern of acceptability for the pronouns is exactly reversed between fragment and full sentence.

(147) A: Who cooked this meal?
B: # I/ Me/ # She/ Her/ etc.
B′: I / *Me/ She /*Her /etc. cooked this meal.

[12]We will offer a similar, but superior analysis below in Section 8.3.1.7. Briefly, we analyse the fragment in (146-b) for example as being contrasted with the 'no' answer, and hence we can explain why "no, but with your carpentry." is a felicitous answer to the question in (146-b).

Finally, as (Ginzburg 1999*b*) notes, it is not at all clear what the correct full-sentence correlate for fragments should be in such an approach. So far, we have implicitly assumed that this is a simple *in situ* substitution of the *wh*-phrase with the fragment phrase, as sketched above in (134), but as the following examples show, sometimes only clefts or pseudo-clefts work. In any case, however, there does not seem to be one construction that consistently works.

(148) a. A: Who stole the phonemic level?
B: Not Bill.
B′: *Not Bill stole the phonemic level.
B″: It was not Bill who stole the phonemic level.

b. A: What did Ivan say?
B: That Bill stole the phonemes.
B′: What Ivan said was that Bill stole the phonemes.
B″: *It was that Bill stole the phonemes that Ivan said.

c. A: What did Jill say?
B: Nothing.
B′: #It was nothing that Jill said.
B″: #What Jill said was nothing.

We come now to what we labelled the ‘argument from generality’ above. (Yanofsky 1978) was the first to put forward this kind of argument against an ellipsis approach. Her data consists of fragments of the type we explicitly do not deal with in this thesis, namely ‘discourse-initial’ or ‘situational’ fragments like those in (149).

(149) a. Fire!
b. A coffee please.
c. The red one.

Nevertheless, her argumentation is interesting, and so we briefly reconstruct it here. It goes roughly as follows: Such utterances, like the fragments we have looked at, seem to be used to perform speech acts with complete ‘messages’ as content.[13] However, there is no linguistic context, and in particular no syntactic context from where material for their reconstruction (or deletion under identity) could come from. Hence, there has to be a part of the grammar that deals with this kind of utterance and that cannot rely on ellipsis rules. But if there is such a part, then theoretical parsimony would dictate that other elements of the theory that only do the same, like fragment ellipsis rules, would have to be removed.

[13] This could be disputed for (149-a), but does seem acceptable for the others.

One doesn't even have to turn to this quite different kind of fragment to make this argument, as (Barton 1990) points out, and as we can reinforce with our data from Chapter 2. The further away one moves from the paradigm question–short-answers, or even only from NP–short-answers, the less plausible a purely syntactic approach appears. If we recall the kinds of fragments we called *resolution-via-inference* above, as for example shown in (150), there are many more types of fragments where syntactic reconstruction isn't plausible since there quite simply is not enough syntactic material in the context.

(150) A: Why did Peter leave so early?
B: Exams.

As a preliminary conclusion, we can say that the situation looks confusing. Many of Morgan's points have been challenged, and for some we have offered satisfying different explanations, but we haven't so far conclusively shown that *no* syntactic properties of full-sentences at all bear on the well-formedness of fragments. But in any case, a syntactic ellipsis-approach can't be the full story, since we also have to explain fragments for which there is no explicit linguistic context. In the next Chapter we will discuss an approach that tries to find a third way, by allowing *some* syntactic influence from full-sentences on fragments. However, as will become clear in the discussion of this approach, it still has some serious shortcomings, which we will try to overcome when we develop our approach in the later chapters. Now, however, we will briefly mention a few purely semantic approaches, and look at two approaches that turn their attention to fragments over and beyond short-answers.

## 4.3 Semantic Approaches

Just as with VPE, the resolution of fragments can be modelled as a recovery of a property from the context. For example, the fragment in the fragmental correction in (151-a) could be represented as shown in (151-b), and the task then is to find a value for the higher-order variable *P*.[14]

(151) a. A: Peter came to the party.
B: No, John.
b. $P(john)$

To find this property, a DSP-style approach could be used, as shown in (152) (we adapt the terminology 'source' and 'target' to fragments and their antecedents for this illustration).[15]

[14] We will argue argue *against* this representation extensively below in Chapter 6.
Moreover, we are not concerned with how such representations could be generated by a grammar; but we note that what Morgan (1973) called 'base generation' would not give us such a logical form.

[15] An approach to correction that indeed starts from this basis is (Gardent, Kohlhase & van Leusen 1996).

(152)
a. Representation source clause: $come\_to\_party(peter)$
b. Representation target clause: $P(john)$
c. Parallel elements: $peter \leftrightarrow john$
d. Equation: $P(peter) = come\_to\_party(peter)$
e. Solution: $P \mapsto \lambda x.come\_to\_party(x)$
f. Substitution of e) in b), $\lambda x.come\_to\_party(x)[john]$
β-reduced: $come\_to\_party(john)$

A semantic treatment of short-answers could be even simpler, at least in some theories of question-semantics. In theories coming from the 'structured meaning' tradition (e.g. (Tichy 1978, Hausser & Zaefferer 1978, von Stechow & Zimmermann 1984, Krifka 1999)), "question meanings are functions that, when applied to the meaning of the answer, yield a proposition." (Krifka 1999, p.2).[16] In such an approach, the equation-step of a DSP-style approach isn't even required, and short-answers are regarded as the natural form of answers. (153) shows how this approach would resolve an NP short-answer (using a simplified logical representation for the NL expressions).

(153)
a. Who saw Mary? $\leadsto \lambda x[saw(x,m)]$
b. John. $\leadsto j$
c. $\lambda x[saw(x,m)]\,j \Leftrightarrow saw(j,m)$

This extends naturally to verb-questions, as the following shows (again using for the purposes of illustration a rather oversimplistic representation):

(154)
a. What did John do? $\leadsto \lambda P[P(j)]$
b. Fishing. $\leadsto \lambda x.fish(x)$
c. $\lambda P[P(j)](\lambda x[fish(x)]) \Leftrightarrow \lambda x[fish(x)](j) \Leftrightarrow fish(j)$

*res-via-inf*-examples like the one we gave above ((150), repeated here as (155)) can be seen as supporting such semantic approaches, since those approaches allow the property that is to be recovered to be only implicit in the context and hence to be inferred.

(155) A: Why did Peter leave so early?
B: Exams.

[16] This view of questions leads to various problems, most notably that it makes polarity questions, single *wh*-questions and multiple *wh*-questions all of different semantic types, and equates questions semantically with relations. For a summary of arguments against such a view see (Groenendijk & Stokhof 1997, p.1107), for a defence see for example (Krifka 1999) or (Ginzburg & Sag 2001, pp. 100–119). In a Groenendijk-Stokhofian approach (Groenendijk & Stokhof 1984, Groenendijk & Stokhof 1997), the resolution mechanism described above would have to be amended by an operation of stripping off the ?-functor that takes such λ-abstracts into sets of propositions.

To summarise, as with VPE, we find ourselves in a dilemma. While some data points towards a semantic approach, there is still the data *pro* a syntactic approach from the previous section. Before we come to possible ways out of this impasse, however, we have to very briefly look at another semantic approach and then at some pragmatic approaches.

The CORE-language engine, developed in the late 1980s by SRI-international, implemented a semantic approach to fragments which is interesting for us here because of their use of *underspecified* representations (i.e., descriptions of logical forms; cf. the introduction in Section 6.3). To give an example, the NP-fragment "John" is represented in their system by the following *Quasi Logical Form (QLF)* (Alshawi 1992, p.252):

(156) `a_form(<t=ell, p=vp>, P, [P, john]).`

This representation expresses that this is an anaphoric formula (`a_form`), where the reason for its anaphoricity is ellipsis (`t=ell`) of type VP (`p=vp`), and the anaphoric variable is `P`, which is the functor of the formula, i.e. the formula is $P(john)$.[17] This is very close to the representation in the DSP-style approach we suggested above, and indeed the authors consider their approach to be 'a particularisation of the general method introduced by [DSP]' (p.254). It is a particularisation because rather than going via equations and higher-order unifications to actually recover a value for `P`, in this approach ellipsis is resolved via direct substitution of terms (cf. also (Crouch 1995)). This substitution acts on an intermediate level of representation (that of *Resolved Quasi Logical Form (RQLF)*), at which scope-relations are already fixed; this implements the requirement for scope-parallelism between source and target noted above. While the use of underspecification in this approach is interesting—our approach will also make use of this technique—the approach is first of all limited to *res-via-id* fragments (it relies on substitution and cannot construct properties P) and also doesn't capture the syntactic constraints discussed above.

## 4.4 Pragmatic Approaches

In this section we discuss approaches to fragments that do not have direct counterparts in the literature on VPE.[18] They are not discourse-based in the same sense as above in Section 3.4, but rather emphasise the importance of pragmatic reasoning on the resolution of fragments. They both use the notion of intention as a starting point, (Barton 1990) using an explicitly Gricean approach and trying to make the notion of conversational maxims fruitful for this application, whereas (Carberry 1990) takes a more direct approach in recovering intentions behind utterances.

[17] It is not clear to us from the presentation in (Alshawi 1992) what exactly the impact of `p=vp` is in this representation. It cannot be meant to restrict `P` to values that are representations of VPs; they themselves later give an example where the short-answer fills an object-position, and hence where the property is not a VP.

[18] Although cf. (Kuno 1987), where informal pragmatic constraints on VPE are stated.

### 4.4.1 Barton (1990)

As already briefly mentioned above in Section 2.4, (Barton 1990) makes an interesting distinction of the phenomenon of fragments which resembles ours between *res-via-id* and *res-via-inf*. Barton motivates this distinction with data like the following:

(157) a. A: What stops the White House staff from visiting Tip O'Neill in his Congressional office?
B: (An) old grudge.

b. A: The White House staff doesn't visit Tip O'Neill in his Congressional Office.
B: Old Grudge.

Data like (157-b) is explicitly offered as evidence against Morgan's (1973) ellipsis approach, with the argumentation we described above, namely that in these examples there is no explicit reconstruction/deletion trigger. As Barton similarly only sees the two alternatives for analysis of fragments offered by Morgan and described above, she hence argues for *base generation* and *direct interpretation*. Again, we are not much interested in the syntactic arguments for *base generation*, and so we concentrate on her arguments for and her explication of a DIT.[19]

According to (Barton 1990), the data motivates the postulation of two distinct modules of interpretation, one of *linguistic context*, responsible for the resolution of fragments like (157-a), and one of *conversation context*, in which fragments like (157-b) are resolved. We turn to the former first. In Barton's model, utterances are represented in the 'linguistic context' in the form of what roughly corresponds to (syntactic) logical form annotated with thematic roles / argument structure, as shown in (158-b) for the first sentence of exchange (158-a).

(158) a. A: Mitchell gave a sworn statement, didn't he?
B: Yes, sir. To the Jury.

b. $[_S$ $[_{NP}$ Mitchell] $[_V$ gave] $[_{NP}$ a sworn statement] ], ...
SOURCE [SRC, PATIENT, (GOAL)] PATIENT

In the 'module of linguistic context' now, fragments or 'independent constituent utterances', as she calls them, are resolved through an operation of *discourse inference*, which starts from a question about intentions of the speaker:

> *"[W]hat hearers do is draw inferences about the intentions of a speaker with respect to fitting an independent constituent utterances into this structure of linguistic context."* *(Barton 1990, p.142)*
> *"If a discourse sequence includes an independent constituent utterance that potentially matches an*

[19] Moreover, much of the book is concerned with (Chomskian) syntactic theory specific considerations of the autonomy of linguistic modules, which do not concern us here. We will later use a mono-stratal linguistic theory which does not have such problems.

> *expansion possibility of a previous element within the structure of linguistic context, then assign the constituent utterance to function as that expansion possibility in an elaboration of the structure of linguistic context." (p.112)*

The 'expansion possibility' is defined there as 'the ability of a lexical item to select an argument role or to occur with a modifier.' This operation would resolve the fragment 'to the jury' in (158-a) above to fill the thematic role GOAL, as shown in the following.

(159) [$_S$ [$_{NP}$ Mitchell] [$_V$ gave] [$_{NP}$ a sworn statement] ], didn't he?
SOURCE [SRC, PATIENT, (GOAL)] PATIENT
[$_{PP}$ [$_P$ to] [$_{NP}$ the Jury] ]
GOAL

Barton (1990) does not further specify how such resolved fragments are interpreted, i.e. she does not offer a formal semantics of such constructs. However, with some goodwill, this part of the approach can be understood as a semantic approach as in the previous section, extended with a representation for optional elements. In any case it offers the interesting observation that fragments can also 'add' optional complements; as shown in Chapter 2.

The module of conversational context on the other hand is based on the notion of cooperative contribution, following (Grice 1969). Fragments like the one in (157-b)—in general, our *res-via-inf* fragments—are resolved in Barton's approach through an operation of 'cooperative inference', via reasoning about their relevance:

> *"In a discourse sequence, the representation of linguistic context for each utterance becomes the representation of an utterance with a potential structure of implicatures explaining its relevance as a cooperative contribution within the conversational context." (Barton 1990, p.166)*

The principle of relevance, however, is notoriously difficult to formalise (but see (Sperber & Wilson 1986)), and Barton does not offer much in this direction, apart from a classification into 'domains of relevance':

(160) Types of Cooperative Inference, according to (Barton 1990, p.167):

a. information from knowledge of the physical context of the situation
b. knowledge of the topic of conversation
c. background knowledge of interlocutors
d. knowledge of the world

Insofar as the theory of discourse in which we realize our resolution mechanism (SDRT) realizes the idea of relevance, our approach can be seen as related to this. However, as will be discussed below, SDRT crucially denies that reasoning about intentions is *always* the starting point for discourse inference; rather it places linguistic and discourse information on the same level.

### 4.4.2 Carberry (1990)

Another interesting pragmatics-based approach is offered by (Carberry 1989, Carberry 1990).[20] In this approach, resolution of fragments is based on recovering the intentions and plans of dialogue participants, and seeing how the fragment might further such plans. Carberry claims that

> *"understanding intersentential ellipsis often depends more on pragmatic knowledge, such as the inferred task-related plan and discourse goals motivating the speaker, than on the syntactic structure or semantic content of the preceding utterances" (Carberry 1989, p.76)*

As support for this view, she gives the following example:

(161) A: I want to cash this check.
Small bills only, please.

For examples like (161), we do not disagree with the claim that reasoning about plans is important to resolve the fragment. In fact, as the reader will recall, we also have in our taxonomy in Chapter 2 speech act types that are individuated with reference to plans. What we disagree with is the claim that one *always* has to reason about intentions when resolving fragments.[21] We claim that there is a large class of fragments where such plan-based reasoning can be avoided: for example short-answers, or elaborations. Moreover, there seems to be no room for *linguistic* structures like discourse structure in Carberry's model, whereas (Moore & Pollack 1992) show that this information is equally important. The following example (repeated from Chapter1) illustrates this:

(162) Peter: Let's meet the weekend after next.
Mary: OK, but not Saturday.
Peter: Right. #2pm?

If only goals are recorded, the model cannot predict that the last utterance in this example is infelicitous; it cannot mean "how about we meet at 2pm on Sunday?", even though meeting on Sunday can be assumed to be the goal at that point. A plan-based approach restricts the flow of information too much, in that it always puts intentions at the beginning of the process of computing implicatures. Following SDRT, we think this model is not general enough.[22]

[20] We have already discussed her taxonomy of fragments above in Section 2.4.

[21] Even though the statement quoted above is qualified with 'often', Carberry does not show how her model would make use of other knowledge sources, and it is not clear whether she suggest that only other kinds of ellipsis (e.g. VPE) can be resolved using 'syntactic structure or semantic content of the preceding utterances', or some kinds of fragments as well. In any case, her algorithm always begins with inferring intentions, and that is what we are arguing against here.

[22] (Fernández & Ginzburg 2002) also argue against Carberry on the grounds that such complex reasoning does not always seem to be required. However, as we will argue in the next chapter, their approach does not seem to be capable of *ever* allowing it. We simply want to argue for a more flexible set-up where information can flow either way.

Moreover, there is amassing psycho-linguistic evidence that costly reasoning about intentions is only the last resort in dialogue processing, and if there are cheaper ways of interpreting material then they are chosen (Pickering & Garrod in press). However, we do not argue systematically for psycholinguistic plausibility of our model.

Lastly, we disagree on what the task of fragment resolution is. This can be illustrated with the following examples from (Carberry 1989). Carberry contrasts these dialogues to show that information about shared beliefs is needed to resolve fragments. In (163), the fragmental question "at night?" is used to express surprise, according to Carberry, whereas it is simply used to enquire about more details in (164). This difference arises from the fact that the queried information is already mutually known in the former dialogue, but not in the latter.

(163) A: When does CS400 meet?
B: CS400 meets on Monday from 7:00pm until 9:00 pm.
A: Who's teaching it?
B: Dr. Brown.
A: At night?

(164) A: Who's teaching CS400?
B: Dr. Brown.
A: At night?

While we do not dispute that these two utterances carry different implications, we do not want to model these directly in the resolution of the fragment. In our approach, the question will in both cases resolve to something paraphraseable as "is Dr. Brown teaching this course at night?"; additional implicatures might then be computed on the basis of this. Our (more modest) goal is simply to predict the truth conditional content of fragments, including the (truth conditional) resolution of the underspecified bits.

## 4.5 Summary

In this chapter we have reviewed some of the literature on fragments. We have discussed Morgan's (1973) data in support of syntactic approaches, and concluded that even though some of his arguments can be rebutted, there seems to be evidence both for and against such an approach. We have then briefly reviewed purely semantic approaches, and argued against approaches that take conversational intentions as their only starting point.

---

The SDRT-arguments against plan-based methods in general are summarised in (Asher & Lascarides 2003, Ch.3).

# Chapter 5

# Previous Approaches to Fragments II: A Grammar-Based Approach

In this chapter we discuss an approach to fragments that tries to unify elements of syntactic and semantic approaches. It uses a semantic resolution strategy, but combines it with a syntactic matching condition. The approach is based on a combination of a discourse structure theory (KOS, discussed in Section 5.2) and a grammar, and hence we will dub it the *Grammar-Based Approach* (GBA) here to contrast it with the *Coherence-Based Approach* (CBA) we will later develop. We discuss the GBA in a bit more detail than the previous approaches because it contains valuable elements which we will take up in our approach. As we will show, it also has some shortcomings, however.

The presentation begins with a short look at the grammar formalism that both this approach and ours uses, namely HPSG. We then review the theory of discourse context used in the approach, and introduce and discuss the approach.

## 5.1 Background: HPSG

The approach to fragment interpretation discussed in this chapter is framed in a syntactic theory called *Head-Driven Phrase Structure Grammar* (HPSG). We now highlight a few features of HPSG, focussing on its connection to ideas developed elsewhere in the thesis.[1]

The basic linguistic entity that HPSG is concerned with is, following (de Saussure 1916), the *sign*, which is a pairing of form and meaning. It is modelled in the theory by typed feature structures, which in turn are (possibly partially) described by statements in a logical language, typically notated as attribute value matrices.[2,3] The descriptions or constraints are organised into *types*, so that each type is a (partial) description of a kind of linguistic object, e.g. a phrase or a word. The operation of *unification* provides a means of combining types to form other types, e.g. to combine words into phrases. Figure 5.1 shows an excerpt of the type-definition from (Ginzburg & Sag 2001). The novel element that will interest us most here is the specification of the feature CONTXT (context), with its sub-features SAL-UTT (salient utterance) and MAX-QUD (maximal question under discussion). These features will be described shortly.

Unlike Chomskian approaches (c.f. e.g. (Haegeman 1994)), HPSG does not need to posit empty, i.e. 'invisible/inaudible' elements in analyses. There are two reasons for this. First, many phenomena were shown to be amenable to a lexicalised approach, where the role played in other theories by those 'phantom' elements is satisfied by special demands of the lexical items. Where earlier variants of HPSG strived to derive *all* of the behaviour of sentences from the lexical items in it, using only very simple combinatorial schemata, later versions (Sag 1997, Ginzburg & Sag 2001) emphasised the usefulness of letting phrasal types introduce semantic material additional to what is encoded in the lexical entries. In other words, *constructions* were added to the analytic arsenal of the theory. The approach to fragments described in the present chapter and our approach both make use of this tool.

## 5.2 Background: KOS, a theory of discourse

The dialogue-semantic backbone of the GBA is a theory called KOS (Ginzburg 1996*a*, Ginzburg 1996*b*).[4] It evolved out of Ginzburg's investigation of the semantics of questions (Ginzburg 1995*a*, Ginzburg 1995*b*). In this latter theory, questions are not only individuated by their set of full answers, as in more traditional approaches (e.g. (Groenendijk & Stokhof 1997)), but also by their 'resolvedness condition', which is an agent-relative pragmatic notion. KOS (or the QUD-model, as it is also called) bases on this treatment of questions a theory of dialogue semantics.

[1] A textbook introduction to HPSG is (Sag & Wasow 1999), the 'classic' presentations are (Pollard & Sag 1987, Pollard & Sag 1994).

[2] There is a customary, but sloppy manner of speaking in which the descriptions (or rather, the notation format for descriptions) are also called feature structures. Description and model should however not be confused.

[3] The idea of using partial descriptions / constraints is what connects this syntactic theory nicely to the semantic technique of underspecification which we will later use.

[4] Apparently, KOS is not an acronym, just the name for the theory (Ginzburg 1996*a*).

$$
\begin{bmatrix}
\textit{sign} & \\
\text{PHON} & \textit{list(form)} \\
\text{SYNSEM} & \begin{bmatrix} \textit{synsem} & \\ \text{LOCAL} & \begin{bmatrix} \textit{loc} & \\ \text{CAT} & \begin{bmatrix} \textit{cat} & \\ \text{HEAD} & \textit{part-of-speech} \\ \text{SUBJ} & \textit{list(synsem)} \\ \text{COMP} & \textit{list(synsem)} \\ \text{SPR} & \textit{list(synsem)} \end{bmatrix} \\ \text{CONT} & \textit{sem-object} \\ \text{STORE} & \textit{set(scope-object)} \end{bmatrix} \\ \text{SLASH} & \textit{set(local)} \\ \text{WH} & \textit{set(scope-object)} \\ \text{BCKGRND} & \textit{set(fact)} \end{bmatrix} \\
\text{CONTXT} & \begin{bmatrix} \text{C-INDICES} & \begin{bmatrix} \textit{c-ind} & \\ \text{C-SPKR} & \textit{index} \\ \text{C-ADDR} & \textit{index} \\ \text{U-LOC} & \textit{index} \end{bmatrix} \\ \text{SAL-UTT} & \textit{set(local)} \\ \text{MAX-QUD} & \textit{question} \end{bmatrix}
\end{bmatrix}
$$

Figure 5.1: An excerpt of the specification of the type *sign* from (Ginzburg & Sag 2001)

The model uses a game metaphor to describe dialogue,[5] according to which the participants (players) make moves in a game whose public effect is recorded on the gameboard and its private effects on private scorecards. KOS combines this with an information state approach, which emphasises the *update effects* of utterances on a dialogue context.[6] The information state in the theory consists of two attributes, the ‘unpublished mental situation’ (UNPUB-MS) and the ‘dialogue game board’ (DGB):

(165)
$$\begin{bmatrix} \text{UNPUB-MS} & \textit{(goals, beliefs etc.)} \\ \text{DGB} & \begin{bmatrix} \text{FACTS} & \textit{set(proposition)} \\ \text{QUD} & \textit{question} \\ \text{LATEST-MOVE} & \textit{sign} \end{bmatrix} \end{bmatrix}$$

The field ‘unpublished mental situation’ stores the private intentions and beliefs of the DP relative to which questions get resolved, while the dialogue game board is a complex attribute with the following sub-attributes:

- FACTS, which is the common ground, a set that collects mutually believed propositions.
- QUD, the questions under discussion, a set consisting of the currently discussable questions, partially ordered according to the relation of “conversational precedence”, ‘$\prec$’.
- LATEST-MOVE, which stores the content of the latest move made, e.g. “A asserted that p”.

The QUD-stack is the most important device, since it determines what can coherently be uttered at any point in a discourse.[7] It carries almost all the burden of structuring the discourse. The structure arises on the basis that the question maximal on QUD licenses only information that stands in a certain relation to the question. This follow-up to the question can be realised as a fragment, in which case the QUD determines how this fragment is resolved. Before we come to this, however, we have to briefly list in what ways information can stand in a relation to the QUD according to KOS. First, it can be about the question (roughly speaking, by partially answering it), or the information decides the question by providing an exhaustive answer, or it resolves it by providing information that is sufficient, relative to the ‘mental situation’ of the DP, i.e. relative to her goals and beliefs; the last notion thus is context (i.e. agent)-dependent. Questions can stand in the relation depends-on, where a question $q_1$ depends-on on question $q_2$ if everything that decides $q_1$ also decides $q_2$. Utterances that stand in one of these relations to the QUD are considered coherent in this theory.

Dialogue is dynamic, and the issue at hand can change during the course of a conversation. This poses

[5]This metaphor can be traced back to (Wittgenstein 1953/84), via (Lewis 1969, Carlson 1983) *inter alia*.

[6]Information-State approaches are described in (Traum, Bos, Cooper, Larsson, Lewin, Matheson & Poesio 1999). Such approaches are related to *dynamic semantics*, where the notion of update likewise is central ((Kamp 1981, Kamp & Reyle 1993, Groenendijk & Stokhof 1991); see also the introduction below in Section 6.3).

[7]The idea of using questions to structure discourse is based on (Carlson 1983); similar models have been proposed independently by Klein & von Stutterheim (1987) and Roberts (1996) (who even uses the same name, QUD). For a critical discussion of this way of structuring the discourse context, see (Asher 1998).

the question of how the QUD changes over time, i.e. how questions are put on and removed from the QUD, and how its partial ordering comes about. A question gets removed from QUD (QUD gets *downdated*) when information is provided that resolves it; the simplest way to *update* the structure is simply to ask a question. If a question is asked and accepted for discussion it is put on the QUD. The next example shows two updates and two downdates. The stack-architecture of QUD ensures that A's first answer is related to the second question, and the second answer to the first—i.e., the QUD structures the discourse into segments.

(166)

| | | |
|---|---|---|
| A: Who shall we invite for tomorrow? | $Q_1$ put on QUD | |
| B: Who will agree to come? | $Q_2$ put on QUD | (Ginzburg 1997, p.69) |
| A: Joe and Peter and maybe Carl. | $Q_2$ removed from QUD | |
| B: I see. Peter then. | $Q_1$ removed from QUD | |

Now certainly not all exchanges in dialogue follow this question-response model. But since QUD is the only structuring device, all moves have to have an influence on it. In this model, assertions also give rise to updates of the QUD.[8]

An assertion "that *p*" puts the question "whether *p*" on QUD. The focus partitioning of the assertion can also trigger further updates, as shown in (167).[9]

(167)
a. BILL likes John. → "Who likes John?" on QUD
b. Bill likes JOHN. → "Who does Bill like?" on QUD
c. Bill LIKES John. → "In what relation does Bill stand to John?" on QUD

These are the questions with which QUD is automatically updated when an assertion is made. Other questions are concerned with the 'clarification potential' of utterances, i.e. the possibilities for clarification they give rise to. We will discuss this separately below in Section 5.4.

However, often contributions in a dialogue address issues that have not been explicitly raised, as for example in the short dialogue (168). In such cases coherence is preserved under this approach if the dialogue participant is able to construct a question which is relevant at that point in the dialogue, and then accommodates this question onto the QUD. This operation of accommodation must resort to plans and mental states. So far, however, to our knowledge there exists no detailed theory of how to do this.[10]

[8] To our knowledge, so far there is no detailed account of how requests would affect the QUD.

[9] After (Engdahl, Larsson & Ericsson 2000).

[10] But cf. the first attempts in this direction developed in the TRINDI project, e.g. (Cooper, Engdahl, Larsson & Ericsson 2000, p.2). The "as cheap as possible" in (168) would in this approach force a question on the stack which hasn't been raised explicitly. Cooper et al. (2000) admit that it is rather unintuitive to consider the offering of additional information as answers to as yet unraised questions, and propose to call this "issues raised" instead of questions. Which issues can be raised presumably has to do with the underlying goals the DPs follow.

(168) A: When do you want to travel?
B: Early April. As cheap as possible.
(Cooper et al. 2000, p.2)

We perceive this need for accommodation as a limitation of the model; in our approach, we would simply analyse the relation between B's two utterances above as an *Elaboration*, without the need to construct any questions. While it might be possible to always compute and accommodate questions that encapsulate the semantics of the rhetorical relations introduced in Chapter 2 ("and then what happened?" for *Narration*, "why?" for *Explanation*, etc.), using the same techniques of defeasible reasoning described below in Chapter 8, this step of accommodation of questions then becomes superfluous. Hence, it seems to us that the QUD-model would at least have to be extended with a notion of speech acts / rhetorical relations independent from question-answering. We will return to this issue when we discuss the GBA in the next section.

## 5.3 The Basic Approach

Before we discuss the technical details of the GBA, we have to return briefly to the data given by Morgan (1973) in favour of a syntactic approach. Following (Ginzburg 1999*b*), we claimed that the data shows that neither *all* syntactic properties of full-sentences bear on the well-formedness of fragment, nor *no* properties at all. For example, the required choice of complement in examples (136) and (138) from Section 4.2 (repeated here as (169) and (170)) suggests syntactic constraints are necessary.

(169) a. A: What does John want?
B: To come over after dinner.
B′: *Come over after dinner.

b. A: What did John help you do?
B: *To wash my car.
B′: Wash my car.

(170) a. A: On whom can we rely?
B: On Sandy.

b. A: Who did you see?
B: #On Sandy.

In light of this conflicting evidence, (Ginzburg 1999*b*) suggests that the syntactic influence is more localised, namely as a syntactic 'parallelism' between elements of source and target. The pattern in

(169) can then be explained by the need to match the sub-categorisation requirements of the verb in the question with the category of the fragment-phrase. This is the basic idea that is implemented in the approach.

The characterisation of this approach as 'grammar-based' already emphasises its distinctive feature, namely that it attempts to integrate the resolution of short answers into the grammar component of linguistic knowledge. In a nutshell, the GBA works as follows: fragments are resolved by functional application of the semantic representation of the question to the representation of the meaning of the short-answer, under the condition that certain *syntactic* features can be unified. The resolution is purely semantic in that no syntactic structure is reconstructed and then interpreted; however syntactic information is used to provide additional matching conditions.

As mentioned above, the grammar formalism in which this GBA is realized is HPSG. One design feature of this formalism that makes it attractive for a grammar based implementation especially of this combined syntactic/semantic strategy is that it offers access to various different kinds of linguistic information in one structure, the *sign*. This means that constraints involving both semantic and syntactic information are easy to define. For the resolution of fragments, however, just the syntax and semantics of the fragment alone are not enough, obviously, and further information must be taken from the wider linguistic context: for example from the question to which the fragment is a short answer. Contextual information of this kind is standardly not present in HPSG signs, and hence, as described in the previous section, (Ginzburg & Sag 2001) introduce a representation of the linguistic context into HPSG feature structure (descriptions) by integrating elements of KOS. How this contextual information is utilised in the resolution of fragments will now be shown first with the example of short answers.[11] We deal separately with their treatment of short answers to *argument questions* (that are questions where the *wh*-element fills a complement position, as in "who saw Mary?"), and to *adjunct questions* (where the *wh*-element fills an adjunct position, e.g. "when did Peter see Mary?").

### 5.3.1 Short Answers to Argument-Questions

(Ginzburg & Sag 2001) represent question meanings with (the HPSG equivalent of) λ-abstracts. The attribute-value-matrix (171-b) shows the situation semantics equivalent of the first-order representation (171-c) (leaving out the representation of "Mary"), both representing the meaning of (171-a). The feature PARAMS in these AVMs collects what in the first-order representation are the λ-bound variable(s), whereas the effect of using variables in first order logic is achieved by re-entrancy (e.g., [1] below is the value both of the INDEX in PARAMS and of the SEE-ER feature of the predicate). *Wh*-words can additionally impose semantic (sortal) restrictions—in (171-b) this is the restriction that the index must be that of an animate entity.

[11] The description is mainly based on (Ginzburg & Gregory 2001), which uses a simpler semantic ontology than (Ginzburg & Sag 2001), but shares the same overall architecture.

*bare-arg-ph & bare-decl-cl:*

$$\left[\begin{array}{ll} \text{CAT} & s \\ \text{CONT} & \left[\begin{array}{ll}\text{SOA}\mid\text{NUCL} & \boxed{1}\\ \text{QUANTS} & \boxed{2}\oplus\boxed{3}\\ \text{RESTR} & \boxed{4}\cup\boxed{5}\end{array}\right] \\ \text{CTXT} & \left[\begin{array}{ll} \text{MAX-QUD} & \left[\begin{array}{ll}\text{PARAMS} & \left\{\boxed{6}\right\}\\ \text{SOA} & \left[\begin{array}{ll}\text{QUANTS} & \boxed{3}\\ \text{NUCL} & \boxed{1}\end{array}\right]\end{array}\right]\\ \text{SAL-UTT} & \left[\begin{array}{ll}\text{CAT} & \boxed{7}\\ \text{CONT} & \boxed{6}\left[\begin{array}{ll}\text{INDEX} & \boxed{8}\\ \text{RESTR} & \boxed{4}\end{array}\right]\\ \text{STORE} & \{\}\end{array}\right]\end{array}\right]\end{array}\right] \rightarrow \mathbf{H}\left[\begin{array}{ll}\text{CAT} & \boxed{7}\ \mathit{nominal}\\ \text{CONT} & \left[\begin{array}{ll}\text{INDEX} & \boxed{8}\\ \text{RESTR} & \boxed{5}\end{array}\right]\\ \text{STORE} & \boxed{2}\ \mathit{set(qf)}\end{array}\right]$$

Figure 5.2: The construction type *short-answers*

(171) a. Who saw Mary?

b.
$$\left[\begin{array}{ll}\text{PARAMS} & \left\{\left[\begin{array}{ll}\text{INDEX} & \boxed{1}\\ \text{RESTR} & \left\{\mathit{animate\text{-}rel}(\boxed{1})\right\}\end{array}\right]\right\}\\ \text{SOA} & \left[\begin{array}{ll}\text{QUANTS} & \langle\rangle\\ \text{NUCL} & \left[\begin{array}{ll}\multicolumn{2}{l}{\mathit{see\text{-}rel}}\\ \text{SEE-ER} & \boxed{1}\\ \text{SEEN} & \boxed{2}\end{array}\right]\end{array}\right]\end{array}\right]$$

c. $\lambda x.[animate(x) \wedge see(x,m)]$

Now, if according to KOS this question (171-a) is the ‘question under discussion’ at a certain point in a discourse, then (171-b) is consequently accessible as the value of CONTEXT|MAX-QUD. A fragmental answer to this question then is resolved by a special *construction type*, whose specification is shown in Figure 5.2.[12] This unary ‘rule’ takes a nominal daughter to a sentential (i.e. verbal with all valence-requirements saturated) phrase.[13]

[12]For the purposes of presentation, the specification is given in the form of a phrase structure rule, but keep in mind that in fact it is just a constraint on feature structures.

[13]Although we use the simpler semantic ontology of (Ginzburg & Gregory 2001) in the presentation here, we divert from the presentation there in that we define the daughter to be a head-daughter, as in (Ginzburg & Sag 2001). This means that a default or ‘generalised’ head-feature principle has to be assumed, which copies over from the HEAD of the daughter to that of the mother only values that aren’t explicitly specified on the mother.

The workings of this rule is best explained with an example. Figure 5.3 shows an instance of this rule; it presents the sign for an utterance of ‘John’ as a reply to the question ‘Who saw Mary?’. The value for CAT of the NP ‘John’, [7], is identified with the value for the CAT of the SAL-UTT. The value of SAL-UTT is specified as “the wh-phrase utterance associated with the PARAMS set of MAX-QUD” (Ginzburg & Sag 2001, p.301), i.e. here it is “who”. Through this co-indexation, the syntactic matching condition between fragment and question is enforced. The semantic resolution is done by identifying the INDEX of the fragment, [8], with that of the SAL-UTT, whose CONTENT in turn is one of the PARAMS of the question under discussion. This means that the index of the fragment ‘John’ is identified with the SEE-ER role of the question (which was held in PARAMS, or, in other words, was $\lambda$-abstracted over). This amounts to functional application plus $\beta$-reduction. The content of the resolved short answer finally contains this matrix verb of the question, plus an amalgamation of the restrictions from the question and the new restrictions the fragments brings with it.

As this has shown, the claim made by the GBA is that interrogatives, or, more specifically, their *wh*-phrase, and short answers show *categorial congruence*. *Wh*-interrogatives are analysed in (Ginzburg & Sag 2001) as extraction constructions, and, beginning with (Pollard & Sag 1994, Ch.9), and further developed in (Sag 1997) and (Bouma, Malouf & Sag 2001), extraction in HPSG roughly works by using ‘realisation rules’ that decrease the valency requirement of verbs and in place of the non-realized complement put a special kind of sign, a so-called ‘gap-synsem’, on the argument structure of the matrix verb. This gap gets filled ‘higher up’ in the structure by a constraint which requires filler and gap to be co-indexed. This ensures that the filler gets its category specification from the subcat-frame of the verb. And it means in our case that ultimately this, the subcat-frame, is where the constraints on the form of admissible short answers (to argument-questions) come from.

We are now in a position to explain how the GBA rules out the PP as short answer in (170-b) above, repeated here as (172).

(172) A: Who did you see?
B: #On Sandy.

Since the rule in Figure 5.2 requires that the category of the fragment and that of SAL-UTT be identical (see coindexation [7]), it rejects ‘on Sandy’ in that example as a short answer to the question, even though the *content* is possibly identical to that of ‘Sandy’: being a PP[on], ‘on Sandy’ simply doesn’t meet the *syntactic* requirements.[14]

Finally, the constraint in Figure 5.2 makes another prediction. If there are quantifiers present in the fragment, they will always end up outscoping those present in the question. This means that the GBA

[14] Note that ‘on Sandy’ therefore is not regarded as a well-formed fragment at all in this case; the GBA doesn’t make a difference between fragments and their *use* (as short answers for example), as we will discuss in detail below in Section 5.5. In that sense, the GBA is bound to regard ‘on Sandy’ in that context as an ungrammatical expression.

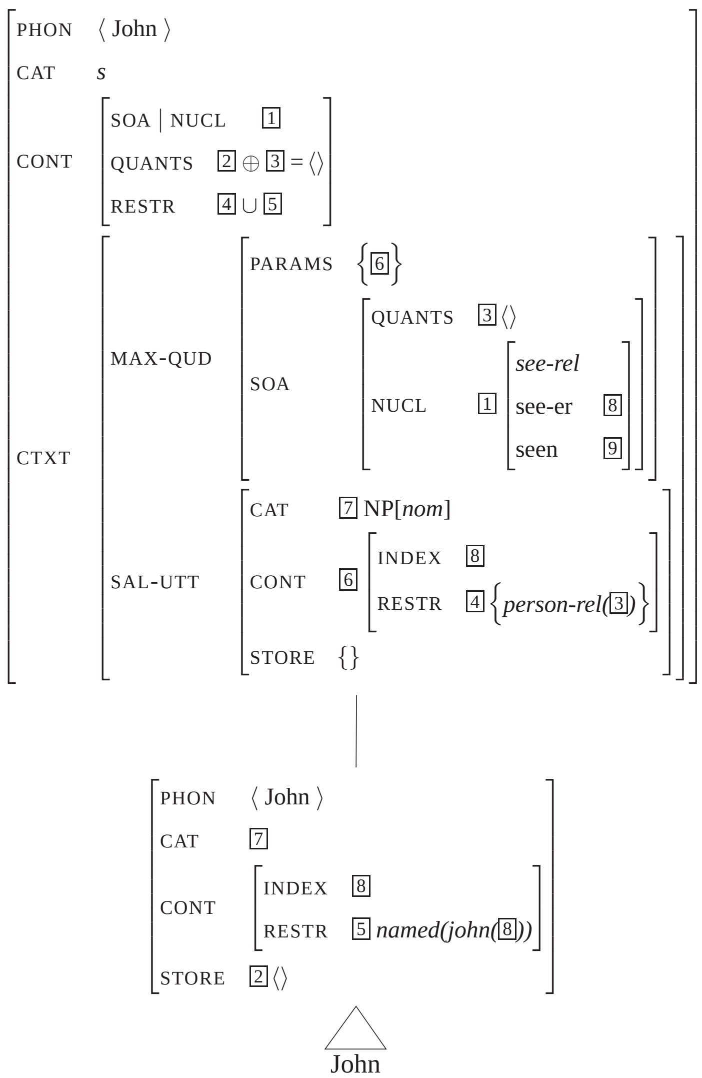



Figure 5.3: ‘John’ as an answer to ‘Who saw Mary?’

stipulates a semantic difference between short-answers and what *prima facie* seems to be a full-sentence paraphrase. For example, in (173), GBA predicts the full answer (173-b) to have a reading that (173-a) doesn't have, namely one where there is a Christmas carol that is hated by everyone. We will discuss this prediction below in Section 5.5.

(173) A: Who hates a Christmas carol?

a. B: Everyone.
b. B: Everyone hates a Christmas carol.

This concludes the description of the GBA account of short-answers to argument-questions; it forms the basis for the treatment of the other kinds of fragments we will describe in the rest of this section.

### 5.3.2 Short Answers to Adjunct Questions

The GBA to short-answers to adjunct questions is not as well developed as that detailed in the previous section. (Ginzburg & Sag 2001) explicitly excludes short answers to adjunct questions from their investigation, and only (Ginzburg & Gregory 2001) offers some ideas as to how to treat them.

In general, (Ginzburg & Gregory 2001) claims, the mechanism described above can be re-used. Adjunct questions are represented in their approach as shown in (174). A short answer to such a question has to provide an adjunct relation (ADJ-REL), which takes the main relation of the antecedent question in MAX-QUD as argument, resulting in the desired resolution.

(174) a. When did John see Mary?

b.

$$\begin{bmatrix} \text{PARAMS} & \left\{ \begin{bmatrix} \text{INDEX} & \boxed{1} \\ \text{RESTR} & \textit{time}(\boxed{1}) \end{bmatrix} \right\} \\ \text{SOA} & \begin{bmatrix} \text{ADJ-REL} & \boxed{1} \\ \text{SOA-ROLE} & \begin{bmatrix} \textit{see-rel} & \\ \text{SEE-ER} & \boxed{2} \\ \text{SEEN} & \boxed{3} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

There is a problem here, however, as (Ginzburg & Gregory 2001, p.10) acknowledge: "the categorial parallelism requirements associated with adjuncts [...] appear to be somewhat freer than with arguments". They give the following example, where a *when*-question can be answered with a PP, an NP or an adverb:

(175) A: When did Joe leave?
B: At 2 o'clock / Yesterday / Recently.

(Ginzburg & Gregory 2001) do not discuss this further; one obvious solution in their framework, however, would be to claim that the *wh*-word "when" is syntactically ambiguous, making the question ambiguous. The short-answer then resolves this ambiguity.

While this could work for this particular kind of adjunct (temporal), it seems that this strategy would still undergenerate for other kinds of adjunct-questions, such as the *why*- or *how*-questions shown in (176). The problem here is not the syntactic matching, but rather the fact that semantic material that resolves the fragmenty is not linguistically explicit in the context at all, as we will argue below.

(176) a. A: Why did Joe leave? — B: Exams tomorrow.
b. A: How can I get downtown? — B: Bus number 14.

## 5.4 GBA and short-questions

Ginzburg *et al.* also offer an approach to 'short questions', i.e. fragmental interrogatives as shown in (177) (capitalisation in this example denotes intonational stress).

(177) A: I met several of your students.
a. B: Who?
b. B: WHO?
c. B: When?

The approach basically uses the same techniques as described in the previous section, namely resolution via MAX-QUD and SAL-UTT. Given that it is not at all obvious what the *question under discussion* could be relative to which these fragments are resolved, in this approach the problem of resolving these fragments is shifted to that of accommodating appropriate questions onto the QUD. As far as we can see, Ginzburg and collaborators allow three basic ways in which this can happen.[15] Which way leads to the desired MAX-QUD depends on the type of *sluice* (as Ginzburg *et al.*, following (Ross 1969), also call this type of fragment).

The first type we will discuss here is what the authors call the *direct sluice*, instances of which in the example above would be (177-a) and (177-c). Contrasted with this is the *reprise sluice*, as in (177-b).

[15] The presentation of the GBA to this kind of fragments is less well systematised, and scattered over several papers with partially conflicting terminology (Ginzburg 1997, Ginzburg 1999*a*, Ginzburg & Cooper 2001). Hence what we present here is our attempt at a rational reconstruction of the *classification* that we think is attempted; in the technical details the presentation follows (Ginzburg & Cooper 2001).

The difference between them is whether the utterance to which the fragmental question is a reply has been fully understood and only additional information is requested, or whether it has *not* been fully understood and some elements are being clarified.[16] In the example above, full comprehension seems to be indicated by (177-a) and (177-c), but not by (177-b).

In the class of *reprise sluices*, Ginzburg makes a further distinction between *constituent readings* and *clausal readings*. Fragmental questions like the one shown in (178) below, so Ginzburg claims, are ambiguous between a reading that can be paraphrased as "Who is that Bo person?" (the constituent reading) and another reading "Did you just say that BO (of all people) called?" (the clausal readings).[17]

(178) A: Bo called.
B: BO?

Since the resolution of the *reprise sluices* relies on more complex operations, and also requires rather drastic additional changes to the standard inventory of HPSG, we will postpone their discussion and begin with the more straightforward type of *direct sluices*.

### 5.4.1 Direct Sluices

Within the class of direct sluices, it is useful to also distinguish between those referring to *arguments*, as (179-a) and (179-b) below do, and those referring to *adjuncts*, like (179-c).

(179) A: A thesis was approved.

a. B: Which one?
b. B: By the committee?
c. B: When?

As we will see, only the first type can be dealt with relatively naturally in the GBA.

Now, what is the question on MAX-QUD that allows the techniques from the previous section to be re-used, and how do these techniques have to be adapted? According to KOS, assertions of a proposition *p* trigger the accommodation of a polar question '*whether p*'; e.g. for (179) this would be something para-

[16] Hence, Ginzburg *et al.* also call fragments of the latter type *clarification ellipsis*. We have commented above in Section 2.4 already on the fact that the GBA often seem to mix classifications referring to *form* (like 'sluice') with those referring to *use* (like 'clarification ellipsis').

[17] While we agree with the judgement that there is a constituent reading, it seems to us that the way the authors paraphrase the 'clausal reading' conflates two issues. First, such a reading can express simply surprise at the content (as in "BO? But I thought she's away?") but also something like "did I *hear* you correctly?" The authors give conflicting information about whether such readings are supposed to be included or not. (Ginzburg & Cooper 2001) states that this type of fragment can be instantiated by "any apparently co-referring phrase"; for example in (178) something like "my cousin?". Earlier in the same paper, however, they give an example where it is more likely that acoustic understanding is the issue. We will return to this later.

$$\begin{bmatrix} \textit{slu-int-cl} & \\ \text{CONT} & \begin{bmatrix} \textit{question} & \\ \text{PARAMS} & \boxed{\Sigma_1} \\ \text{PROP} & \begin{bmatrix} \textit{proposition} & \\ \text{SIT} & \boxed{1} \\ \text{SOA} & \begin{bmatrix} \text{QUANTS} & \boxed{A} \\ \text{NUCL} & \boxed{3} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{STORE} & \{\} \\ \text{MAX-QUD} & \begin{bmatrix} \textit{question} & \\ \text{PARAMS} & \{\} \\ \text{PROP} & \begin{bmatrix} \textit{proposition} & \\ \text{SIT} & \boxed{1} \\ \text{SOA} & \begin{bmatrix} \text{QUANTS} & \langle \textit{pquant-rel} \rangle \oplus \boxed{A} \\ \text{NUCL} & \boxed{3} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \rightarrow \mathbf{H} \begin{bmatrix} \text{H-DTR} & \begin{bmatrix} \text{STORE} & \boxed{\Sigma_1}\textit{neset(param)} \end{bmatrix} \end{bmatrix}$$

Figure 5.4: The type *slu-int-cl*

phraseable as "Was a thesis approved?". Given this question on MAX-QUD, the fragment can be resolved by the type-constraint *slu-int-cl* (*sluice-interrogative-clause*), which is shown below in Figure 5.4.

In this constraint, SAL-UTT is identified with a "sub-utterance" of the utterance that is being replied to and about which more information is requested, and the fragment phrase (in this case, a *wh*-phrase) supplies the parameter for the resulting content, which is that of a *wh*-question. Put differently, what this construction achieves is simply λ-abstraction over an argument which is picked out by (or for which must hold) a syntactic matching condition. An example shall make this clearer.

Figure 5.5 shows the resolution of the fragment in "A student called. — Who?" according to the GBA.[18] As explained above, the MAX-QUD is the polar question (i.e., a question with an empty set of parameters) "did a student call?", and the SAL-UTT is the sign of the part of the utterance giving rise to the content *some student* (i.e., "a student"). In (Ginzburg & Sag 2001), the restriction on the *wh*-element is taken from the fragment, and hence the content of the resulting clause can be paraphrased as "who called?"; in (Ginzburg & Gregory 2001), on the other hand, the restriction is copied over from the content of the SAL-UTT, and hence in this case the resolved content would be "which student called?".

[18]This figure is Ginzburg & Sag's (2001) (62) on page 326.

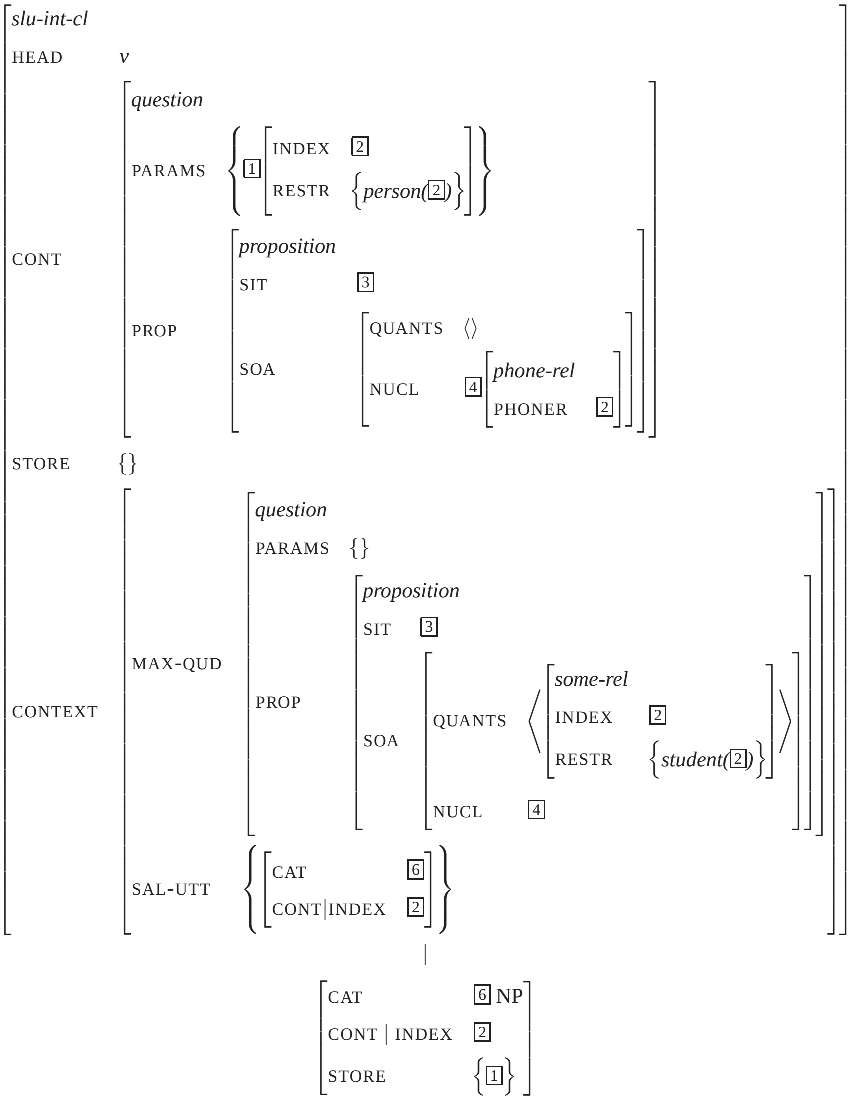



Figure 5.5: Resolution of “A student phoned. — Who?”

The use of SAL-UTT in this constraint leads to two obvious problems, which we have already alluded to above: First, this approach cannot straightforwardly explain direct sluices like (179-b) above, repeated here as (180).

(180) A: A thesis was approved.
B: By the committee?

The problem is that for such fragmental questions there is no explicit antecedent which could be the SAL-UTT; they inquire about optional arguments of the previous sentence. (Ginzburg & Sag 2001, p.330) note this problem, and speculate on a solution in which optional arguments are represented by some special kind of sign, which can then be picked up by SAL-UTT, but they don't follow up on this (rather unattractive) solution.

The second problem is related to this, and concerns what one might call adjunct-questions as in (179-c) above, repeated here as (181).

(181) A: A thesis was approved.
B: When?

Again, the problem is that there is no antecedent that could be the SAL-UTT. (Ginzburg & Sag 2001) explicitly exclude this kind of fragment by restricting their approach to nominal fragments, but (Ginzburg & Gregory 2001) point at ways in which the approach could be extended to deal with this. Their solution is the addition of another clause-type which does not use SAL-UTT at all and simply puts the appropriate adjunct-relation in the parameters set, wrapped around the propositional content of the MAX-QUD, as it were. We forgo showing the type-specification here and only note that the detour via MAX-QUD seems especially superfluous for this kind of utterance.

We will present later an approach which does not have these problems and does not have to assume empty elements for optional arguments but still can explain syntactic phenomena like the case-concord in the following German example.

(182) A: Wir haben gestern ordentlich gesoffen.
B: [$_{NP}$ Den guten Likör]$_{ACC}$?
(roughly: *We drank properly yesterday. — The good liquor?*)

### 5.4.2 Reprise Sluices

We have mentioned above that the GBA distinguishes between a *clausal* and a *constituent* reading of reprise sluices. Corresponding to these two types of reading, (Ginzburg & Cooper 2001) define two

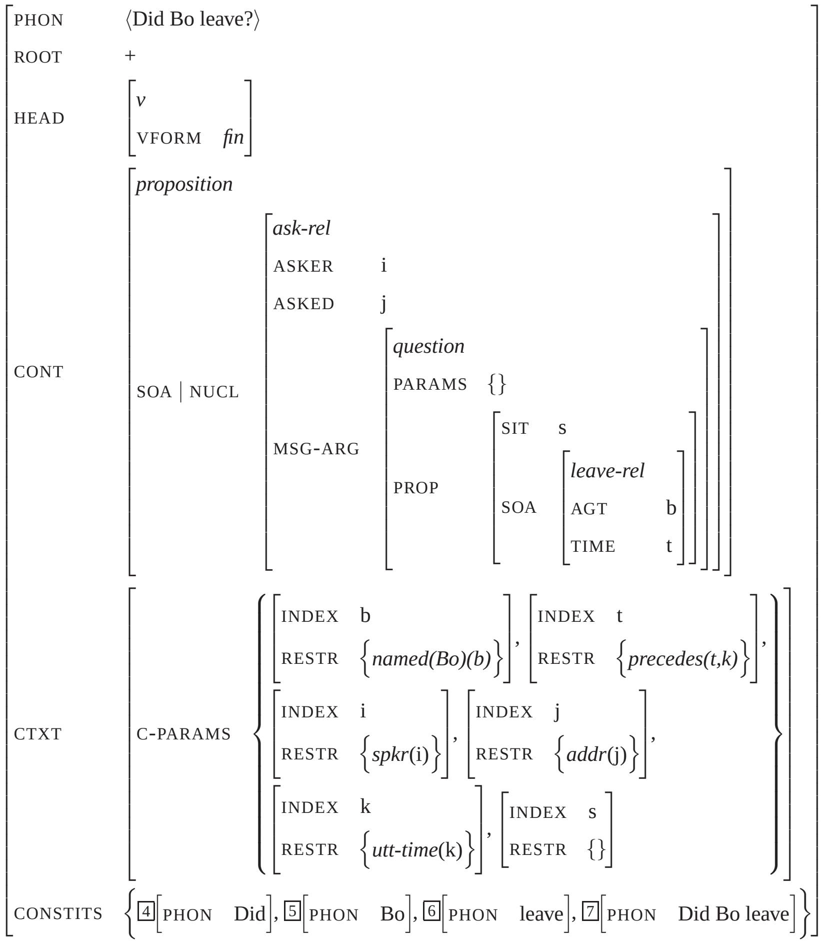



Figure 5.6: The sign for "Did Bo leave?" according to (Ginzburg & Cooper 2000)

'coercion'-operations on signs which make available appropriate questions on MAX-QUD. Before we can describe these operations, however, we have to explain the significant additions to the repertoire of HPSG the authors assume.

The three most drastic changes are illustrated by the sign for "did Bo leave" shown in Figure 5.6. First, all complete utterances are now construed to contain an illocutionary force operator, whose argument is what in more standard versions of HPSG is the content of the utterance. In this example this means that the content of the interrogative now is something paraphraseable as "I ask you whether Bo left." Secondly, in a more technical addition, the authors assume a feature that holds the signs of all constituents of a clause, not just its immediate ones which can be found in (NON)HEAD-DAUGHTER(S). This is required because all sub-elements of a clause can be clarified and must be accessible, as we will see shortly.[19]

Finally, the contextual feature C-INDICES is promoted to hold 'parameters' for most context-sensitive elements of an utterance—proper names, deictic pronouns, indexicals.[20] The task of a hearer then, as Ginzburg and Cooper claim, is to instantiate these parameters. Partial failure to do so can be remedied by asking a clarification question, and the potential to do so is modelled by the coercion operations to which we now turn.

As we said above, Ginzburg assumes two readings for a question like the one in (183); one that clarifies the reference of a constituent ("Who is that Bo person?", the constituent reading) and one that clarifies whether the utterance was comprehended properly ("Did you ask whether BO called?").

(183) A: Did Bo call?
B: BO?

We begin with Ginzburg & Cooper's (2001) analysis of the 'constituent reading'. The MAX-QUD required to resolve the fragment in this way is produced by an operation on the sign of "Did Bo call?" which the authors dub 'parameter identification'. (We will discuss below the theoretical status of such rules.) The respective rule is shown in (184); it is read as "if an utterance whose sign unifies with the left-hand side of the rule was produced, then an appropriate response is an utterance which unifies with the right-hand side of the rule."[21]

(184) Parameter identification:

$$\begin{bmatrix} \text{CTXT} \mid \text{C-PARAMS} & \left\{ \ldots \boxed{i} \ldots \right\} \\ \text{CONSTITS} & \left\{ \ldots \boxed{2} \begin{bmatrix} \text{CONT} & \boxed{i} \end{bmatrix} \ldots \right\} \end{bmatrix}$$

$$\Rightarrow$$

[19] The feature is also needed for the direct sluices discussed in the previous section, where a short-question like (i) can ask for an elaboration of the content of the embedded clause or of that of the matrix clause.

(i) A: John said that Mary has cheated.
B: When?

[20] These changes are all rather controversial, but we postpone discussion to Section 5.5 below and continue here with the presentation of the approach.

[21] As an aside, these rules could be re-interpreted as speech act recognition rules, in the sense that they give sufficient conditions on (form and content of) two utterances to infer that the second is a clarification of the first.

$$\begin{bmatrix} \text{CONTENT} \mid \text{MSG-ARG} & \begin{bmatrix} \textit{question} & \\ \text{PROP} & \boxed{3} \end{bmatrix} \\ \text{SAL-UTT} & \boxed{2} \\ \text{MAX-QUD} & \begin{bmatrix} \textit{question} & \\ \text{PARAMS} & \left\{ \begin{bmatrix} \text{INDEX} & \boxed{4} \end{bmatrix} \right\} \\ \text{PROP} & \boxed{3} \begin{bmatrix} \text{SOA} & \begin{bmatrix} \textit{content-rel} & \\ \text{SIGN} & \boxed{2} \\ \text{CONT} & \boxed{4} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

This rule grabs a constituent of the utterance ($\boxed{2}$ in (184)) whose content ($\boxed{i}$) it offers for clarification by syntactically parallel fragments (since the sign is put on SAL-UTT). In other words, it licenses a question with the content "what is the content of the sign $X$, which was part of your previous utterance?" as a follow-up to an utterance—for (183) this would be, as intended, something like "what is the content of 'Bo' in your previous utterance?".[22] (Note the introduction of a *content-rel* for this purpose.)

The coercion rule for the clausal reading makes use of the presence of illocutionary force relations in the content of clauses; it is shown in (185).

(185) Parameter focussing:

$$\begin{bmatrix} \text{CTXT} \mid \text{C-PARAMS} & \boxed{1} \left\{ \ldots \boxed{i} \ldots \right\} \\ \text{CONSTITS} & \left\{ \ldots \boxed{2} \begin{bmatrix} \text{CONT} & \boxed{i} \end{bmatrix} \ldots \right\} \\ \text{CONTENT} & \boxed{3} \end{bmatrix}$$
$$\Rightarrow$$
$$\begin{bmatrix} \text{CONTENT} \mid \text{MSG-ARG} & \begin{bmatrix} \textit{question} & \\ \text{PROP} & \boxed{3} \end{bmatrix} \\ \text{SAL-UTT} & \boxed{2} \\ \text{MAX-QUD} & \begin{bmatrix} \textit{question} & \\ \text{PARAMS} & \left\{ \boxed{i} \right\} \\ \text{PROP} & \boxed{3} \end{bmatrix} \end{bmatrix}$$

The content of questions licensed by this constraint is quite literally "did you just ask/assert/command whether/that $p$?", where in $p$ the clarified element is focussed—for (183) this would be "did you just

[22] We forgo showing which type builds this question out of the fragment phrase and MAX-QUD and SAL-UTT.

ask whether BO left?".[23] The rule achieves this by making the content of the antecedent sign (recall that this will always be a proposition involving an illocutionary force relation) the 'body' of MAX-QUD and one of the sub-constituents the parameter; i.e., it abstracts over this sub-constituent.

(Ginzburg & Cooper 2001) also discuss differences in the constructions needed for constituent readings expressed by repetition sluices as in our example or by *wh*-phrases, but the general principle is the same, and so we conclude our description of GBA and come to the discussion of its merits and shortcomings.

## 5.5 Discussion

In the following sections we highlight the strengths and weaknesses of the GBA.

### 5.5.1 Syntactic *Parallelism*?

Both in (Ginzburg & Sag 2001) and in (Ginzburg & Gregory 2001), fragments are restricted to be nominal phrases (see the restriction on CAT in Figure 5.2). While this category includes prepositional phrases, we have so far only seen examples of NP fragments, and indeed this is what Ginzburg *et al.* mostly discuss. In this subsection we will see what predictions are made by the categorial congruence condition about the behaviour of prepositional phrases and other kinds of fragment-phrases. The data used here is mostly repeated from the section on Morgan's approach in the previous chapter.

#### 5.5.1.1 Prepositional Phrases

Example (186) shows short-answers to interrogatives which ask for the NP object of a prepositional complement. In these interrogatives, the preposition is left 'stranded', i.e. only the *wh*-element has been extracted out of the PP. This is, at least for non-formal registers, the standard form of such interrogatives in English.

(186) a. A: Who can we rely on? { B: On Sandy. / B: Sandy. }

b. A: Who shall we give the money to? { B: To Kim. / B: Kim. }

c. A: What did the message appear on? { B: On the screen. / B: The screen. }

[23] Again we forgo showing which clause-type composes this content out of fragment and context.

It seems that as short answers to such interrogatives, both NPs and PPs are acceptable, regardless of whether the preposition is 'case-marking' as in (186-a) and (186-b) or 'predicative' as in (186-c).[24] (We have already seen examples of this kind.)

Now recall that GBA requires categorial identity between short-answer and an element of the antecedent, through token identity of SAL-UTT and fragment phrase. But what exactly is this 'element of the antecedent' that is on SAL-UTT? The value of this feature is specified informally in (Ginzburg & Sag 2001, p.301) as "the *wh*-phrase utterance associated with the PARAMS set of MAX-QUD". The data in (186) above suggest that one has to allow some degrees of freedom in what counts as the *wh*-phrase that is 'associated'—either just the NP-argument of the preposition, or the whole PP—to explain the optionality and rescue the syntactic matching condition.

Interestingly, the condition of categorial identity seems stronger for free relatives, which otherwise have the same distribution as NPs:

(187) A: What does it depend on? { B: On what they decide<br>B: ?#What they decide }

While Ginzburg & Sag (2001) do not discuss interrogatives like the ones above, they do talk about interrogatives where the whole PP has been 'pied piped' (Ross 1967). They make two claims, both of which, however, seem to be too strong:

- "predicative prepositions [...] do not appear in pied piping structures" (Ginzburg & Sag 2001, p. 196).
- interrogatives with pied-piped case-marking prepositions accept only PPs as short answers. (Ginzburg & Sag 2001, p.301, n.9)

Example (188) is taken from (Ginzburg & Sag 2001, p. 196), where the question is marked as ungrammatical. My informants seem to disagree there; while they do find the question sounding rather formal they nevertheless do not regard it as ungrammatical:

(188) A: For whom does the museum run special tours? — { B: For young children.<br>B: Young children. }

The following examples are taken from (Ginzburg & Sag 2001, p.301, n.9) where the NP-answer is marked as ungrammatical; according to my informants again both variants of the answer are acceptable:

[24] Generally, my informants were in agreement that always both variants are acceptable, even though there seem to be preferences, depending on whether the question is pied-piped or not; this might be explained with 'syntactic priming' effects (Branigan, Pickering & Cleland 2000). For example, (Levelt & Kelter 1982) found that subjects tend to mirror the surface form of the question when asked either "what time do you close?" or "at what time do you close?".

(189) A: To whom did you give the book? — { B: To Joe / B: Joe }

(190) A: On what does the well-being of the EU depend? — { B: On a stable currency. / B: A stable currency. }

Finally, we should point out that this optionality holds not only for short-answers, but apparently for sluices as well, as (191) shows.[25]

(191) A: Peter relies on Sandy.
B: On who? / Who?

To summarise, it seems that interrogatives which ask for objects of prepositional phrases do in general, be they pied-piped or not, allow both NP- and PP-short-answers, and that hence the syntactic matching condition must be relaxed accordingly.[26] As already mentioned above, this is in contrast to German, where this optionality is not licensed.

(192) A: Auf wen können wir uns verlassen? { B: Auf Sandy / B: #Sandy }

#### 5.5.1.2 Sentential Complements

Sentential complements pose another problem for the syntactic matching condition.[27] First of all, the GBA has to assume that "what" in certain constructions (e.g. (193-a) below) can be of category S[*comp*], to retain the assumed syntactic parallelism. Further, a number of verbs that take sentences as complements accept both full sentences and sentences with a complementizer, e.g. (193-a).[28] However, if those verbs are used in an interrogative, a short-answer can only be a complement sentence, as (193-b) shows. Hence, without further changes, the syntactic matching condition would overgenerate here.

(193) a. They believed { that they will be reassigned / they will be reassigned }

b. A: What did they believe? — { B: That they will be reassigned. / B′: #They will be reassigned. }

[25]Interestingly, there seems to be an influence of register on optionality: if the, nowadays more formal, variant 'whom' is chosen, then the PP variant ("on whom?") seems to be much preferred over the NP variant ("whom?").

[26]Either through allowing flexibility in determining SAL-UTT, or through allowing a map between PP and NP.

[27]Note that, as mentioned above, (Ginzburg & Sag 2001) and (Ginzburg & Gregory 2001) explicitly restrict the scope of their theory to nominal short answers. It will nevertheless be interesting to see how the approach could be extended beyond this kind of short answer.

[28]Why and for which verbs this is so is not yet well understood; cf. (Bolinger 1972) for an early study of this phenomenon.

However, there are other constructions besides short answers in which the complementizer becomes compulsory:

> *"It is commonly observed that in numerous syntactic environments a clause must appear with a complementizer. This set of environments includes at least sentence fragments, subject position, and the focus position in various copular constructions." (Pollard & Sag 1994, p.125)*

Since (Ginzburg & Sag 2001) analyse interrogatives as extraction constructions similar to focussed clauses, any solution for the mismatch in focus-construction would also solve this for the syntactic matching condition.

#### 5.5.1.3 Category Mismatches

It has often been noted that some extraction constructions allow different syntactic categories in extracted positions than in non-extracted positions, or to phrase it more technically, allow the filler to be of a different category than the gap. These 'movement paradoxes' have been extensively discussed in the LFG literature,[29] with examples like the following:

(194) a. (i) *That he was sick* we talked about for days.
(ii) *We talked about *that he was sick* for days.
((Higgins 1973), quoted in (Bresnan 1994))
b. (i) *That languages are learnable* is captured by this theory.
(ii) *This theory captures *that languages are learnable*.
((Grimshaw 1982), quoted in (Bresnan 1994))

These examples can be directly adapted to the question-answer environment (as (Morgan 1973) already noted, see discussion above in Section 4.2).

(195) a. A: What did you talk about for days? — B: That he was sick.
b. A: What does this theory capture? — B: That short answers are weird.

Again, one can speculate about possible ways to rescue the syntactic congruence condition. First, one could locate this phenomenon in the operation of extraction, and somehow allow a filler to be an S[*comp*] even though the gap wants an NP, coercing the S[*comp*] to denote a fact. This would capture the data in (194) as well. The second possibility is to relax the matching to allow an NP requirement to be met by an S[*comp*], similarly restricted to verbs selecting for facts. Lastly, one could argue that these kinds of complement-sentences are really NPs, which then however must be restricted not to occur in certain positions. We will not decide on any of the possibilities and simply note that examples like this don't necessarily contradict the syntactic matching condition.

[29] See for example (Bresnan 1994, Bresnan 2001).

#### 5.5.1.4 Short Answers to Adjunct Questions

We have already briefly mentioned in Section 5.3.2 Ginzburg & Gregory's (2001) proposal for a treatment of adjunct questions. We noted that the extent of syntactic parallelism in these cases is not clear, but in any case less strict than with argument-questions. The following example showed this.[30]

(196) A: When did Joe leave?
B: At 2 o'clock / Yesterday / Recently.

This flexibility could be taken as evidence that for adjunct questions there is no syntactic influence at all, and the only requirement is semantic. While this might explain (196), it is not enough to deal with short answers to adjunct questions likes those in (197-a). As discussed at length above, in these cases *semantic* material seems to be 'missing' as well; function application of the question to its fragment answer is too simplistic to resolve the content of these short answers, and more powerful reasoning is required. We will return to this issue presently.

(197) a. A: Why did Joe leave? — B: Exams tomorrow.
b. A: How can I get downtown? — B: Bus number 14.

(198) a. (i) *John left exams tomorrow.
(ii) John left *because he has to (take? supervise? ... do some unspecified action on)* exams tomorrow.
b. (i) *You can get downtown bus number 14.
(ii) You can get downtown *by taking* bus number 14.

### 5.5.2 Extensibility

In this section we discuss whether the *techniques* used by the GBA could be adapted to deal with kinds of fragments not yet discussed; we will still not say anything about whether we think these techniques are theoretically sound. We begin with a look at kinds of short-answers we have neglected so far, and then briefly look at speech acts other than answering and elaboration / clarification.

#### 5.5.2.1 Other types of short answers

There is a type of short answer that has been given short shrift so far, namely VP-answers as in (199) below.

[30] Interestingly, the PP/NP problem shows up here again; "2 o'clock" is also fine as a short answer in (196), although it isn't in its full-sentence counterpart "*Joe left 2 o'clock".

(199) A: What did Peter do?
B: Wash the car.

Again, both (Ginzburg & Sag 2001) and (Ginzburg & Gregory 2001) do not deal with fragments of this kind. While we do not want to speculate here what a GBA to this kind of fragment would be, we point out here that the rule in Figure 5.2 would not suffice, since additionally to identifying the INDEX of the phrase with the abstracted-over variable in the question, the variable filling the subject-position of the VP in the fragment has to be identified with that in the question (B in (199) is only an answer if it is resolved so that Peter is the one who's washing the car). Recall that the GBA is committed to giving a fully specified semantics for the fragment, and hence cannot rely on discourse techniques for identifying antecedents of anaphora.

Finally, we have not discussed short answers to multiple *wh*-questions, as in (200).

(200) A: Who relies on whom?
B: Sandy on Carl, and Peter on Mary.

Apart from the required changes to the grammar rules, we think the techniques offered by the GBA are probably sufficient to resolve answers of this kind. However, an additional explanation would have to be found for the fact that multiple NP-answers do not seem to be licensed in English.[31]

(201) A: Who saw whom?
B: ?#Peter Sandy, and Carl Mary.

#### 5.5.2.2 Other Types of Speech Acts

In the formulation described above, the GBA only deals with short answers and short questions (used to elaborate and to clarify).[32] In this section we speculate whether it could offer an approach also to the other kinds of speech acts that can be performed with fragments which we listed in Chapter 2.

One severe limitation has already been noted: as a matter of principle, GBA can only handle what we called *res-via-id* fragments. More powerful reasoning that could infer content from the context and other information sources don't seem to be available to it. But even in the field of *res-via-id* fragments, the approach seems to miss possible generalisations. What Ginzburg *et al.* call 'direct sluices' (an example is repeated here as (202-a)) is called *Elaboration*$_q$ in our approach, to highlight the similarity between this kind of fragment and that shown in (202-b).[33]

[31] But note that we will also not attempt to model the constraint illustrated by (201).
[32] Extensions are very briefly described in (Fernández & Ginzburg 2002).
[33] Recall from Chapter 2 that $_q$-versions of relations connect a question to an antecedent in such a way that their answers would connect to that antecedent with the 'normal' version of the relation. We will see later when we formalise this semantics that we

(202) a. A: Peter's thesis was approved.
B: When?

b. A: Peter's thesis was approved.
B: Yeah. At midnight.

Adding fragment types like that illustrated by (202-b) to the GBA should in principle be possible. However, the new types would all require their own syntactic type, and so their addition would lead to a further extension of the grammar. The problem of course is that the GBA does not distinguish between the speech-act that is performed with an utterance and the utterance itself, or in other words, the use made of an utterance and the utterance itself; for the GBA, the speech act type is encoded in the grammar. (Even though Ginzburg *et al.* do not use the term, it is clear that answering and clarifying are the illocutionary effects of *speech acts*.) This observation leads us to the discussion of the theoretical assumptions made by the GBA.

### 5.5.3 Constraints on Scope

Ginzburg & Sag (2001, p.305) assume the following constraint on the scope ordering of resolved fragments: "if a fragment is or contains a quantifier, that quantifier must outscope any quantifier already present in the contextually salient question." (See earlier discussion on page 99.) However, examples like the following contradict this claim.

(203) A: What did every lawyer buy?
B: A flashy sports car.

In this example, a reading where the quantifier from the fragment outscopes that in the "contextually salient question" seems to be available—it is presumably even the preferred one.[34]

Following (Pollard & Sag 1994), the GBA uses Cooper storage (Cooper 1983) to deal with quantifier scope ambiguities. While this technique allows one to overcome the problem of having to postulate

can often infer those $_q$-relations using the same information as for the 'normal'-versions.

[34]Incidentally, this observation points to a further complication in extending the GBA to other speech acts. Consider the following short discourse.

(i) Every farmer bought a cow.
Well, some sort of farm animal.

Here again it seems reasonable to assume that the fragment can be resolved in such a way that the material from the fragment receives narrow scope. The speech act *Correction* isn't worked out in the GBA, but as mentioned in Section 5.2, the underlying discourse theory assumes that the declarative sentence "every farmer bought a cow" triggers an update of QUD with "what did every farmer buy?". What example (i) shows is that this update must be constrained in its scope order: any decision for a particular scope of the declarative sentence must fix the scope of the question, and via that the scope of the correction. We will later see that we can model this influence more directly without recourse to an accommodated question.

several different syntactic analyses for sentences containing more than one quantifier, it still means that the grammar associates several logical forms with such sentences. Here is not the place to discuss the problems of such an assumption (see for example (Kamp & Reyle 1991, Poesio 1994)), and we only point out that it seems odd to assume that something like the resolution of fragments is handled by the grammar, whereas a similar problem like the determination of the contextually appropriate quantifier scope isn't.

### 5.5.4 Theoretical Problems

In the last couple of subsections we have examined certain empirical predictions the GBA makes and explored whether GBA could possibly offer a unified theory of fragments. In these sections, we have argued from 'within' the GBA, in the sense that we accepted the general strategy of using CONTEXT for resolution. We will now discuss this set-up, starting with a look at its way of constructing the intended meaning of fragments in a non-compositional fashion. We will then return to a question we have so far not answered, namely where the contextual information could possibly come from, that is how the theory of dialogue, KOS, interfaces with an HPSG. Finally, we will discuss the repercussions adopting a GBA has on the overall design of a theory of dialogue interpretation.

#### 5.5.4.1 (Non-)Compositionality, and Speech Act Recognition

We have seen above how the content of a fragment, for example a short answer, is resolved in the GBA by combining in the grammar information from the context and from the fragment itself. This, by very definition, is a non-compositional form of meaning-construction. In the GBA, the meaning of an expression is *not* composed exclusively of the meanings of its parts considered in abstraction from particular occasions in which they are uttered; the whole point of extending the context feature was to get access to information about the occasion of utterance. Non-compositional meaning construction, however, conflates two tasks: that of determining the contribution of syntax to the meaning, and that of determining the contribution of the context.[35,36] In effect, this move blurs the boundaries between grammar and pragmatics. While this might be intended—one fundamental assumption in HPSG is that in theory it can encompass all kinds of linguistic knowledge—we will argue in this section that this is not an advantageous move, since grammar and pragmatics rely on different types of information that are best modelled with different logics.

Let's for the moment still assume that SAL-UTT and MAX-QUD are available for the fragment-rules. We

[35] Indeed, the way the GBA composes the content of fragments is strikingly similar to the non-compositional pronoun-rule in DRT (in the formulation of (Kamp & Reyle 1993); cf. the brief introduction to dynamic semantics in Chapter 8). There, pronouns in the input sentence get replaced during the DRS-construction with an accessible antecedent from the context (i.e. the DRS that has already been built); this construction thus equally hides from the logical form the specific contribution of the context to meaning.

[36] For a detailed general discussion of the merits of compositionality see e.g. (Groenendijk & Stokhof 1991), (Janssen 1997).

have already pointed out above that the fragment rules in the GBA implicitly encode information about the performed *speech act*: fragments in the GBA are always resolved as short-answers, sluices, corrections, etc. Due to the non-compositionality of the meaning composition, there is no independent notion of 'fragment-hood' in this approach; the construction already resolves them as "full" message-bearing entities. (And hence there is no distinction between infelicitous and un-grammatical fragments.) To reinforce this point, the fragment 'Peter' in (204-a) doesn't have a compositional semantic contribution *as a fragment* according to the GBA; semantically, it is indistinguishable from the full sentence in (204-b).

(204) A: Who will come?

a. B: Peter.

b. B: Peter will come to the party.

We would like to argue that this is problematic. A fragment is for example a short answer by virtue of the performed *speech act*. Answering a question (using either short or long form, non-sentential or sentential expressions) means performing a certain speech act, means (in the default case) *using* a declarative clause to *do* something. The alignment between linguistic form and performance of a speech act, however, is generally seen as the product of highly complex default reasoning; default reasoning that is too powerful to be expressed via default inheritance for example.[37] As pointed out above, in the GBA the grammar fixes the speech act of a fragment. Why this is a problem can be illustrated with the following mini-dialogue.

(205) a. A: Who was in the car with Sam?

b. B: Peter.

c. I mean, that's not the answer, I don't know who else was in the car, but it was Peter who I saw, not Sam.

In (205-c), B corrects herself, or rather a possible mis-interpretation of her communicative intentions behind (205-b). She explicitly corrects the impression that the speech act performed was one of answering (205-a), rather, it was one of correcting it. This shows that the reasoning that leads A to infer that (205-b) was a short answer must be defeasible; under the light of the new evidence in (205-c), A will have to revise his understanding of (205-b).[38]

This latter point is exemplified in (206) as well:

(206) a. A [to C]: Who did this?

b. B [to C]: Who saw it happening?

[37] See *inter alia* (Perrault 1990, Hobbs et al. 1993, Lascarides & Asher 1993).

[38] See also (McRoy & Hirst 1995) for a discussion of such speech-act misunderstandings.

c. C: Peter.

Let's say that after the exchange (206) took place, both A and B think that C has answered their question, but not that of the respective other. (Technically, this amounts to saying that they think that their own question respectively was QUD-maximal.) Then, if C's reply is not per chance an answer to both questions, there can be a point later in the conversation where one of them has to realize their mistake, which in the GBA would have to trigger a re*parse* of the fragment.

The GBA has to claim that utterances like (205-b) and (206-c) are ambiguous, i.e. that the grammar generates more than one reading for them. Similarly, in (207) the GBA has to assume that the fragment is three-way ambiguous (and not just that its use is); i.e. the GBA would generate three different questions.

(207) A: Peter sold Spartacus to Mary.
B: Who?

All this, however, shows that another module besides the grammar is necessary which filters out unwanted readings; this decision must be based on contextual information and be defeasible (i.e., refutable in the light of further evidence). Moreover, the readings must be discernible by what speech act they represent, since direct reference to that seems to be possible ("that's not the answer" in (205-c)).[39]

We will not discuss here the feasibility of such an approach. Suffice to note that it puts a heavy load on the grammar, which always has to produce all possible readings of a fragment (given SAL-UTT and MAX-QUD) and also on the pragmatics module, which then has to filter out readings on the grounds of whether the constructions used (which means the speech acts assumed to have been performed) seem appropriate. In a way, speech acts come into play twice here, once in the grammar where readings for all possible uses are generated, and once in the pragmatics, where one reading is chosen over the others by taking information from diverse (possibly extra-linguistic) sources into account. We will see in the coming chapters how our approach avoids this duplication of work by making the pragmatic task of determining the speech act responsible for finding a resolution of the meaning of the fragments.

This discussion has shown that it is not clear how a GBA can be integrated into a wider theory of dialogue interpretation. In any case it seems that a GBA doesn't make a pragmatic module superfluous; one would still need it to decide on the appropriate readings. In the next section we return to the question of how the pragmatic theory assumed by the GBA can be interfaced with the grammar (if at all).

[39] A related issue is the usefulness of explicitly representing illocutionary force relations in the grammar, as described in Section 5.4.2 above; we defer the discussion of this to a separate section below.

#### 5.5.4.2 What's in a CONTEXT?

In Section 5.3 we have seen how the resolution of short answers works in the GBA, given that a certain amount of contextual information is available as value of the feature CONTEXT. Just like (Ginzburg & Sag 2001, Ginzburg & Gregory 2001) do in their presentation, we have left one rather important question unanswered, namely where exactly this contextual information comes from. Or, to put it more pointedly, where does KOS live? Ginzburg *et al.* describe informally a procedure that computes the value for SAL-UTT from MAX-QUD, and the coercion rules from (Ginzburg & Cooper 2001) explain a little bit about how certain values for MAX-QUD are computed,[40] but how this fits in with the general set-up of an HPSG is left unexplained.

In the following, we will speculate about three possible ways of conceptualising what the contextual information means, and we will see how they lead to substantial theoretical problems. We will see how GBA requires a radical departure from traditional concepts of the division of labour between different modules of linguistic competence.

We begin with a reminder of what HPSG set out to be: it uses feature–structures as mathematical models of *types* of linguistic entities, i.e. words, phrases, sentences.[41] The admissible feature structures that are meant to model types of well-formed linguistic entities of a language are described by constraints formulated in a logic of typed feature structures. HPSGs are monostratal, in that the constraints can jointly constrain what is usually seen as different kinds of linguistic information, e.g. syntactic, semantic or phonological. No representation is derived by whatever operations from another representation; rather, they exists in parallel; the constraints imposed on them are order-independent.

With that in mind, we will now try to make sense of the use of CONTEXT in GBA. We will investigate three possible different interpretations of the set-up described above.

**Restrictions-on-use interpretation** The feature CONTEXT has previously been used in HPSG to collect "contextual constraints on the appropriate use of a linguistic expression" (Green 1996, p.2) or "appropriateness conditions associated with an utterance of a given type of phrase" (Pollard & Sag 1994, p.332) It has been used to account for pragmatic effects of the choice of certain words (German polite 'Sie', Korean honorifics (Pollard & Sag 1994)), register variation in English (Wilcock 1997)) and pragmatic effects of certain topicalisation constructions and intonation contours (Engdahl & Vallduví 1996), *inter alia*. What these models have in common is that CONTEXT is used as a 'repository' into which information can be put by lexical items or by constructions. The value of this feature then can restrict the situations in which an utterance of this sign, i.e. the production of a token of this type, would be considered appropriate.

Let's see whether we can reconstruct the GBA as doing something similar. Can the sign in Figure 5.3,

[40] Although the paper does not say anything about where these rules are applied, in the grammar or externally.
[41] Cf. e.g. (Pollard & Sag 1994, p.9).

representing ‘John’ as an answer to ‘Who saw Mary?’, be seen as carrying the contextual restriction ‘I can only be used to convey the message ‘John saw Mary’ appropriately in a context where the question ‘Who saw Mary?’ has just been asked, and I’m meant to provide an answer.’? This is a rather specific contextual restriction, and there’s a catch: the restriction regarding the content of the question cannot conceivably come from the lexical entry ‘John’, neither can it come from the construction. The information flow cannot be the same as in the analyses of the phenomena mentioned above, from constituents of the sign to restrictions on its use. Rather, here it seems to be the context that restricts the content of the sign. We are looking at a novel use of this feature, then, and we have to explain where its value comes from.

**CONTEXT-as-Discourse-Representation Interpretation** Another option is to take the name CONTEXT literally and assume that this attribute indeed contains a representation of the linguistic context, from which the *grammar* then computes the values for MAX-QUD and SAL-UTT.[42,43] This, however, seems rather implausible. One of the main relations which in KOS structures the context, resolves (cf. Section 5.2 above), is defined agent-relative, while the other notions need access to the semantics of question and answer (to decide whether the answer decides the question). The process of question accommodation, as developed in (Cooper et al. 2000), likewise relies on detailed information, in this case about plans. The upshot of this is that the computation of QUD must rely on information about content and about the cognitive state of agents; hence this information has to be represented in a logic that is at least first order (Kamp 1991). Moreover, reasoning about cognitive states should be non-monotonic, since it is inherently fallible (see references above). Using a first order logic (and a non-monotonic one at that) as the underlying logic, however, makes it undecidable which *discourses* are well-formed according to the grammar. Since under that interpretation there would be no distinction between the sentence-grammar and the discourse-“grammar”, the former would be affected by this problem.[44] We conclude that this also cannot be the intended interpretation of the role of CONTEXT.

**External-Module Interpretation** (Ginzburg & Gregory 2001) offers an implementation–oriented suggestion for how the values for CONTEXT are determined. They divide the task of fragment interpretation into four components (Ginzburg & Gregory 2001, p.163):

1. An HPSG grammar;
2. a dialogue representation module, which takes on the tasks of structuring and storing a representation of the previous discourse, and of computing the values for MAX-QUD and SAL-UTT;

[42] Note that this would have to be a representation of *types* of contexts, since a HPSG is supposed to model *types* of linguistic expressions, not tokens.

[43] Under this interpretation KOS would ‘live’ inside the grammar.

[44] There are less costly ways to deal with pragmatic information. SDRT, as we will see in detail later, uses a sophisticated system of logics that encapsulate information and provide information to a so-called ‘glue-logic’, a decidable logic in which the discourse structure is computed. This approach, however, relies crucially on this modularity; the model of grammar/pragmatics we describe here is fundamentally unmodular.

3. these values are then assigned to the features in the sign of the current clause,
4. whose content is then resolved.

If this is to be read as a description of the steps of the process of fragment interpretation, it seems that they are using an unusual notion of grammar here. It is not fully clear to us what the output of the grammar in Step 1 is in the case of fragments, but from the description of Step 3 it seems that the grammar outputs a sign consisting of a full parse for the head daughter (the fragment-phrase) and co-indexations according to the fragment rule (Figure 5.2), but lacking actual values for MAX-QUD and SAL-UTT. The task of computing these values (for example using the coercion-rules described above) is handed to another module, which for all practical purposes could be called a pragmatic discourse management module. These values are then assigned to the features, and the content of the fragment is then resolved, presumably by unifying these values into all places that are coindexed.

This description of the implementation hints at what probably is the intended interpretation of CONTEXT in the GBA. Much like the C-INDICES (representing speaker, hearer, and spatio-temporal location) are meant to be instantiated 'at run-time', MAX-QUD and SAL-UTT could be seen as being interface-points between pragmatics and grammar, whose values get instantiated at the point of *using* the sign. (This would have to allow for ambiguity, though, as shown with example (207) above, where there are several possible values for SAL-UTT.)

We have several objections against such a view. First, from a very technical stance, it makes it difficult, if not impossible, to use parsing systems that are not purpose-built, since it requires either additional input (the values for MAX-QUD and SAL-UTT), or a call to a pragmatics module 'mid-parse', as it were. More substantially, by reversing the information flow in that way—from context, which is out of reach to the constraints of the grammar, to content—the grammar isn't strictly generative anymore. Where previously the lexical items could be seen as the base of the recursion that could enumerate all well-formed representations of the grammar, this is now not possible anymore. Lastly, the information always flows from the QUD to the resolution; i.e., the (underspecified) content of the fragment cannot be used to determine what the QUD is.

#### 5.5.4.3 Other Changes

In this section we will very briefly discuss the repercussions of the further changes the GBA introduces which were illustrated with Figure 5.6.

The criticism we want to put forward here is related to an issue we raised before, namely that in our opinion it is not advisable to let the grammar decide on speech acts that have been performed. As explained above in Section 5.4.2, (Ginzburg & Sag 2001) assume that the illocutionary force of an utterance is represented in its sign; the example above was that of an interrogative that has as content an *ask-rel*. As Ginzburg & Sag (2001) themselves note, this is rather close to the 'performance hypothesis' which was pursued in the 1970, according to which all sentences contained a (normally unexpressed) operator for illocutionary force (see references and discussion in (Levinson 1983)). The authors claim that they can escape certain negative consequences of this hypothesis since they do not have to assume that this operator is syntactically represented. However, it is not quite clear how they intend to escape the semantic problem that the illocutionary force relation that is 'wrapped around' the original content has an influence on the truth-value of the utterance: "I assert that *p*" is true just by uttering it, whereas the truth of *p* might be a contingent matter. But even if they can avoid this problem, our main objection is another one. As mentioned above, the alignment between linguistic form and performance of a speech act is generally seen as the product of highly complex default reasoning that is beyond the reach of the grammar. Simple information about sentence mode is certainly not enough. (Lascarides 2001) cites the following two sentences as an example that not all (syntactic) imperatives are commanded:

(208) a. Go straight on and you'll fall down the stairs.
b. Come home by 5pm and we can go to the hardware store before it closes.

The information that makes the difference here, however, is not normally available to the grammar (as discussed above), and so the assignment of an *imperative-rel* by the grammar either can't be taken to entail that its argument is commanded, or if it is, it will sometimes make the wrong predictions about what speech acts are performed.

This concludes our discussion of the GBA; we will summarise the main points in the following summary section.

## 5.6 Summary

In this long chapter we have introduced and discussed in detail the interesting grammar–based approach to the resolution of fragments put forward by Jonathan Ginzburg and colleagues. We have seen that the general techniques used (functional application and a syntactic matching condition) can relatively successfully model the resolution of *res-via-id* fragments. However, we have also seen that the framework in which the authors use these techniques has certain shortcomings, which we summarise in the following:

- Functional application is too simplistic to resolve *res-via-inf* fragments.
- The non-compositionality of the approach leads to a duplication of work: the grammar must produce content corresponding to all speech acts that could be performed with a certain fragment, out of which then the actual contextually appropriate one has to be filtered out by a discourse interpretation module.

- The non-compositionality also leads to a rather complex and not very well described interaction between grammar and discourse modelling.

As we hope to show in the next chapters, these changes are unnecessary and not justified by the data; using the device of *underspecification* and allowing the discourse modelling limited access to syntactic information we can achieve better empirical coverage while at the same time offering a clean, modular architecture that is also independently motivated. Our approach can be seen as covering the middle ground between Ginzburg's GBA and Carberry's (1990) plan-based approach, in that it makes use of 'cheap' linguistic information like the GBA does whenever it can, but can recourse to plan-based reasoning of the type Carberry describes when it must.

# Chapter 6

# A Coherence-Based Approach I: Semantics

With this extensive review of the extant approaches done, it is now time to introduce our own approach, which attempts to overcome some of the problems we highlighted in earlier approaches. Rather than extending a grammar to include representations of the context of utterances, like the GBA does, we aim to use a more conservative grammar and account for fragments by interfacing this grammar with an accompanying theory of discourse interpretation that is designed to analyse anaphoric phenomena of various kinds (SDRT, (Asher 1993, Asher & Lascarides 2003)). The nature of the interface between grammar and pragmatics that we assume is by now almost standard in computational semantics: we construct (possibly partial) *descriptions* of logical forms on the syntax/semantics interface and use discourse information to extend these partial descriptions to more complete ones (Alshawi 1992, Blackburn & Bos 2000). This allows us to restrict the task of the grammar to constructing representations that express exactly that information that can be acquired from the syntax of an utterance, possibly leaving certain semantic distinctions, which are ultimately made on the basis of contextual information, *underspecified*. Adding further specifications to these representations is left to a discourse update module that aims to integrate new utterances into the context in such a way that makes the contribution *coherent* with the context. (Hence, we will call our approach "coherence-based", CBA, to contrast it with the grammar-based approach (GBA) described in the previous chapter.)

We begin our presentation of the CBA with the definition of the syntax and semantics of the representations of the compositional semantics of fragments we will use.

## 6.1 Overview

We have seen in the previous chapter that the GBA does not have a semantic notion of fragments as such: in the case of NP-short-answers for instance, the semantics of the NP is directly combined with contextual material (in this case the question they are meant to answer) to yield a propositional content. At no point is there a representation of the semantics of the fragment independent from its context. In our set-up, we want to separate grammatical and contextual analysis, while allowing the two sources of information to interact during discourse update. Hence, we will have to provide a *compositional* semantics for fragments. This semantics will have to represent both the semantic information that comes from the phrase as well as the fact that there is information "missing" that has to be filled in by the context. The compositional semantics should also ensure that this 'missing' information is such that, were it resolved, the result would be a *message*—i.e., either a proposition, a question or a request.

This chapter is structured as follows. In Section 6.2 we motivate the use of semantic underspecification. In Section 6.3 we describe a particular underspecification formalism, $\mathcal{LL}$ (after (Asher & Lascarides 2003)). This formalism is introduced with scope-ambiguities as examples; in 6.4 we present our extensions that allow for the expression of the underspecification arising from the use of fragments. In a nutshell, our extension is a constraint that ensures that all specifications of the representation of the fragment use the semantic material of the fragment phrase in a certain way. E.g., if the fragment is "Peter.", then our representation demands that all resolutions denote an event which involves Peter.

Readers who are familiar with underspecification formalisms can skip directly to Section 6.4, possibly after scanning the definitions given in 6.3. To only get an idea of our approach to underspecifying the compositional semantics of fragments reading Section 6.4.1 is sufficient; further examples are discussed extensively in Section 6.4.2. In Section 6.5 some final revisions to the formalism are introduced.

## 6.2 Why underspecification?

Suppose we choose a language of predicate logic to represent the meaning of sentences of English. As an illustration, sentence (209-a) would then get assigned the logical form (LF) (209-b) (ignoring tense information, and following a neo-Davidsonian approach to the representation of the meaning of event-sentences).

(209) a. Peter saw Mary
b. $\exists e \exists x \exists y (\mathit{named}(x, \text{“Peter”}) \wedge \mathit{named}(y, \text{“Mary”}) \wedge \mathit{see}(e, x, y))$

Now, what can we say in this logic about the semantics of an NP-fragment like "Peter."? One thing we know about all meanings that can possibly be intended to be conveyed with such a fragment is that they will be of propositional type.[1] That means that once we decide that a certain utterance is a fragment (and not just a phrase), we have more semantic information than what is represented in just the translation of the phrase it consists of.

We have to answer a possible objection here. Why should we not let the pragmatic module work directly on the 'normal' meanings of the fragment-phrases? There are several reasons why this is not the optimal strategy, which will be discussed in detail later on, so we only highlight some of the arguments here. Firstly, working on meanings (which will be represented at least as first-order formulae) directly rather than on descriptions means that a logic with an intractable notion of validity has to be used for pragmatic processing; this should be avoided. Secondly, it would mean that we cannot use the same definitions for rhetorical relations as for full sentences, since they expect propositions, questions and requests as arguments, and so we would have to introduce relations specifically for fragments. Moreover, since the same speech act can be performed with many different syntactic fragments (e.g., short-answers can be PP-fragments, NP-fragments, VP-fragments etc.) whose meanings all have different semantic types, we would need several relations just for fragments realising one speech act. Lastly, there is also data that suggests that fragments have a different syntactic status than phrases—this at least suggests giving them a different semantics as well.

As we said, we will discuss these points again later on, and so we just conclude for now that for our example "Peter.", the LF (210)—the translation of the NP—isn't sufficient as a representation of the semantics of the *fragment*.[2]

(210) $\lambda P.\exists x(\mathit{named}(x, \text{“Peter”}) \wedge P(x))$

What we have to express in our semantic representation, then, is that the translation of the phrase—in our case, a quantified variable that is in a "named"-relation—is an *argument* of some (perhaps complex) event-predicate; which predicate, the grammar can't say. (211) shows two attempts at expressing this in the same language of predicate logic we have used in (209-b).

(211) a. $\exists P \exists e \exists x(P(e,x) \wedge \mathit{named}(x, \text{“Peter”}))$
b. $\exists e \exists x(P(e,x) \wedge \mathit{named}(x, \text{“Peter”}))$

The formulae above simply introduce a second order variable $P$ into the LF (informally speaking, in place of the verb-phrase translation in (209-b)). This variable is bound by a quantifier in the first formula, and free in the second. Now, does this express what we want? Let's first look at (211-a).

[1]That is, if we make the simplifying assumption here that there is an un-ambiguous 'declarative' intonation. We will later see how this decision about the type of message can be left underspecified in the compositional semantics as well.

[2]Note that, as is customary in Montagovian approaches to formal semantics, we assume the use of higher-order logic ($\lambda$-calculus) during construction of LFs here, even when the final representation for the sentence might still be first order. However, for (211) we will need higher-order variables as well, and so we can assume that the language for meaning representation used here is higher-order as well.

The formula (211-a) is true if there exists a predicate *P*, an eventuality *e* and an individual *x* named "Peter", and *P* is predicated of both. One now could imagine a strategy, following the abduction-based approach to discourse representation of (Hobbs et al. 1993), in which interpreting the fragment "Peter." consisted in finding a proof of the formula (211-a), possibly drawing on additional information. Finding a proof involves instantiating the variable *P* with a 'concrete' instance—a predicate constant. In the context of a question like "who saw Mary?", *P* could accordingly be resolved to something like '$\lambda z \exists y(named(y, \text{"Mary"}) \wedge see(e,z,y))$',[3] so that the fragment would be resolved to (209-b).

We have some objections to such an approach, however.[4] First, this would entail that all the reasoning necessary to resolve underspecification would have to be done within the higher-order logic used for meaning representation; a logic with an intractable notion of validity. We will later see that we can use a much less expressive logic for this contextual reasoning if we use *descriptions* of (logical forms representing) the content. Second, such an approach also fails to distinguish situations where a coherent interpretation requires one to identify what instantiates the existentially quantified variables and where it doesn't. Put differently, it does not distinguish between values that are unspecified for linguistic reasons and those that are for other reasons. Last but not least, such an approach would force a certain order of quantifiers (in cases where *P* includes a quantifier). We have seen in the discussion of the GBA that this is problematic, given the observation that fragments can still exhibit scope ambiguities.

For the same reasons, we will not use representations like (211-b), where the predicate-variable is left free. One could imagine an approach using such representations where resolving fragments consists in updating an assignment function that provides a value for the free variable. But again, such an approach would not make it clear in the representation where there is information that is underspecified by the syntax.

We will instead use a device that is independently motivated, well studied, and can easily be integrated into modern grammar formalisms (Copestake, Flickinger, Sag & Pollard 1999, Dalrymple 1999): underspecification. Rather than trying to express directly in the logical form that information is "missing", we climb to a 'meta-level' and let the grammar produce a *description* of logical forms. This description only has to be as precise as the linguistic information affords; if there are semantic differentiations the grammar can't make, the description simply describes a *set* of logical forms containing representations for all different ways of making the distinction. In other words, the description produced by the grammar is a *partial* one. So what we want to express for our fragment example "Peter." is that it will denote an eventuality, but one about which we only have very partial information (namely that the entity denoted by "Peter" is somehow involved in it). In (212), this is glossed as '$P = ?$' with the *x* being an argument to that unknown relation.

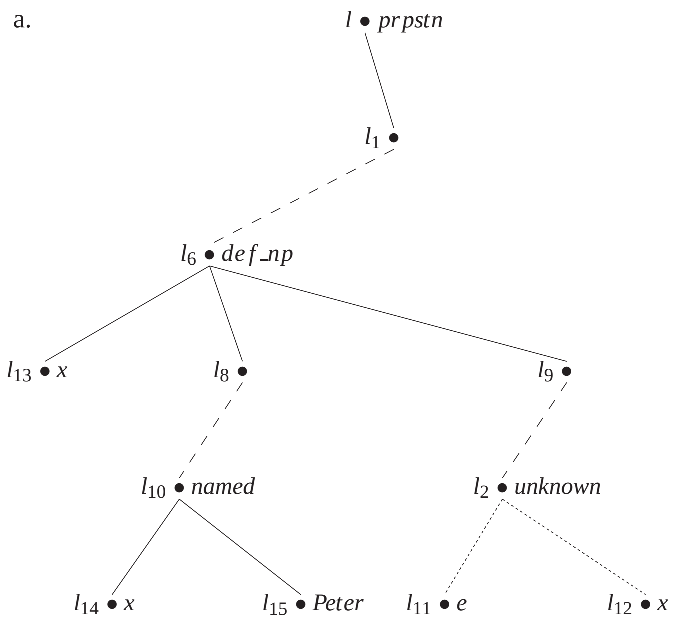
(212) $\exists e \exists x P(e,x) \wedge named(x, \text{"Peter"}) \wedge P = ?$

[3] Glossing over how the event variable should be dealt with.
[4] For a detailed argument contra such a strategy see also (Asher & Lascarides 2003).

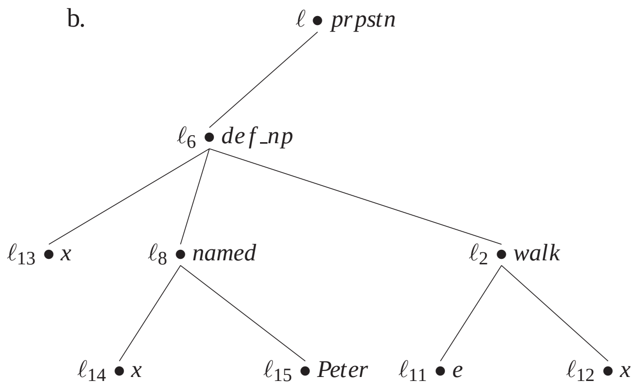
Note that (212) is not intended to be a formula of the language of LFs; rather, it is a description of the *form* of such formulae. More precisely, the description will be satisfied by any 'real' logical form in the language we choose for meaning representation that contains "$named(x,$"Peter"$)$" and a (perhaps complex) predicate relating $e$ and $x$. In other words, these normal logical forms will be *models* of the description.[5]

To give a visual impression of this approach, (213-a) shows a graphical representation of the compositional semantics of "Peter." that we will eventually adopt, while (213-b) shows a graphical representation of one of the logical forms that are described by (213-a) (corresponding to "Peter walks."). Roughly, the role of $P =?$ in (212) is played here by the constraint *unknown*. The various styles of lines express certain *scopal* or structural relations between the various parts of the formula—this will be explained in detail presently.

(213) a.




b.




[5] Recall the relation between HPSG-constraints and the feature structures they describe.

In Section 6.4, we will make formally precise what these representations express, but first we have to introduce the formalism in which they are framed. The strategy for expressing underspecification that is followed in the formalism, namely labelling bits of formulae (see the labels $l_1$ to $l_{12}$ in (213-a) for instance), was originally invented to deal with scopal ambiguities (see e.g. (Reyle 1993, Bos 1996)), but has since been extended to other sources of semantic underspecification (plural NPs, lexical ambiguity, anaphora, VP-ellipsis; see for example the papers collected in (Deemter & Peters 1996), also (Egg, Koller & Niehren 2001)). The particular formalisation of the labelling-idea we describe here is from (Asher & Lascarides 2003), which in turn is an extension of (Asher & Fernando 1999), combined with ideas from CLLS (Egg et al. 2001); we have made some minor changes to it that will be indicated below. The background section in which this approach is described is a bit longer than previous ones, because later we will need to make fairly detailed extensions which rely on the definitions of the basic approach.

## 6.3 Background: A Formalism for Underspecified Semantics

Sentence (214) exhibits a well-known quantifier scope ambiguity, which results in its having the two readings represented by (215-a) and (215-b):[6]

(214) Every fish needs a bicycle.

(215) a. $\exists(y, bicycle(y), \forall(x, fish(x), need(x,y)))$
b. $\forall(x, fish(x), \exists(y, bicycle(y), need(x,y)))$

We can describe in words what both formulae have in common: there is an existentially quantified variable $y$ with the restriction $bicycle(y)$, a universally quantified variable $x$ with restriction $fish(x)$, and the formula $need(x,y)$ is in the scope of both of the quantifiers. In effect, the only difference between the two readings is the order of the quantifiers (and hence the values of their nuclear scopes). What we want to achieve now is to give descriptions like this in a logical language which allows the exact list of described (first-order) formulae to be computed. This can be done by labelling such bits of formula and defining a language for talking about structural relations between these labels. (216) shows the common structural elements of the two formulae above, tagged with a *label*, and with an underscore in place of the nuclear scope, to indicate that what exactly fills this place is underdetermined by the grammar.

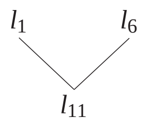
(216) $l_1 : \forall(x, fish(x), \_), \; l_6 : \exists(y, bicycle(y), \_), \; l_{11} : need(x,y)$

This expresses the first part of the description we have given above in words. What is not yet specified is the information that the material labelled with $l_{11}$ is in the nuclear scope of both quantifiers. This can be

[6]Where, for ease of exposition, the quantifiers are represented as generalised quantifiers with separate argument places for restriction and scope.

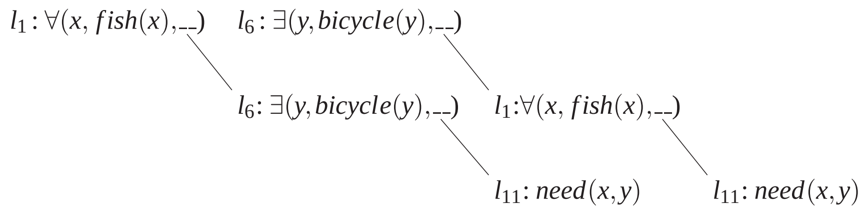
illustrated graphically as in (217), where we interpret subordination in the graph as the relation "being-in-the-scope-of". Note that (217) is one of many equivalent ways of expressing what we want—we will come back to this issue.

(217)




The two formulae (215-a) and (215-b) correspond (in a sense we will make precise in the next section) to the two possible ways of turning this graph into a tree, where the formula tagged with the subordinate label is "plugged" into the scope of the formula immediately dominating it, as shown in (218). In that sense, (216) together with (217) *describes* these formulae.

(218)




In the following, we describe a specific formalisation and generalisation of this labelling approach, namely that by (Asher & Lascarides 2003) (henceforth *A&L*). The general strategy is as follows. Logical formulae in one language are viewed as corresponding to certain structures; and then another language is defined that is interpreted over these structures. The semantics of this latter language is restricted in such a way that satisfiable formulae of this language only have models that correspond to formulae of the first language; and in that sense, these formulae of the underspecification language *describe* these other formulae.

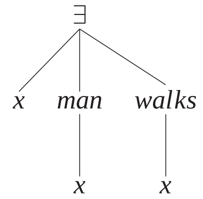
#### A labelled language for scope underspecification

First, some terminology. Underspecification languages, as mentioned before, are *description* languages. They allow us to describe formulae of one language (the base language, henceforth we will write "BLF" for "base language formula") with formulae of another language (the description or constraint language; we will sometimes call formulae in this language *underspecified logical forms* (ULFs)). It is important to stress that these languages can be given different semantics; e.g. the base language can be higher order and dynamic, while the description logic is first order and static. For concreteness, we will use a first-order predicate language as base language in our examples here (we won't need to say anything at this stage about how it is interpreted, dynamically or statically). We will make this separation between the base logic and the description logic explicit by writing the satisfaction relation for the description language as $\models_d$ and that for the base language as $\models_b$.

Logical formulae can be viewed as *trees* built by *constructors*, which are the symbols of the language (i.e. the constants—both logical and non-logical—and the variables). So the simple BLF (219-a) can be represented equivalently by the tree (219-b).[7]

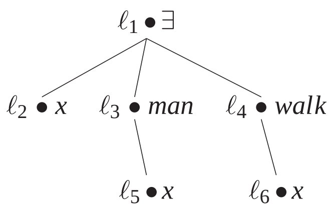
(219) a. $\exists(x, man(x), walks(x))$

b.




We want to be able to talk about *structural* relations between subformulae of BLFs; in the tree perspective, this corresponds to talking about nodes and subtrees. *A&L* take up this idea by defining *labelled-Σ-structures* ($L\Sigma S$s). Intuitively, these structures are built out of constructor trees by viewing the nodes as being *labels* for parts of the formula. The following shows a graphical representation of a $L\Sigma S$ corresponding to (219-a) / (219-b). We use bullet-points and write the $\ell_n$s to their left just for readability; note that the $\ell_n$s are supposed to *be* the nodes of the graph.[8]

(220)




$L\Sigma S$s are defined as follows (note that they, being intended as description-structures, are defined relative to the signature Σ of a given base language):[9]

**Definition 1** *Labelled Σ-Structure*
Given a set Σ of constructors $f$ (in the base language) with specified arities $n_f \in \{0, 1, \ldots\}$, a *labelled Σ-structure* is a triple $\langle U, Succ, I\rangle$ where

(L1) $U$ is a non-empty set (of labels).

(L2) $I$ is an interpretation function which is defined on Σ, such that for every $f \in \Sigma$ of arity $n$, $I(f)$ is an $(n+1)$-ary relation on $U$ (i.e., $I(f) \subseteq U^{n+1}$). (We will call these relations *labelling relations*.)

[7] Representing quantifiers in a generalised way, as above, and forgetting about eventualities for the moment, (219-a) is the LF of the sentence “a man walks”.

[8] *A&L* do not use this representation format; we have adapted it from (Egg et al. 2001) for purposes of illustration.

[9] This definition is taken from (Asher & Lascarides 2003, p.129).

(L3) The binary relation *Succ* on $U$ is well-founded,[10] where (by definition) *Succ* consists of exactly the pairs $\langle \ell', \ell \rangle \in U \times U$ such that for some $f \in \Sigma$ of positive arity $n > 0$, and some $(n-1)$-tuple $\ell_1, \ldots, \ell_{n-1}$ from U:
$\langle \ell, \ell_1, \ldots, \ell_{n-1}, \ell' \rangle \in I(f)$ or $\langle \ell_1, \ell, \ell_2, \ldots, \ell_{n-1}, \ell' \rangle \in I(f)$ or ...
or $\langle \ell_1, \ldots, \ell_{n-1}, \ell, \ell' \rangle \in I(f)$

(L4) $U$ contains a unique supremum $\ell_0$ in the partial order defined by *Succ*; i.e., $\forall \ell \in U, Succ*(\ell_0, \ell)$, where *Succ*∗ is the transitive closure of *Succ* (we will write $\ell \succeq \ell'$ for $\langle \ell, \ell' \rangle \in Succ*$).

(L5) For all $\ell \in U$ except for $\ell_0$, there is a unique label $\ell'$ such that $Succ(\ell', \ell)$; i.e., each label except for $\ell_0$ has a unique parent. This means in effect that every label in $U$ except for $\ell_0$ is an argument to a unique constructor symbol. (L3-L5) ensure that the labels $U$ form a *tree* under *Succ*.

Observe that the relation *Succ* can be reconstructed from $U$ and $I$; for certain definitions however it is more convenient to represent *LΣS*s as $\langle U, Succ, I \rangle$ rather than just $\langle U, I \rangle$.

We can see that (220) is a graphical representation of an *LΣS* if we view the set $\{\ell_1, \ldots, \ell_6\}$ of nodes in (220) as our $U$, and read the (parts of the) interpretation function $I$ off the tree by adding a tuple $\langle \ell_1, \ldots, \ell_n, \ell_k \rangle$ to $I(f)$ iff $\ell_k$ labels $f$ in the tree and $\ell_1, \ldots, \ell_n$ are the daughters of $\ell_k$. (This illustrates why we call the $I(f)$s *labelling*-relations.) (221) shows (220) in a non-graphical notation as a triple as in the definition given above. Note how every occurrence of '*x*' for example has its own label.

(221)
$$\langle \quad \{\ell_1, \ldots, \ell_6\},$$
$$\{\langle \ell_1, \ell_2 \rangle, \langle \ell_1, \ell_3 \rangle, \langle \ell_3, \ell_5 \rangle, \langle \ell_1, \ell_4 \rangle, \langle \ell_4, \ell_6 \rangle\},$$
$$\{I(\exists) = \{\langle \ell_2, \ell_3, \ell_4, \ell_1 \rangle\}, I(x) = \{\ell_2, \ell_5, \ell_6\}, \ldots\} \quad \rangle$$

In fact, every *LΣS* corresponds to a unique formula in the base language, and every formula in the base-language corresponds to a unique *LΣS*, up to alphabetic variance on variables and labels, respectively.

*A&L* now define a language that gets interpreted over these structures; we will call this the "labelled language" ($LL$). This is the language in which the grammar will specify the descriptions of the logical forms of the sentences. Its construction is basically a straightforward implementation of the strategy outlined at the beginning of this section: the syntax of $LL$ rests on that of the base language, so that for every constructor $f$ of arity $n$ of the base language we have a relation symbol $R_f$ of arity $n+1$ in $LL$, where the additional argument place is filled with what can be viewed as the label of the constructor. For example, the three-place constructor '∃' of the base language becomes $R_\exists(l_1, l_2, l_3, l)$ in $LL$, where $l_1, l_2$ and $l_3$ are the labels of the arguments of this constructor (the bound variable, the restrictor, and the nuclear scope, respectively), and $l$ is the label of the constructor itself; a variable $x$ of the base language becomes $R_x(l)$. So all constructors of the base language become *predicates* in the labelled language, and in that sense lose the meaning they have in the base language. E.g., from $R_{every}(l_1, l_2, l_3, l_4)$ does

[10] I.e., every non-empty linearly ordered subset of this partial order has a minimal element; i.e., there are no cycles.

not follow in the $LL$ that $R_{not}(l'_4, l_4) \wedge R_{all}(l_1, l_2, l_3, l'_4)$. In that sense, the logic of the labelled language 'knows' about the *form* of the WFFs of the base language, but not about their *interpretation*.

We also sort our labels $l$, depending on what kind of base language expression they label. This is defined in the following definition. The definition of the $L\Sigma S$s has to be adapted accordingly to sort the universe $U$, which we omit here.

**Definition 2** *Sorts*
Base language variables of type *individual* are labelled by labels of the sort $i^l$, those of type *event* by those of sort $e^l$;[11] base language formulae are labelled by labels of sort $t^l$.

The resulting language is first-order without function symbols and quantifiers, but with (sorted) variables.[12]

**Definition 3** *The Syntax of the Labelled Language ($LL$)*

- **Vocabulary:**
  - the logical constants $\wedge, \vee, =$ and $\neg$;
  - a set $L$ of sorted variables $l$ (with sub- and superscripts)—we call these variables *label-variables*; and
  - a set of predicate symbols with arity $m \geq 0$. Specifically, there is
    - for each $n$-place constructor $f \in \Sigma$ a corresponding $n+1$-place predicate symbol $R_f$ in the labelled language, taking labels of the appropriate sorts as arguments; and
    - the two-place predicate '*outscopes*.'
- **WFFs:**
  - Let $P_m$ be an $m$-place predicate symbol in the vocabulary, and let $l_1, \ldots, l_m$ be labels. Then $P_m(l_1, \ldots, l_m)$ is a WFF.
    (For the predicates that are derived from constructors of the base language, we will also allow an infix-notation $l_m : f(l_1, \ldots, l_{m-1})$, where $f$ is the constructor.)
  - If $l_1$ and $l_2$ are labels, then $l_1 = l_2$ is a WFF.
  - If $\phi, \psi$ are WFFs, then so are $\phi \wedge \psi, \phi \vee \psi, \neg\phi$.

The model-relation $\models_d$ between $L\Sigma S$s and variable assignments on the one side and formulae of $LL$ on the other is defined in the usual Tarski-style, with the exception of the addition of a clause for *outscopes*;

[11] Note that we assume an ontology of the base language that differentiates between individual variables and events, and also that we deviate from orthodoxy here and use $i$ for the type of individuals, so that we can use $e$ for events.

[12] Here we deviate from *A&L*, who distinguish in the syntax between labels and holes, the former being constants of $LL$ and the latter variables that are bound by an existential quantifier. We prefer to uniformly have all $l$s as variables, and so let the assignment function do all the work.

we will omit the general definition and only show that additional clause in (222) below. We will call pairs of $L\Sigma S$s $\Lambda$ and variable assignments $g$ that are such that $\Lambda \models^g_d \phi$, where $\phi$ is a formula of $LL$, *solutions* of $\phi$ (again this is our and not Asher & Lascarides's (2003) nomenclature; inspired by (Egg et al. 2001)).

(222) $\langle U, Succ, I\rangle \models^g_d outscopes(l_1, l_2)$ (where $g$ is an assignment function) iff $[\![l_1]\!] \succeq [\![l_2]\!]$.

We are now in a position to show how the ideas from the introduction to this section can be realized with this language and the structures over which it is interpreted. Above we said that the labelled bits of formula (216) (repeated here as (223-b)), together with the scope-relations represented graphically in (217) (shown non-graphically with a predicate *outscopes* in (223-c)), describe the two readings of sentence (214) (which is repeated here as (223-a)).

(223) a. Every fish needs a bicycle.

b. $l_1 : \forall(x, fish(x), \_), l_6 : \exists(y, bicycle(y), \_), l_{11} : need(x, y)$

c. $outscopes(l_1, l_{11}), outscopes(l_6, l_{11})$

To turn the pseudo-formula in (223-b) into a formula of $LL$, we have to do two things. First, we have to label all constructors of the base-language, and second, we have to specify how the 'holes', signified by the underscores, are to be represented. For the latter task, we will simply re-use label-variables, where the only difference between these 'holes' and other label-variables is that they do not label anything, i.e. do not occur as the last argument of any relation. And so (223-b) can be represented in $LL$ as shown in (224-a) below. Note that since we now have labels for the underscores in (223), we use those in the outscopes statements. This makes the graphical presentation we will later show simpler; these two ways of expressing the structural relations are equivalent. We will discuss in a minute how, given further assumptions, the outscopes-relations can be expressed even more economically.

Since these formulae become unreadable quite quickly, we will sometimes collapse formulae where labels do not participate in ambiguities, and so allow (224-a) to be written as (224-b) (the infix notation for labels was defined in Definition 3).

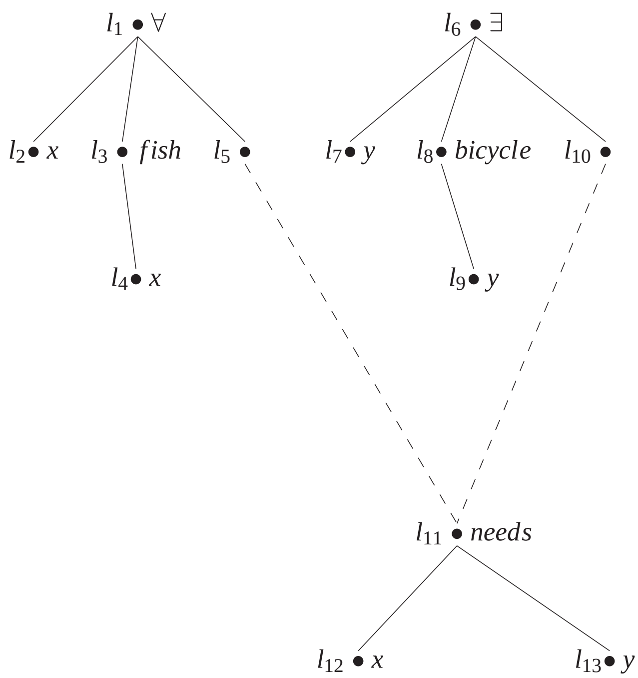
(224) a.

$$
\begin{array}{llll}
R_\forall(l_2, l_3, l_5, l_1) & \wedge & R_\exists(l_7, l_8, l_{10}, l_6) & \wedge \\
R_x(l_2) & \wedge & R_y(l_7) & \wedge \\
R_{fish}(l_4, l_3) & \wedge & R_{bike}(l_9, l_8) & \wedge \\
R_x(l_4) & \wedge & R_y(l_9) & \wedge \\
\multicolumn{4}{l}{\quad R_{needs}(l_{12}, l_{13}, l_{11})} \\
R_x(l_{12}) & \wedge & R_y(l_{13}) & \wedge \\
outscopes(l_5, l_{11}) & \wedge & outscopes(l_{10}, l_{11}) & \wedge
\end{array}
$$

b. $l_1 : \forall(x, fish(x), l_5) \wedge l_6 : \exists(y, bicycle(y), l_{10}) \wedge l_{11} : need(x,y) \wedge outscopes(l_5, l_{11}) \wedge outscopes(l_{10}, l_{11})$

As even the collapsed constraints can become rather complex, and to bring out the closeness between the formulae and their intended models, we introduce a third—graphical—notation.[13] In this notation, the formula looks as shown in (225). Label variables are written to the left of their node of this graph; if they label a base-language constructor, this is written to the right of the node (i.e., 'holes' do not have anything on their right side). Solid lines in the graph represent the argument-relation, while the dotted lines stand for the *outscopes*-relation between nodes.

(225)

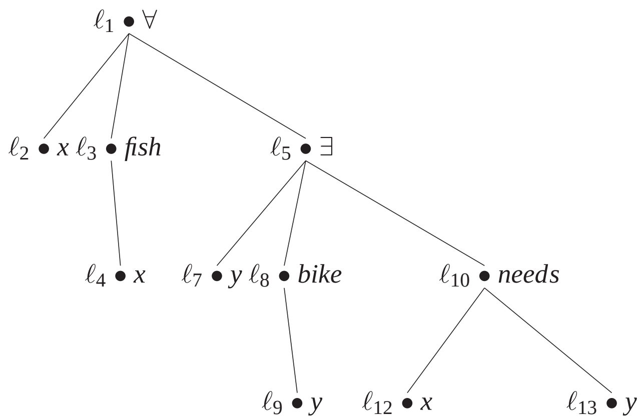



The similarity between this representation and the graphical representation of *L*Σ*S*s introduced above in (220) is quite intentional. The one main difference is that whereas *L*Σ*S*s are always trees, $LL$ constraint graphs don't have to be. The connection between the two graphs (or, equivalently, between formulae of $LL$ and *L*Σ*S*s) is made, as mentioned above, via $\models_d$. (226) is a variable assignment *g* that shows that the *L*Σ*S* shown in (227) is a *model* for (224-a)—i.e., the pair (227) and (226) forms a *solution* of (224-a). And in that sense does (224-a) describe the base-language formula (215-b): the *L*Σ*S* corresponding to (215-b) is a solution for the constraint-formula (224-a).[14]

[13] This again is slightly different from the presentation in *A&L*.
[14] We omit showing how this works for (215-a), where basically just the quantifiers swap position.

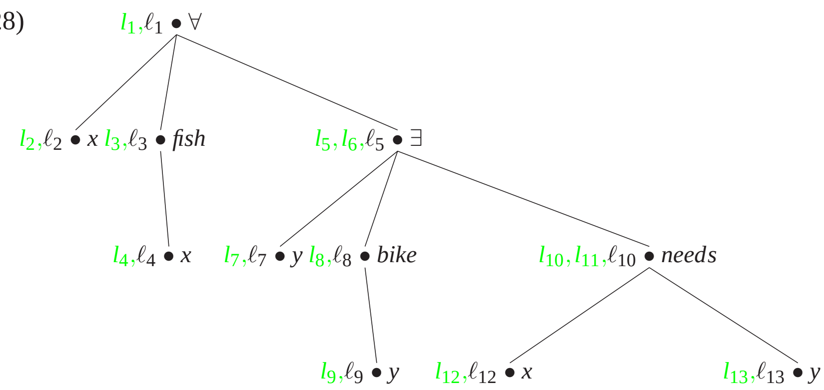
(226) $g(l_1) = \ell_1, \quad g(l_2) = \ell_2, \quad g(l_3) = \ell_3, \quad g(l_4) = \ell_4,$
$g(l_5) = g(l_6) = \ell_5, \quad g(l_7) = \ell_7, \quad g(l_8) = \ell_8,$
$g(l_9) = \ell_9, \quad g(l_{12}) = \ell_{12}, \quad g(l_{10}) = g(l_{11}) = \ell_{10},$
$g(l_{13}) = \ell_{13}$

(227)




To represent solutions we now introduce a more compact graphical representation, illustrated by (228). In this representation, the $L\Sigma S$ is shown together with the intended variable assignment superimposed in coloured print.

(228)




We're not done yet, though. We would like it to be the case that (226) and a similar pair of $L\Sigma S$ and assignment corresponding to the other reading given above in (215) are the only two solutions for the $LL$ formula. But at the moment they aren't. In fact, there is an infinite number of solutions, and the reason for that is that there is nothing in the definitions above preventing nodes that aren't referred to in the constraint from occurring in the $L\Sigma S$, as long as all constraints (for example on *outscopes*-relations) are satisfied. So for example an $L\Sigma S$ that is like (227), except that the subformula at $\ell_{10}$ is not a direct daughter of $\ell_5$ but rather part of a complex formula that is—e.g. an $L\Sigma S$ paraphrasable as 'for every fish there is a bicycle such that John thinks that fish needs that bicycle')—would still be a model of (224-a).

This property of not restricting the size of the model (and hence the 'size' of the described formula) is something we require for the representation of fragments, as we will see in the next section, but here for scope-underspecification we must make one more assumption to reach our goal of showing that such formulae are indeed descriptions of all and only the readings of a scopally ambiguous sentence: we require models in solutions to be *minimal*, i.e. to use a minimal number of entities in $U$. If we make this restriction, then (224-a) does indeed only describe the two readings we want.

It's time to take stock now. We have described how certain structures, the $L\Sigma S$s, can be seen as representing logical formulae of one language, the base language. We then have presented the language $LL$ that is interpreted over such structures, and we have shown how a formula in this constraint language can be seen as describing formulae of the base language, namely via the relation $\models_d$ between formulae of $LL$ and their models. Following *A&L*, we can now overload $\models_d$ in the usual way to denote logical consequence, as shown in the following definition:

**Definition 4** *logical consequence*
For WFFe $\phi, \psi$ of $LL$, $\phi \models_d \psi$ iff for all $L\Sigma S$s $\langle U, Succ, I\rangle$ and assignments $g$, if $\langle U, Succ, I\rangle \models_d^g \phi$, then $\langle U, Succ, I\rangle \models_d^g \psi$

If $\phi$ in $\phi \models_d \psi$ contains holes (i.e., represents some form of ambiguity), and $\psi$ is a disjunction of formulae of $LL$ that do *not* contain holes (for this we say that they are *fully specified*), then the disjuncts in $\psi$ represent all described readings. It is simple to define a translation function that directly reads off the base language formula from such a description formula, and so this notion of logical consequence connects an underspecified formula with all its base-language readings (via the translation function), and (via the definition of $\models_b$) with all its base-language models as well.[15] Note, however, that this definition does not provide us with a productive way of actually *computing* all disambiguations of an ULF. All this tells us is when something counts as a complete disambiguation (namely if it is a disjunction of all readings). The literature on CLLS gives many interesting results about the complexity of the problem of enumerating these readings which, given the similarity between $LL$ and their language can possibly be transferred.[16] But in any case, we're not interested in computing *all* readings (which, as we will see, will be impossible for fragments anyway, since they will have an infinite number of readings if viewed independently of the context in which they are uttered). Rather, we're interested in computing the one *preferred* reading (or maybe the $n$ most-preferred ones), given the discourse context. In Chapter 8 we will show how this can be done for our application. We simply note here, however, that the definition of discourse update will not require us to list all possible solutions of an $LL$-formula for a fragment.

### Stating Scope-Constraints

[15]In contrast, (Reyle 1993) has to define a complex supervaluation for his ULFs, whereas here we simply use the classical entailment relation.

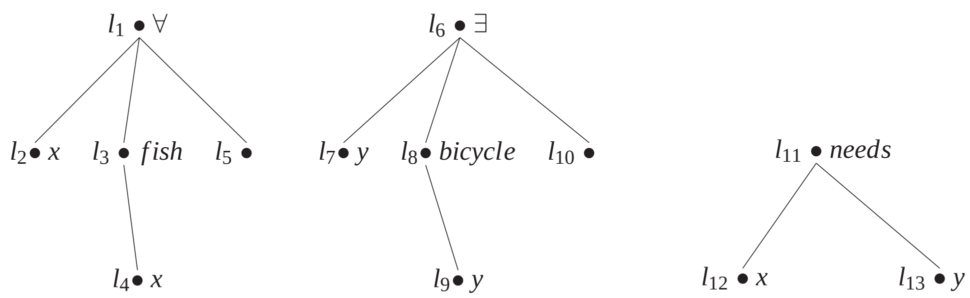
[16]Some references to papers related to complexity issues: (Niehren & Koller 1998, Egg et al. 2001, Erk, Koller & Niehren 2002).

Now that we have defined the semantics of our descriptions, we can briefly come back to the issue of how the scope-constraints in the descriptions should be specified. The question is how we can formulate (or let a grammar generate) descriptions that express what we want in a minimal way, or, to phrase it differently, how we pick out from a set of equivalent descriptions the one that expresses the desired constraints in the most efficient way (where efficiency can be defined as minimal number of literals, or even in relation to a particular semantic construction mechanism).

So, which outscopes-statements are minimally required? It should be clear that one does not have to fully specify the relation *Succ** via *outscopes* statements in order to successfully describe BLFs. E.g., if $R_\forall(l_1,l_2,l_3,l_4)$ is a conjunct of a ULF, a statement of $outscopes(l_4,l_1)$ is superfluous, since this is entailed by the former literal. So one principle guiding the generation of descriptions should be not to state outscopes-constraints that are entailed by other literals of the description in the same way.

However, the constraints that express the scopal-underspecification in the examples above are *not* entailed by other elements of the description. We have also already noted that certain ways of expressing those scope relations that underspecify the order of quantifiers are equivalent: (224) above expresses as a relation between the nuclear scope label and the VP-representation what (223) expresses as a relation between the label of the quantifier and that of the VP-representation. But are they *minimal* descriptions of the intended solutions? Given the definition of the structures that interpret those descriptions, they are not. Both representations explicitly state that the VP-representation is in the scope of the quantifier, but this isn't actually necessary: recall that *LΣS*s have to be trees, and the only way to form trees out of the material in the descriptions is to let the VP-representation be outscoped by the quantifiers. Hence, the final representation we adopt for our example is as shown in (229)—simply three unconnected bits of material.

(229)




Following (Copestake et al. 1999), we will assume one additional constraint on solutions for descriptions, namely that they must ensure that all occuring variables are bound. In the example we have seen so far (i.e. (229)), this constraint does not have an effect, but we will later see examples where the restrictor of quantifiers is underspecified as well, and where thus there is a possibility of quantifiers 'floating' to a position which disconnects them from occurrences of the variable they are supposed to bind. We will provide the definition below in Section 6.5.1, where we will need the necessary auxiliary definitions, and only note here that it defines a subclass of the solutions for a given description (the so-called *normal* solutions).[17] Finally, and for the same purpose of restricting where quantifiers can 'float' to, we adopt the convention from (Copestake et al. 1999) to explicitly denote $\ell_0$, the supremum of the partial order *Succ*, with $l_\top$. Again, for the example above this is not necessary, but we will need it in cases where we want to ensure that certain material outscopes all quantifiers.

This concludes our description of how scope-underspecification can be represented in a logical language. In the next section we will extend this language to deal with the underspecification needed to represent the compositional semantics of fragments. We are only concerned with the semantics here; we deal with details of the syntax/semantics interface in the next chapter.

## 6.4 The Logical Form of Fragments

### 6.4.1 A Constraint for Fragments

We said above that we want our compositional semantics of 'declarative' fragments to express that a) they resolve to propositions, of which b) the main predicate is unknown, but c) (some) of its arguments are specified, though their grammatical roles may not be. We gave (212)—repeated here as (230)—as a pseudo-formal representation of this information.

(230) $\exists e \exists x P(e,x) \wedge named(x, \text{“Peter”}) \wedge P = ?$

So, what does this condition "$P(e,x) \wedge P =\ ?$" translate to in $\mathcal{LL}$? Let's first try a naïve approach. Maybe we can use labels to directly express that the constructor $P$ is unknown, for example as in (231) (assuming that we introduce new elements '=' and '?')?

(231) $R_P(l_e, l_x, l) \wedge l = ?$

Unfortunately, this doesn't quite say what we want. What this means in $\mathcal{LL}$ is that the position of a subformula $P(e)$ in a larger formula is unknown, but it doesn't say anything about the value of $P$. The problem for this strategy is that the constructors do not receive labels on their own, only together with their arguments. But this is actually sufficient for our task and we don't have to extend the expressivity of the language, because we can make use of our knowledge about the *form* of all possible values for $P$, or, to put it differently, our knowledge about what all meanings that could possibly be intended by actual uses of the fragment in context have in common: we know that in all resolutions there will be a subformula in which the variables $e$ and $x$ occur (and occur *freely*, as we will see presently). These are notions that we can define straightforwardly in $\mathcal{LL}$, and so accordingly we introduce a new constraint

[17] Having this global constraint on solutions, however, might make it more complex to compute solutions. CLLS, in contrast, deliberately only puts local constraints on the structures (Erk et al. 2002).

into the $LL$ which encapsulates these demands on the subformula. We will call this constraint *unknown*, because it stands in for an unknown subformula.

We should add a word of caution here. Unlike formulae of $LL$ that only contain underspecification of scope, those that contain *unknown_rel*-constraints will describe, if they are satisfiable at all, a countably infinite number of BLF formulae. This is what we want, since it reflects the fact that viewed in isolation from the context there is an infinite number of propositions (or, depending on the sentence mood of the fragment, questions, or requests) that can be conveyed with such fragments. The compositional semantics specified here simply constrains the readings to make use of the material from the fragment in some way. But this of course makes it impossible to specify a procedure that actually *generates* the set of all readings in finite time; something which, as we have said, is possible for the fragment of $LL$ without $unknown\_rel$. However, as we mentioned above, we are not actually interested in *all* readings anyway. What we are interested in is a different relation, namely the one that relates to the ULF that particular reading that is pragmatically preferred in the given context. We will define this relation in Chapter 8.

We give a definition of the semantics of the constraint *unknown* below in Definition 7, but first we show here what the representation of the compositional semantics of our example fragment "Peter." now finally looks like. (232) give this representation in all three notations we have introduced (the relation notation; the infix notation; and the graphical notation, which we already have seen as (213-a) above).[18,19]

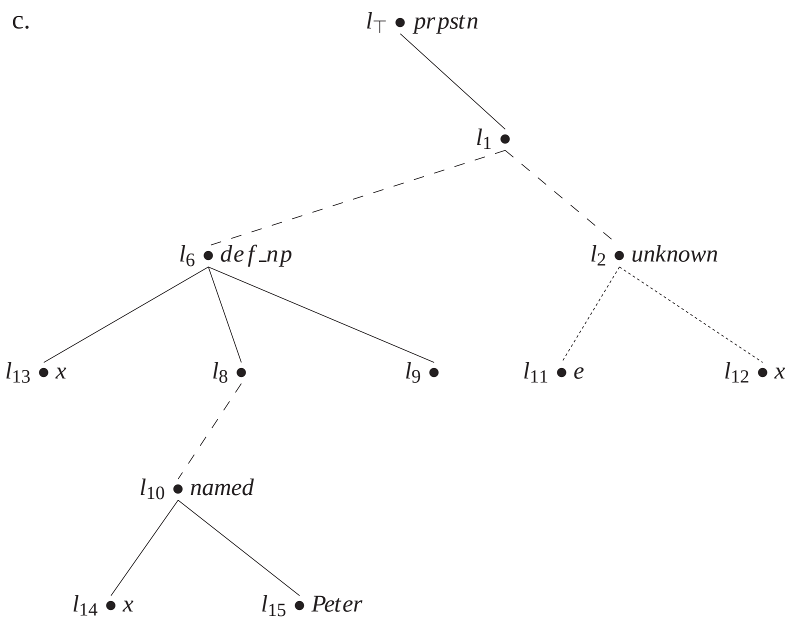
(232) a.
$$\begin{array}{lll} R_{prpstn}(l_1, l_\top) & \wedge\ unknown\_rel(l_{11}, l_{12}, l_2) & \wedge \\ R_e(l_{11}) & \wedge\ R_x(l_{12}) & \wedge \\ R_{def\_np}(l_{13}, l_8, l_9, l_6) & \wedge\ R_x(l_{13}) & \wedge \\ R_{named}(l_{14}, l_{15}, l_{10}) & \wedge\ R_x(l_{14}) & \wedge \\ R_{Peter}(l_{15}) & \wedge & \\ outscopes(l_8, l_{10}) & \wedge & \end{array}$$

b.
$$\begin{array}{l} l_\top : prpstn(l_1)\ \wedge\ l_2 : unknown\_rel(e, x) \wedge \\ l_6 : def\_np(x, named(x, \text{Peter}), l_9) \wedge \\ outscopes(l_8, l_{10}) \end{array}$$

[18]We use as base language here a predicate language that is inspired by the predicates in the "English Resource Grammar" (ERG, (Copestake & Flickinger 2000)). We will say more about this grammar below, so for now just note it uses in its LFs a predicate *prpstn* that outscopes all others and that signals that the formula denotes a proposition (rather than a question or a request), and also that there is a quantifier for definite NPs, aptly named *def_np*.

[19]Recall that, given the additional assumption about scoping possibilities made above, we do not have to explicitly state outscope constraints between $l_1$ and $l_6$ or $l_2$. For clarity, however, we will draw such relations in the constraint-trees.

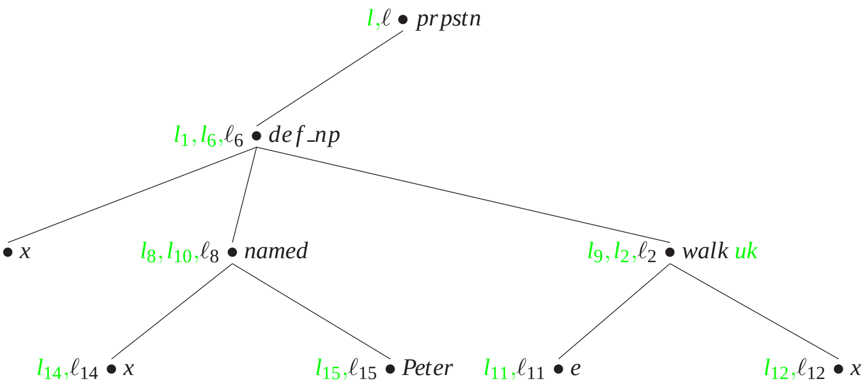
c.




Note that *unknown* is a constraint more like *outscopes* than like for example *def_np*, in that it constrains the configuration of the base language formula and does not get translated into a base language predicate. We draw it in the graphical representation in the same way as the translations of the base-language predicates, with the exception that its arguments are connected with a dotted line. This expresses that the arguments of *unknown* do not have to be immediately outscoped by the label of *unknown* (as arguments connected with solid lines have to), and on the other hand have to satisfy a stricter constraint than simply *outscopes* (which is indicated by dashed lines).

Let us now look at an example that indicates that this indeed expresses what we want, before we formally define the semantics of this new constraint. (233) shows an $L\Sigma S$ corresponding to the base language representation for the sentence “Peter walks.” (this was (213-b) above). This should be one possible resolution of the fragment “Peter.”—for example if the fragment is uttered in the context of the question “who walks?”—and so we’d expect this to be described by (232). (233) shows the $L\Sigma S$ corresponding to the LF of “Peter walks.”, and in colour one possible variable assignment. We also indicate the node which is the root for the *unknown*-subformula by marking it with *uk*. Since the variables $x$ and $e$ do occur unbound in the subformula ‘below’ $\ell_2$ (the denotation of $l_2$), this is indeed a solution for (232). Examples that show that (232) also describes BLFs where “Peter” is an object (e.g. “Marty loves Peter”), or an argument in a PP (e.g. “John made a picture of Peter.”), are given below in Section 6.4.2.

(233)

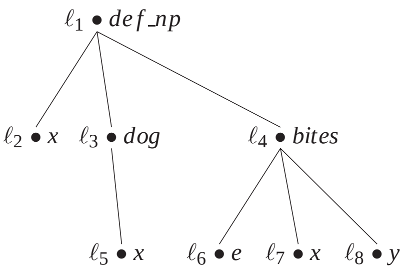



With this motivating example behind us, we can proceed to formally define the semantics of the new constraint. For this, we need a couple of auxiliary definitions. First, *labels*:

**Definition 5** *labels*
Given an $L\Sigma S$ $\langle U, Succ, I\rangle$, we say that $\ell$ *labels* a constructor $R \in \Sigma$ of arity $n$ iff $\langle \ell_1, \ldots, \ell_n, \ell\rangle \in I(R)$ (with $\ell, \ell_1, \ldots, \ell_n \in U$).

Further, we define two relations, *Free_in* and *First_Arg*, that will be used in the definition of the interpretation of *unknown_rel*. Roughly, *Free_in* formalises the base-language notion of free variables on $L\Sigma S$s. Formally, it is the set of all ordered pairs of labels where the former labels a variable that is free in the subformula labelled by the latter. We say that a label $\ell$ is free in a subtree whose root is $\ell'$ if $\ell$ is subordinate to $\ell'$ and does not label a constructor that is the first argument (i.e. the bound variable) of a quantifier which is subordinate to $\ell'$ and superordinate to $\ell$. Or, to put it simpler, $\ell$ is free in $\ell'$ if the variable labelled by $\ell$ occurs in the subformula labelled by $\ell'$, but is not bound in it. A label is in *First_Arg* if it labels an event-variable that is the first argument of some constructor $f$. These definitions are illustrated by (234) (a fragment of for example the LF of "a dog bites sandy").

(234)




$\ell_8$ in (234) is free relative to $\ell_1$, because it is not bound by any quantifier below $\ell_1$; $\ell_6$ is in *First_Arg* relative to $\ell_1$, since it is the first argument of *bites*; $\ell_5$ on the other hand is *not* free relative to $\ell_1$, since it labels the constructor $x$ which is the first argument (i.e., the bound variable) of a quantifier that is subordinate to $\ell_1$ (recall that subordination is reflexive).

We can now give the formal definition of these notions.

**Definition 6** *Free_in and First_Arg*
Given an $L\Sigma S$ $\langle U, Succ, I\rangle$,

- the binary relation *Free_in* consists of exactly the pairs $\langle \ell, \ell' \rangle \in U \times U$ such that
  1. $\ell \in U_{i^l}$; and
  2. $\ell' \succeq \ell$; and
  3. $\ell$ labels $R$; and
  4. there is no $R'$ (with arity $n$) $\in Q$ ($Q$ is the set of quantifiers in $\Sigma$), s.t.
     (a) $\exists \ell'' \in U$ where $\ell''$ labels $R'$, $\ell' \succeq \ell''$ and $\ell'' \succeq \ell$; and
     (b) there are $\ell''', \ell_2, \ldots, \ell_n \in U$, s.t. $\langle \ell''', \ell_2, \ldots, \ell_n, \ell'' \rangle \in I(R')$ and both $\ell$ labels $R$ and $\ell'''$ labels $R$.
- *First_Arg* is the binary relation consisting of exactly the pairs $\langle \ell, \ell' \rangle \in U \times U$ such that
  1. $\ell \in U_{e^l}$; and
  2. $\exists R^n \in \Sigma$ s.t. $\langle \ell, \ell_2, \ldots, \ell_n, \ell'' \rangle \in I(R)$, with $\ell, \ell_2, \ldots, \ell'' \in U$ and $\ell' \succeq \ell''$.

With the help of these relations we can now interpret the predicate *unknown_rel*:

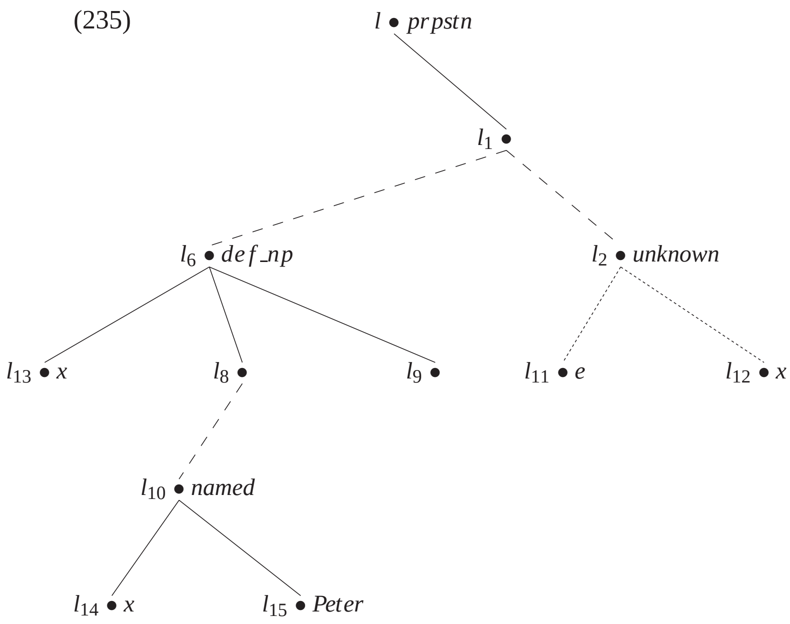
**Definition 7** *Interpretation of* unknown_rel
$\langle U, Succ, I\rangle \models^g_l$ *unknown_rel*$(l, l', l'')$ iff

i) $\langle [\![l]\!], [\![l'']\!] \rangle \in$ *Free_in* (given $\langle U, Succ, I\rangle$ and $g$),
ii) $\langle [\![l']\!], [\![l'']\!] \rangle \in$ *Free_in* (given $\langle U, Succ, I\rangle$ and $g$),
iii) $\langle [\![l]\!], [\![l'']\!] \rangle \in$ *First_Arg* (given $\langle U, Succ, I\rangle$ and $g$).

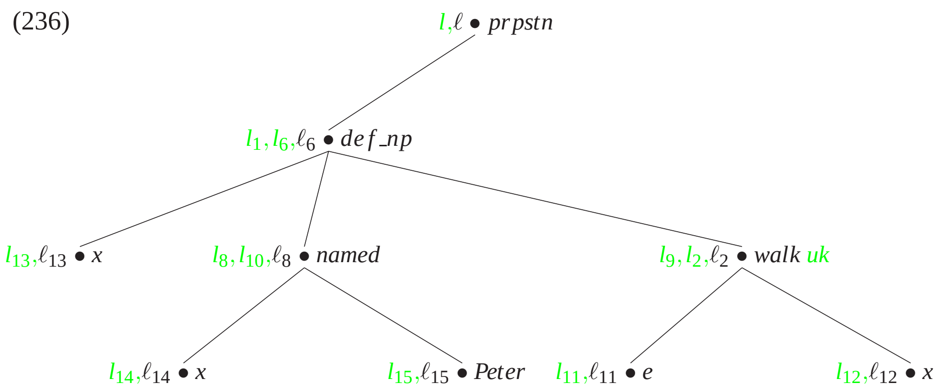
This definition makes sure that *unknown_rel* indeed constrains any subformula ‘below’ $l''$ to contain the variables labelled by $l$ and $l'$. Moreover, it also requires that these variables are free in that subformula. This is necessary, because otherwise we would also describe formulae where these variables are ‘captured’ by other quantifiers (remember that the fragment that introduces the variable will also introduce a quantifier; e.g. the fragment “Peter” will introduce a variable bound by a *def_np*-quantifier). Additionally, we have constrained the first argument, the event variable, to appear as the first argument of some predicate in the subformula. This makes sure that the event variable is indeed the main event of the subformula.

Let’s return to our example (232) now and see whether the $L\Sigma S$ and variable assignment shown in (233) does indeed satisfy it, given the semantics of *unknown* we have now defined. (We repeat the two graphs here as (235) and (236).)

(235)




(236)




(237) shows the relations *Free_in* and *First_Arg* as specified by (236).

(237) $\{\langle \ell_{14}, \ell_8 \rangle, \langle \ell_{12}, \ell_2 \rangle, \langle \ell_{11}, \ell_2 \rangle, \langle \ell_{11}, \ell_6 \rangle, \langle \ell_{11}, \ell \rangle\} = Free\_in$
$\{\langle \ell_{11}, \ell_2 \rangle\} = First\_Arg.$

With a variable assignment as indicated above in (233), that $L\Sigma S$ does indeed satisfy $unknown(l_{11}, l_{12}, l_2)$: (a) $\langle [\![l_{11}]\!], [\![l_2]\!] \rangle \in Free\_in$ is true; (b) as is $\langle [\![l_{12}]\!], [\![l_2]\!] \rangle \in Free\_in$; and (c) also $[\![l_{11}]\!] \in First\_Arg$ holds.

### 6.4.2 More Examples

In Chapter 1 we said that the syntactic form of the utterances we accept as fragments can be described by the rule 'S-frag → (ADV) XP'. That is, we allow all kinds of phrases, possibly modified by an adverb, to stand alone as fragments. So far, we have only seen an example of the semantics of an NP-fragment. In this section we will go through all kinds of fragments our grammar rules will admit and show how we represent their compositional semantics, again just giving a promissory note that the representations can actually be built on the syntax/semantics-interface. We show the $LL$-representation of the compositional semantics of the fragments, and then illustrate with a number of examples that all kinds possible interpretations in context are described by these representations. We will for now fix the message-type to be *prpstn*; later we introduce a way of underspecifying the message-type as well.

For concreteness, we have to align ourselves in these examples more closely than before with particular analyses of linguistic phenomena other than fragments. As mentioned above, we will draw upon the analyses delivered by the ERG, which is the grammar whose implementation we will extend in the next chapter to produce the underspecified logical forms shown here.[20] However, it should be stressed that the particularities of these analyses are quite orthogonal to our point and nothing of principle hangs on choosing them, and so our general method could be just as easily adapted to other grammars as well.

**NP-fragments** We have already seen an example of an NP-fragment, and so we concentrate here on an additional aspect of such fragments that so far hasn't been mentioned. In the following example, it seems to be ambiguous what the intended meaning of the fragment (238-b) is in the context of the question (238-a): it could be either of the readings of the sentence "every fish needs a bicycle" we discussed above where we introduced $LL$ for scope-ambiguities.[21] From this follows that the representation of the fragment must not exclude one of the readings, i.e., it must not make any decisions as to how the fragment-phrase will be scoped in the described BLF.

(238) a. A: What does every fish need?
b. B: A bicycle.

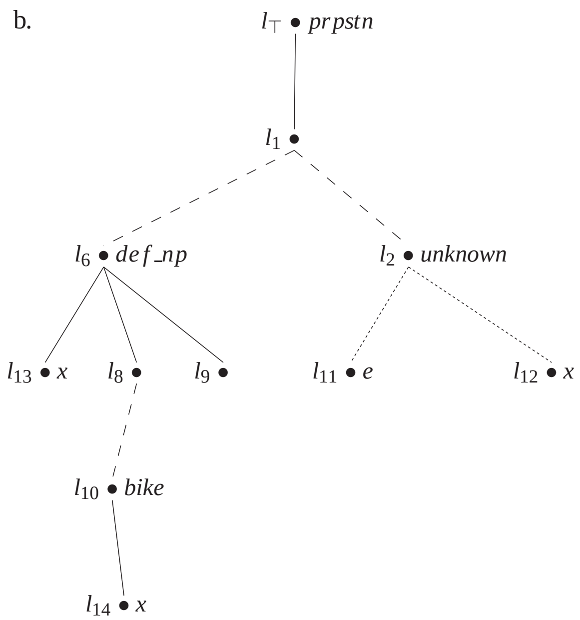
Example (239) shows our representation of the compositional semantics of (238-b) in the familiar formats (we will from now on drop the long notation).

(239) a. $l_\top : prpstn(l_1) \wedge l_2 : unknown\_rel(e,y) \wedge$
$l_6 : a\_quant(y,l_8,l_9) \wedge l_{10} : bicycle(y)$
$outscopes(l_8,l_{10}), outscopes(l_1,l_2)$

[20] For now, we make two major changes: whereas in the representations in the ERG more than one predicate can be labelled with one label, we split such representations into ones using $\wedge$; secondly we continue using *outscopes* rather than the scope-constraint used in the ERG. These differences will be discussed in detail below in Section 6.5.

[21] One reading seems to be strongly preferred here, but that is the case for most single sentences with scope ambiguities as well. As long as there is a possibility to get a different reading, maybe helped by contextual clues, this additional reading has to be represented.

b.



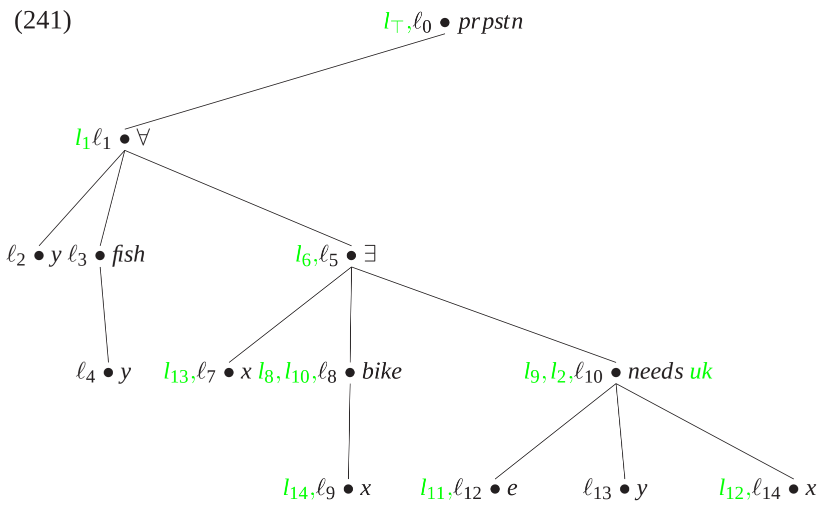

What we want to show is that this formula does indeed describe the two readings repeated in (240)—amongst an infinite number of other BLFs, of course.

(240) a. $\forall(x, fish(x), \exists(y, bicycle(y), need(x,y)))$

b. $\exists(y, bicycle(y), \forall(x, fish(x), need(x,y)))$

We have already shown in the previous section the *LΣS* tree corresponding to (240-a), but since we have changed the base language here a bit (we now have event-variables, and a *prpstn*-predicate), we show the adapted version below in (241). We also put the variable assignment in that tree as well, so that (241) is actually a representation of a *solution* of (239); the constraint $l_6 : unknown(e,x)$ is satisfied, since there is a free occurrence 'below' $l_6$ of both variables, and also $outscopes(l_5, l_6)$ is satisfied, since $\ell_6 \succeq \ell_{10}$. (241) nicely illustrates the fact that there can be nodes in the *LΣS* that are not referred in the $LL$-constraint; in this case, these are the nodes of the tree-fragment representing the universal-quantifier.

(241)




We omit showing here that the other reading of (238-b) is also described by the $LL$ formula (239); one can transfer (241) into an $L\Sigma S$-tree/solution corresponding to that reading simply by exchanging the two subtrees representing the quantifiers and adapting the variable assignment. An assignment of $\ell_5$ to $l_2$ would *not* be a solution of (239), because then $x$ would not be free in the subformula, contrary to what *unknown* demands.

This has shown that the representation of the compositional semantics of NP-fragments does not make any assumptions about scope ordering; these two sources of ambiguity are kept separate in the representation. Below in Chapter 8 we will see that we can define a resolution mechanism that directly works on ULFs and so allows us to resolve these ambiguities separately.

So far we have only seen examples where the label of the *unknown*-constraint has been resolved to the label of that predicate that also has both the variables that are arguments of *unknown* as arguments. For example, in (241) above, the predicate that is at the position $\ell_{10}$ in the BLF (which is denoted by the label of *unknown*, $l_6$) has as its direct arguments both the variables $e$ and $x$ of which *unknown* demands that they must occur unbound in a position subordinate to $l_6/\ell_{10}$. Given these examples, one could be led to assume that what we intend here is some sort of substitution process, where *unknown* is *replaced* by some predicate that has the arguments of *unknown* as its own arguments. This is not the case, as an example of the fragment (238-b) ("a bicycle") in the context of a question like "What did John say Peter mentioned that every fish needs?" shows: here the label of the *unknown*-rel will be resolved by the label of the predicate representing "say" (because that carries the main event-variable), but the individual variable denoting a bicycle will be an argument of the predicate for "needs", which is deeper down in the scope of "say" and not a direct argument of it. We will encounter a similar situation in the next section where we discuss PP-fragments, and so we just summarise the point here by stressing that *unknown* really must be understood as a complex constraint saying something about a potentially very complex (base-language) subformula rooted in a certain position in the overall formula.

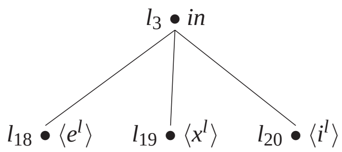
**PP-fragments** We now turn to the representation of PP-fragments like (242).

(242) “in the park.”

We first note that such fragments are (lexically) ambiguous in that the preposition can be of one of two different types, namely it can be a lexical or a functional preposition. We have already encountered these types above in Chapter 2.2, and we won’t argue for making this distinction here;[22] all we need to say here is that we assume that this is a *syntactic* ambiguity, so that we assume two parses and hence two representations for this fragment. Let’s look at the LF for the parse where “in” is a lexical preposition first, and then return to functional prepositions later.

Before the representation of (242) is shown, a few words about the treatment of such prepositions in the ERG are in order. The (lexical) preposition “in” for example is represented in the ULF by a three-place predicate *in*, where the first argument is an event, the second is either an event or an individual, and the third one is an individual, as shown graphically in (243) (we will explain the meaning of the expressions in brackets in a minute).

(243)



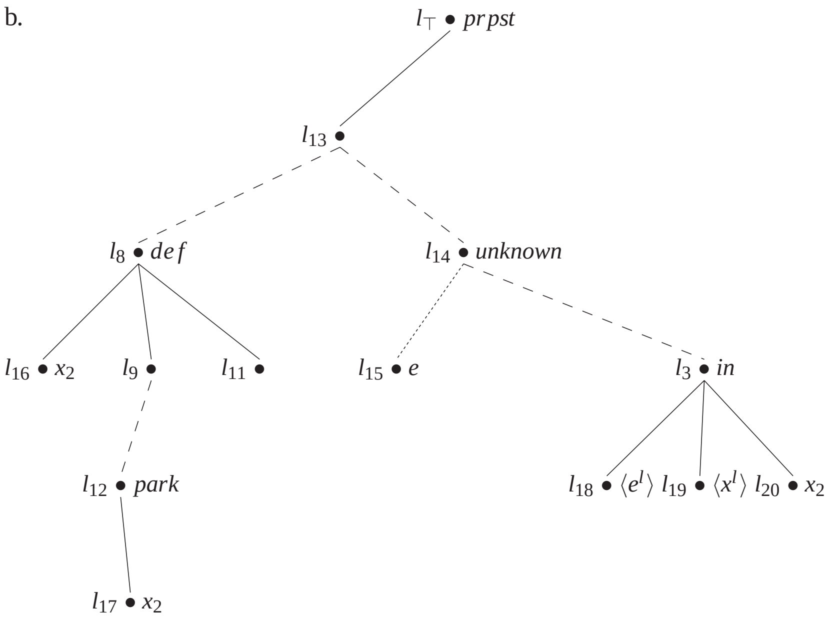

This predicate *in* is used both for prepositions that modify nouns and those that modify VPs. The last argument is always the variable denoting the entity referred to by the object of the preposition—in our example (242) that is the park. In the case where the PP modifies a noun the second argument will be the variable denoting the referent of that noun; in the other case the argument will be the variable denoting the event referred to by the VP that is being modified. This explains how this same predicate can be used for both these cases. Finally, the first argument slot is used in cases where the PP is the complement of the copula (e.g. in “Peter is in the park.”), in which case it is filled with the main event variable. (In other words, in such sentences there is no explicit representation of the copula, but rather the preposition-relation takes on an event-denoting function.) We will encounter all three variants of PP-meanings in the examples below. In our $L\Sigma S$-representations we chose to carry over to the base-language this overloading of the preposition-relations, but one could also follow a strategy where according to number and type of arguments different base-language expressions are used.

Now, (244) shows the $LL$-representation of the semantics of the lexical-preposition parse of (242). (We will explain the new elements presently.)

[22] See for example (Tseng 2000) for a recent review of arguments for and against this dichotomy.

(244) a. $l_\top : prpstn(l_{13}) \wedge l_{14} : unknown\_rel(e) \wedge$
$l_3 : in(?_{e'}, ?_{x'}, x_2) \wedge$
$l_8 : def(x_2, park(x_2), l_{11}) \wedge$
$outscopes(l_{13}, l_8), outscopes(l_9, l_{12}), outscopes(l_{11}, l_{14}), outscopes(l_{14}, l_3)$

b.




The first thing to note is that we use a variant of the *unknown*-constraint here which takes only one argument; the semantics of this constraint is like that given in Definition 7, save that clause ii) is dropped. The object of the preposition—in this case that is "the park"—is 'connected' to the unknown event via the predicate for the preposition, "*in*". This predicate in turn is constrained to be subordinate to the label of *unknown*.[23] The first two arguments of this proposition-predicate are underspecified (i.e., are variables of the underspecification language that do not label anything); we will see below why this is necessary. We make use here of the different sorts of labels for individuals and events, and so demand that $l_{18}$ is of sort **event**, and $l_{19}$ of type **index**, which is a supertype subsuming **event** and **individual**.[24] What we want to express is that at these positions in the BLF, i.e. as arguments to that relation, there are occurrences of variables of this type. We could try to be more specific here, since we know that the second argument of a preposition relation in the base language must be a variable that occurs elsewhere in the formula; as mentioned above and illustrated with the examples below, it is either a variable denoting the event that is being modified, or it is an individual that stands in that prepositional relation to the object of the preposition. However, that would require us to formulate an additional, global constraint (it must say something about the whole formula) that would complicate the semantics of

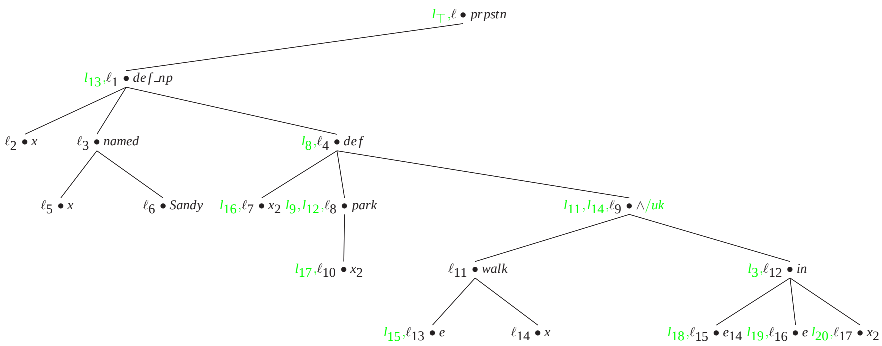
[23] We will see later that we have to adapt this slightly.

[24] To represent this in the tree we have annotated the nodes with the types of the variables, using $x^l$ for the supertype **index**; in the $LL$-formulae in infix-notation this is represented with question marks subscripted with the type.

these representations. We chose not to do this here, and accept the consequence that such descriptions ‘overgenerate’, as it were, in that they describe BLF that do not correspond to grammatical sentences. As discussed above, this is tolerable, as long as our resolution mechanism can be guaranteed to make the right restrictions.

We now show that this constraint describes the *LΣS* (245), which represents the sentence “Sandy walks in the park.”.[25] As before, we represent the intended variable assignment that together with this *LΣS* forms the solution of (244) in the same tree, in a different colour. This combination of *LΣS* and assignment satisfies the constraints in the *LL*-formula (244): the variable *e* occurs freely in the subformula below $l_{14}$ (if $[\![l_{14}]\!] = \ell_9$), and the preposition is subordinate to that label.[26] Note that this is another example where the argument of *unknown* is not a direct daughter of the denotation of the label of *unknown*; i.e., the predicate ‘replacing’ *unknown* is complex.

(245) “Sandy walks in the park.”

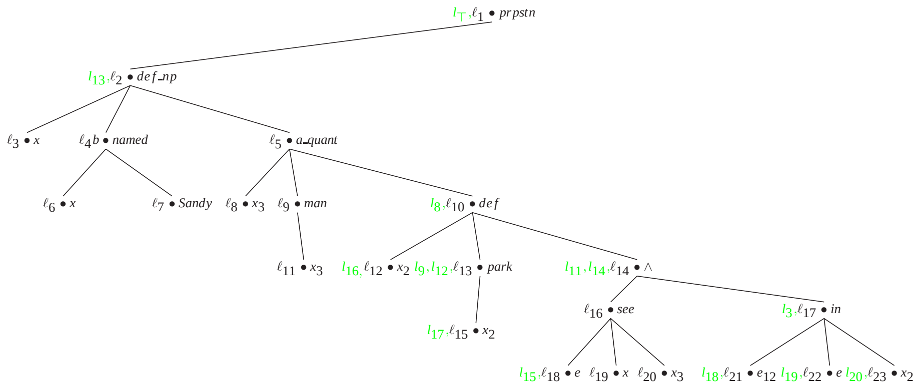



The next two examples show how we satisfy one of the desiderata from Chapter 2. We said there that a fragment like “in the park”—for example uttered as an elaboration of an assertion “Sandy saw a man”—preserves the PP-attachment ambiguity, i.e. is ambiguous between a reading where the event of seeing took place in the park and a reading where the seen man was located in the park. (246) is a representation of the former reading; we show that this is a solution of our constraint for “in the park” in the usual way by indicating the intended variable assignment as well. The reader is invited to check that given this assignment all constraints in (244) are satisfied.

[25] It actually only represents one reading of the sentence, since the quantifiers *def* and *def_np* create an ambiguity, which however is spurious and does not need to concern us here.

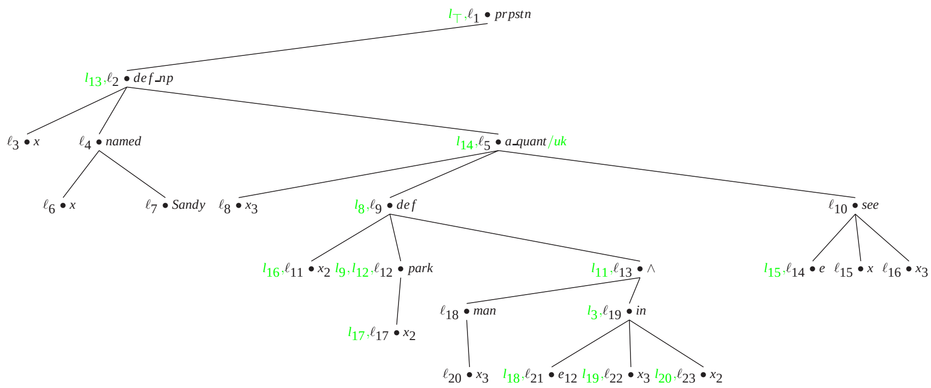
[26] Often there will be more than variable assignment that together with a given *LΣS* forms a solution. Given the *LΣS* in (245) for example, assignments that assign $\ell_4$ or $\ell_1$ to $l_{14}$ also solve the constraint. In such cases we will always show the solution where the handle of *unknown* gets assigned the lowest node in the tree such that the constraint is satisfied.

(246) “Sandy ((saw a man) in the park).”



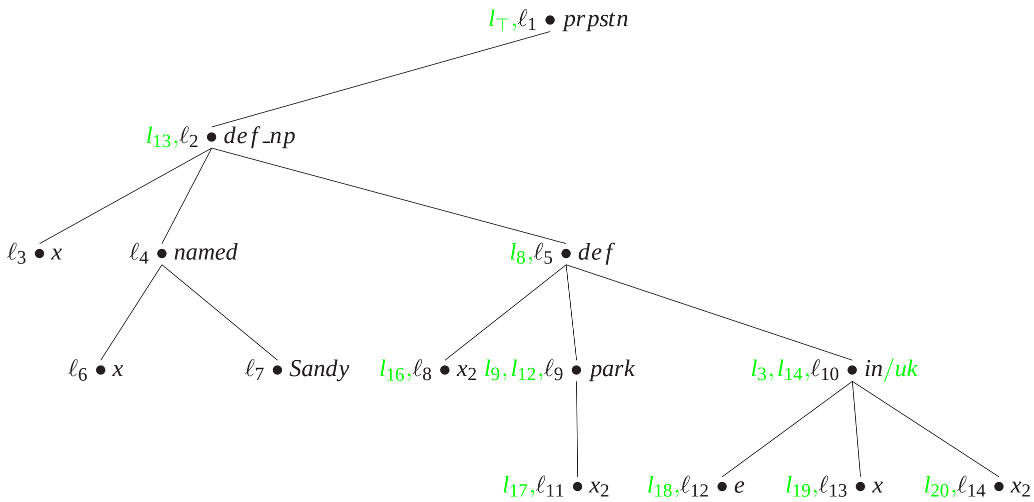

(247) below is an $L\Sigma S$-representation (plus variable assignment) of the other reading, which corresponds to a parse of the sentence “Sandy saw a man in the park” where the noun ‘man’ is being modified, rather than the whole VP.

(247) “Sandy saw a (man in the park).”




The next example shows what could be called a limiting case in the resolution of PP-fragments. As mentioned above, one characteristic of the ERG is that it does not assume a specific predicate for the copula in constructions like “Peter is in the park”, but rather sees the preposition-predicate as the main predicate which carries the event-variable. (248) gives the $L\Sigma S$ for such a reading, again with the assignment that shows that it is a model for (244) printed in colour. The important part here is that the label of *unknown*, $l_{14}$, is identified with that of the preposition-relation; and that this relation carries the main event-variable. Note that *outscopes* is reflexive, so an assignment of $\ell_{10}$ to both $l_{14}$ and $l_3$ will make $outscope(l_{14}, l_3)$ trivially true.

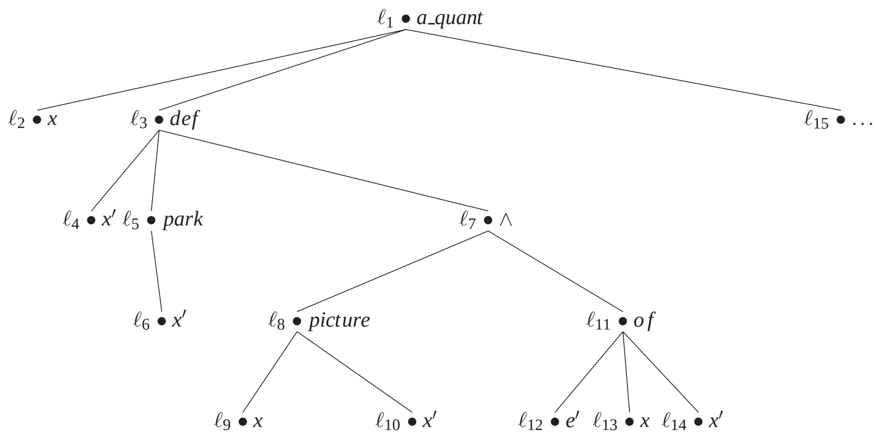
(248) “Sandy is in the park.”




This example shows why we used a supertype for the second argument of such preposition-relations: here, as in (247) above, it is an individual variable that denotes an entity that is in this relation, whereas above in (246) it was an event variable denoting the activity that is modified by the relation.

Finally, as a last example of what lexical preposition-fragments can describe, we now show how our representation for the PP-fragment “of the park” can resolve to readings where the PP is a complement of a noun, as for example in “Peter painted a picture of the park”. We forgo showing the $LL$-representation of the fragment (it is like (244), only with a different preposition-relation), and only show a fragment of the $L\Sigma S$ of (one reading of) the resolution, corresponding to “a picture of the park”. This representation is similar to (247) above (the example where the PP modified a noun), with the difference that here the representation of the noun has a second argument (i.e., $x'$ occurs twice).

(249)



We said at the beginning of this section that prepositions come in two types, lexical and functional. Instances of PP-fragments of the former type have been shown in the last examples, and so we now turn to functional prepositions. Here one might wonder why they should be represented in a *semantic* representation at all; after all, the basis of the dichotomy is the claim that they are only functional. A common position is expressed in the following quote from (Pollard & Sag 1994, p.255): "[...] the head preposition makes no contribution to the CONTENT of the PP [...]". In the ERG, however, such prepositions are represented in the underspecified logical forms. The original reason to do this is that it makes the grammar monotonic in the sense that every lexical item introduces a representation in the LF;[27] we will later make another use of this feature in our resolution mechanism. In any case, having these predicates in the ULF does not mean making a decision about whether they are contentful or not: these predicates can still be seen as describing a tautology in the base language, i.e. as not having any semantic impact.[28] For this reason, we will represent fragments with functional prepositions as shown in (250). We chose a slightly different example here than (242), namely "on Sandy", just because it is more plausible as a fragment with a functional preposition (for example as an answer to "Who can we rely on?").

(250) a. $l_\top : prpstn(l_{13}) \wedge l_{14} : unknown\_rel(e,x) \wedge$
$l_3 : on\_s(?_{e'}, ?_i, x) \wedge$
$l_8 : def\_np(x, l_9, l_{11}) \wedge l_{12} : named(x, Sandy)$
$outscopes(l_{13}, l_{14}), outscopes(l_9, l_{12}), outscopes(l_{14}, l_3)$

Note that functional prepositions are distinguished in these representations by an *_s*-suffix, and also that we use the version of *unknown* that takes two arguments. So apart from the presence of this preposition-relation this representation is not different from that for an NP-fragment; and hence should we choose not to represent the preposition in the BLF the correct set of readings is still obtained.

We close by noting that our resolution mechanism will give us a principled way of choosing between parses of for example "on Sandy" where the preposition is analysed as functional, as above, and parses where it is seen as lexical, since the appropriate one will result in a more coherent update than the other.

**VP-fragments** This section deals with the compositional semantics of fragments consisting of VPs. Here we make a distinction between VP[*bse*] and VP[*inf*], because their representations will differ slightly. We will also present an example of an S[*comp*]-fragment here.

We begin with the ULF of a VP[*inf*]-fragment. In the ERG, the complementizer "to" introduces a relation

[27] More precisely, it allows one to underspecify the type of preposition during construction, without then having to 'remove' the (functional-)preposition-relation from the LF, as for example an approach following (Nerbonne 1996) would do. For the claim that monotonicity is a desirable property for a grammar cf. for example (Shieber 1986).

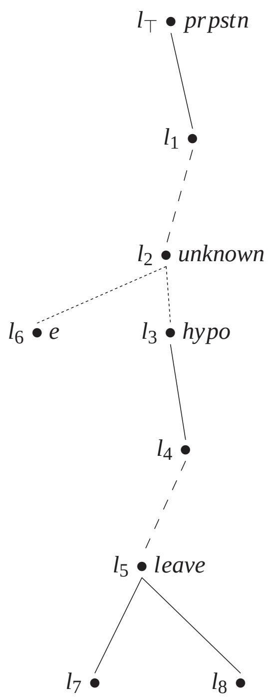
[28] But note that this makes the translation function $\nu$ from ULFs to BLFs many-to-one, and hence non-invertable.

*hypo* (roughly, this describes a modal operator meaning that something holds hypothetically), which is superordinate to the representation of its argument VP. We make this *hypo*-subformula the second argument of *unknown*.

(251) “to leave”

a. $l_\top : prpstn(l_1) \wedge l_2 : unknown\_rel(e, l_3) \wedge$
$l_3 : hypo(l_4) \wedge$
$l_5 : leave(?_{e'}, ?_x) \wedge$
$outscopes(l_1, l_2), outscopes(l_4, l_5)$

b.

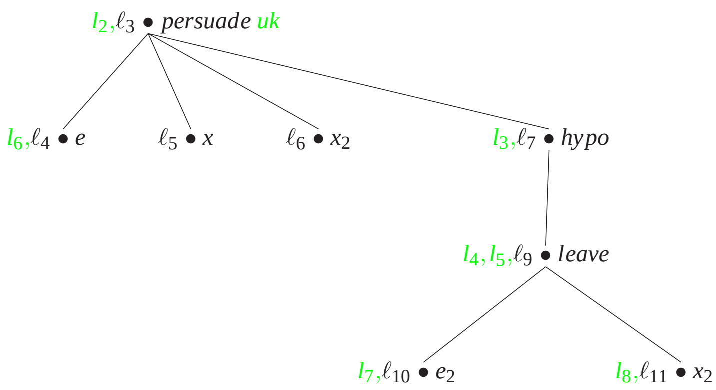



This second argument of *unknown*, in this example $l_3$, is of sort $t^l$ (because it labels a formula of the base language), and so we see here a third variant of the *unknown*-constraint. Its interpretation again is very similar to what is given in Definition 7, all we do is drop clause ii).[29] Again we have holes in this representation—the arguments of the predicate representing the verb—, they are of type **event** ($l_7$) and **individual** ($l_8$), and again we could choose to express additional constraints on the values of these variables, in this case that $l_8$, the individual, must resolve to a variable that occurs elsewhere in the formula: it is the subject of this verb, which must be provided by the same sentence. But as before will rather leave it to the resolution-mechanism to only produce resolutions that satisfy this constraint.

[29]This overloading of *unknown* is harmless, since the number and type of argument will always disambiguate which version is intended.

The next example now shows a fragment of the $L\Sigma S$ for “Peter persuaded John to leave.”, again adorned with a variable assignment, which turns it into a solution of (251).[30] (For reasons of space we have left out the two quantifiers corresponding to the proper names.)

(252)



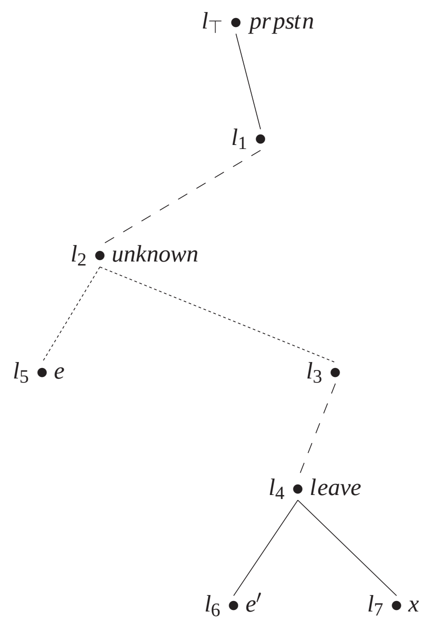

The following $LL$-formula represents the VP[*bse*]-fragment “leave”. Note that an utterance of this string will be ambiguous between a fragment-reading and a reading as imperative. As with functional- and lexical-prepositions, here again our resolution mechanism will offer a principled symbolic way of deciding between the parses, because given a particular context presumably one reading will produce a more coherent update than the other.

(253) “leave” (fragment, not imperative reading)

a. $l_\top : R_{prpstn}(l_1) \wedge l_2 : unknown\_rel(e, l_3) \wedge$
$l_4 : R_{leave}(?_{e_2}, ?_x) \wedge$
$outscopes(l_1, l_2), outscopes(l_3, l_4)$

[30] Note that getting right the control exhibited in (252) (the presuadee is the agent of the leaving) is a question for the resolution mechanism; here we simply show that an LF where the control is present is a solution for (251).

b.




The main difference between this $LL$-representation and that of the VP[*inf*] above is that here the VP-representation is a direct argument of *unknown*, and there is no intervening *hypo*-relation. We won't show here how this constraint describes for example the LF of "Peter made Sandy leave.", since that would be very similar to what we have seen above.

Finally, we subsume under this heading "VP-fragments" also fragments that consist of complement-sentences:

(254) "That Sandy likes Kim."

In the ERG, such complement sentences are distinguished by introducing an additional *prpstn*-relation which is the argument of the verb-relation that takes them as complement. This means that our representation will be very much like that for VP[*inf*]s given in (251), except that the *hypo*-relation is substituted by a *prpstn*-relation.

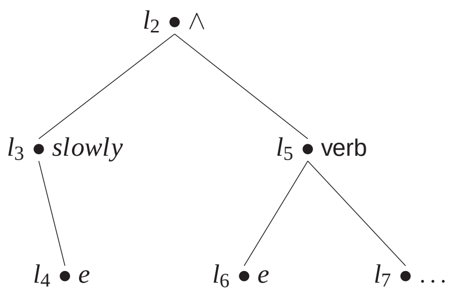
**ADV-fragments** Amongst fragments consisting of adverbs we distinguish between intersectively modifying adverbs and those that modify scopally. A representation of an instance of the former type is the following.

(255) "Slowly"

$l_0 : prpstn(l_1) \wedge$
$l_2 : unknown\_rel(e) \wedge$
$l_3 : slowly\_rel(?_{e'}) \wedge$
$outscopes(l_1, l_2) \wedge outscopes(l_2, l_3)$

The relation representing the adverb will in all described formulae be a sister of the verb-relation that is being modified and a daughter of $\wedge$, as in the schematic representation of an $L\Sigma S$-fragment in (256). The modified verb, however, does not have to be the matrix-verb of the resolved meaning (the one that has the main event-variable *e* as an argument), as the two readings in (257) show, and so we have to underspecify its event-argument with a hole of appropriate type.

(256)

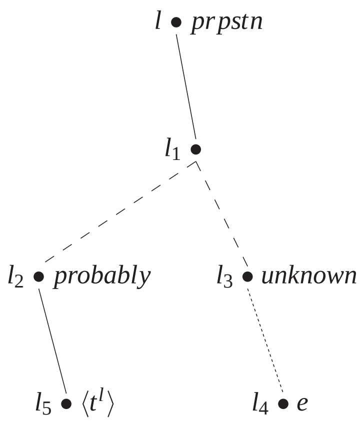



(257) A: How did Peter think John talked?
B: Slowly.

a. Peter slowly thought that John talked.
b. Peter thought that John talked slowly.

We represent scopally modifying adverbs as follows (here we only show the tree-representation, because it will illustrates a point we will make later more clearly).

(258)



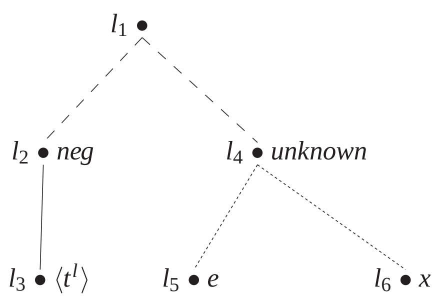
Note that the relation corresponding to the adverb is a *sister* to *unknown* here rather than its mother or daughter, as in most representations we have seen so far. Why is that? Because we want the *probably* to be able to be outscoped by material which comes from the resolution. For example, in a comment-fragment like in the following exchange (259), we do not want the representation of the fragment to make any restrictions on where the *probably* ends up scopally. Now, the semantic difference between a reading where the probably modifies the "can" and one where it does the "come" might seem slight here—we will see below when we discuss modification of fragment-phrases an example where the difference is much more prominent—but as a matter of principle we do not want to exclude anything at this stage, in the representations, that is at all possible.

(259) A: Can Peter come?
B: Probably.

**Modification** We now come to fragments that are modified by adverbs. Let's begin with a special case, namely with negation. The representation of the fragment in (260), of which the part containing the *unknown*-relation is shown in (261), looks very much like that of stand-alone scopal adverbs shown above in (258). Again we do not want to restrict the scope-possibilities of the adverb-relation, and so we must make sure that it can be outscoped in the resolution.[31]

(260) A: Who can work?
B: Not Peter.

(261)




There might be in most cases a strong preference for the reading where the negation has widest scope, but that is just a preference, and so the other reading (where the negation is outscoped by "can", but outscopes "work") should not be excluded from the set of described BLFs. The following example shows that this preference can be overridden (a reading where Alastair is good at not working seems to

[31] One could argue that the latter example, being of the form 'ADV VP' should be treated differently from the former ('ADV NP'), and only in the VP-fragment should the negation be allowed to float. This seems to capture the fact that for (260) it is exceedingly difficult, if not impossible, to imagine a reading that can be paraphrased as "Peter can (not work)". In the interest of generality in the syntactic analysis (see next chapter), we will however tolerate this possible overgeneration of readings here. Again we should point out that our resolution mechanism should make sure that only intended readings are produced; all we have to make sure here is that the intended readings are described by the fragment-LF.

be readily available).

(262) A: What can Alastair do really well?
B: Not work.

**Underspecifying the message type** In all the examples so far we have assumed that we have information about the semantic type of the resolved fragment, namely that it has to resolve to a proposition. This strategy is not implausible, since intonation for example will serve to make this clear, and can be considered syntactic information in the widest sense; or at least information to which the grammar has access. However, it is no problem to underspecify the message-type as well in our representations, if that is desired. In the ERG, questions feature an *int*-relation instead of $prpstn$, imperatives an $imp$-rel. We can now simply introduce a predicate *message* that goes where in our examples so far we had $prpstn$, and we can see this as a supertype which can be specified to either $prpstn, int$ or $imp$. Representations that use this kind of message will describe formulae of all these types.

However, there are lexical items that can occur as fragments and that fix the semantic type of their resolutions, namely question words like “who”, “where” etc. For them we will have to fix the message to $int$. As we will see in the next Chapter, we will treat “who”, “what” and “which N”-phrases as NPs syntactically, and the other *wh*-words as PPs with special preposition-relations; “how” can in certain uses be an AP. We give here two more explicit examples and a list of the preposition-relations and restrictions used in the other *wh*-words.

(263) “who?” (similarly, “which boy?” etc.)
$l_0 : int(l_1) \wedge$
$l_2 : unknown\_rel(e,x) \wedge$
$l_6 : which\_rel(x,l_8,l_9) \wedge l_{10} : person\_rel(x) \wedge$
$outscopes(l_1,l_2) \wedge outscopes(h_8,h_{10})$

(264) “when?”
$l_0 : int(l_1) \wedge$
$l_2 : unknown\_rel(e,x) \wedge$
$l_3 : unspec\_loc\_rel(e',x) \wedge$
$l_6 : which\_rel(x,l_8,l_9) \wedge l_{10} : temp\_rel(x) \wedge$
$outscopes(l_1,l_2) \wedge outscopes(l_8,l_{10}) \wedge outscopes(l_2,l_3)$

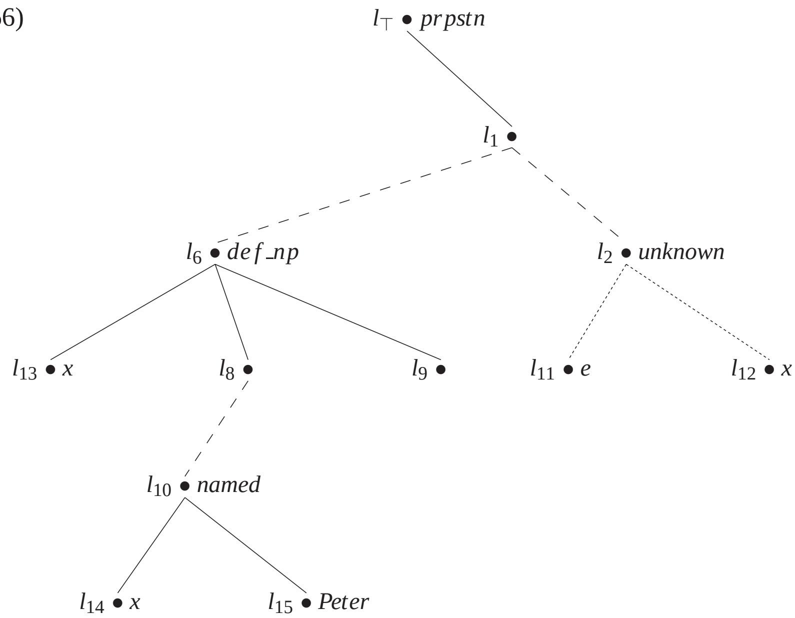
(265) “where”: $unspec\_loc\_rel, which\_rel, place\_rel$
“how”: $unspec\_manner\_rel, which\_rel, way\_rel$
“why”: $for\_rel, which\_rel, reason\_rel$

## 6.5 Revisions, and an alternative notation for $LL$-formulae

### 6.5.1 Linking variables to their quantifier

There is one problem with our representations which we haven't mentioned so far. Whereas for scope-underspecification one can safely consider only minimal solutions, our approach to the underspecification of the meaning of fragments crucially relies on the possibility of having nodes in the $L\Sigma S$ that aren't referred to in the $LL$-formula. (See earlier discussion.) Only this allows us to have an infinite number of solutions for these formulae. However, so far we do not exclude BLFs where a quantifier that isn't mentioned in the ULF intervenes and 'captures' variables. To illustrate this problem, let's look at the constraint for NP-fragments from above, which we repeat here as (266).

(266)

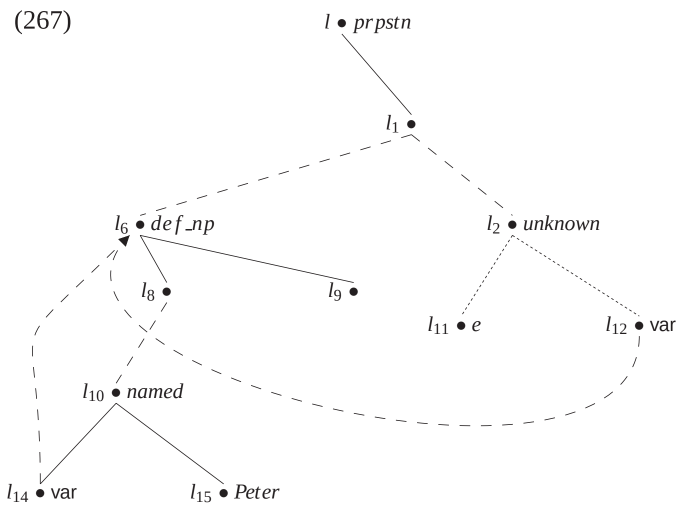



The problem is that this formula happily describes a BLF where there is a quantifier between (the denotations of) $l_9$ and $l_2$ that binds the variable $x$, thereby severing the link between the quantifier described by this constraint and the argument of *unknown*.[32] To avoid this, we take on board an idea from CLLS (Egg et al. 2001), namely to use an explicit 'linking-constraint' to represent variable binding, rather than to rely on variables. This linking constraint will connect an argument place that is to be filled by a variable with the quantifier that binds this variable. (267) below shows a graphic representation of this for

[32]How severe this problem is seen to be depends on what we expect from the ULFs of fragments. We have said at several places that the main desideratum is that all possible resolutions of fragments in all possible contexts are described by them—but this doesn't mean that there can't be more in the set of described BLFs. As we have set up the semantics, we also ensure that we only describe well-formed BLFs. However, the problematic readings *are* well-formed, but can never be resolutions of fragments. What we indicate in the following is a way of excluding them from the set of described BLFs. As a side-effect, it will give us the option to use CLLS parallelism-constraints later as one alternative to define the resolution of *resolution-via-identity*-fragments.

NP-fragments. The quantifier-relation now has one argument less, and where we had a base-language variable before we now have a symbol var; the link between this and the quantifier is symbolised with a dashed arrow going from the var-symbol to the label of the quantifier. Informally, this constraint will be satisfied by all $L\Sigma S$ where the quantifier outscopes the variable that fills the position of the label of the var-symbol, and this quantifier does actually bind this variable (i.e., the variable is not 'captured' by another quantifier).

(267)




We are now done with the description of the revisions, and can formalise them. First, we have to explicitly express in an $LL$-formula what the top-label is; i.e., we now define these formulae as *pairs* now, of which the first element is this label and the second is a formula as given by Definition 3. The semantics of the new constraint *binds* and the further constraints on the class of $L\Sigma S$are defined as follows.

**Definition 8** *labels a bound variable*
Given an $L\Sigma S$ $\langle U, Succ, I\rangle$, we say that an $\ell_v \in U$ labels a bound variable if $\ell_v \in I(\xi)$ for some variable $\xi \in \Sigma$, and there is an $\ell$ (with $\ell_\top \succeq \ell$) s.t.

1. $\ell \succeq \ell_v$; and
2. $\ell$ labels a quantifier $Q$ of $\Sigma$, $\langle \ell', \ell'', \ell\rangle \in I(Q)$, and $\ell'$ labels $\xi$.

**Definition 9** *Interpretation of binds*
$\langle U, Succ, I\rangle \models^g_l binds(l', l)$ iff

1. $[\![l]\!]$ labels a base-language variable $\xi \in \Sigma$; and
2. $[\![l']\!]$ labels a quantifier $Q \in \Sigma$; and
3. $\ell' \succeq \ell$; and
4. there are $\ell', \ell'' \in U$, s.t. $\langle \ell', \ell'', [\![l']\!]\rangle \in I(Q)$ and $\ell'$ labels $\xi$;

I.e., $Q$ binds $\xi$

and

5. there is no $\ell''' \in U$ s.t. $[\![l']\!] \succeq \ell''', \ell''' \succeq [\![l]\!]$ and $\ell'''$ labels a quantifier $Q'$ that binds $\xi$.

   I.e., there is no intervening quantifier that 'snatches away' $\xi$.

**Definition 10** *Normal Solution*

We say that an $L\Sigma S$ $\langle U, Succ, I\rangle$ is a *normal solution* with respect to an $LL$-constraint $\langle l_\top, \Gamma\rangle$ and a variable assignment $g$, iff

1. $\langle U, Succ, I\rangle \models^g_l \Gamma$; and
2. for all $\ell \in U$, $[\![l_\top]\!] \succeq \ell$; and
3. For all $\ell \in U$ s.t. $\ell$ labels a variable of $\Sigma$, $\ell$ labels a bound variable.

We have to put in this third constraint, because even though our *binds*-constraint will take care that all variables that are explicitly mentioned in the constraint are properly bound, we have to make sure that in $L\Sigma S$s that contain nodes which aren't denoted by the $LL$-constraint there are no unbound variables.

### 6.5.2 MRS: an alternative notation for $LL$-formulae

We now introduce the representation format our grammar will use, MRS (Copestake et al. 1999). (268) shows such a representation for the sentence "Every dog barks." What we called label-variable above is called *handle* in MRS, and is notated as $h_n$. An MRS is a quadruple consisting of the 'top-handle' (which correspondes to our $l_\top$); a variable of type **event** which denotes the main event; a bag called the LISZT, containing labelled relations; and a bag of 'handle constraints', which are the constraints on scope. Note that MRSs uses a constraint $=_q$ (read as 'qeq'), we will discuss the difference of this to the *outscopes* we have so far used in a minute; for now we will assume that this is just a different notation for *outscopes*.

(268)
$$\begin{array}{llll} \langle & h, e, & \{ & h_1 : prpstn(h_2), \\ & & & h_3 : bark(e,x), \\ & & & h_7 : every(x, h_9, h_{10}),\ h_{11} : dog(x) \quad \}, \\ & & \{ & h_2 =_q h_3,\ h_9 =_q h_{11} \quad \} \quad \rangle \end{array}$$

The similarity of this representation to our collapsed notation for $LL$-formulae is of course intended; this MRS-style notation is what we have been working towards in this chapter. We can mechanically derive formulae of $LL$ from such MRS, by following the steps in (269), and hence we will from now on regard MRSs just as a notational-variant for our $LL$-formulae.[33]

(269) 1. Introduce new labels for non-handle arguments;

[33] With the caveat that a single MRSs can stand for more than one $LL$-formula, as explained in (269). In any case, MRSs never contain *less* information than we need, and so we can safely move from them to $LL$-formulae.

2. relying on the fact that variable names are unique, read off the *binds*-constraints from the MRS;
3. MRSs underspecify the order of conjuncts by putting conjoined relations on the same label, this has to be translated into $n! - 1$ $LL$-formulae (with $n$ being the number of conjuncts) with binary-trees rooted in $\wedge$; and finally,
4. in MRSs, the scope-constraints are notationally separated from the EPs; we simply conjoin these.

One final thing that is different in MRS, however, is that (Copestake et al. 1999) uses a stricter constraint to express scope-relations than the *outscopes* we have used. They define a variant of *outscopes* which they call *qeq* (written $=_q$), which stands for 'equality modulo quantifiers.' To give an example of its intended semantics, for $l_1 =_q l_2$ to be true there can only be quantifiers 'between' $l_1$ and $l_2$, other scope-bearing relations are not allowed to intervene. For our purposes this constraint is too strict. As discussed above in the section of ADV-fragments and modified fragments, we have to allow other scope-bearing elements (for example modals) to intervene. Hence, we must use the more general *outscopes*-constraint.

We close this section with a side remark. (Copestake et al. 1999) claims that the stricter scope-constraint they use simplifies construction; for this claim, however, they do not give any evidence. According to (Flickinger, p.c.), the evidence for having this stricter constraint is given by some German examples which can only be handled if *qeq* is available. We will not further discuss this here, and simply use *geq* (standing for 'greater or equal', i.e. for our *outscopes* relation) in our representations, regardless of of whether the ERG for example really needs *qeq* or not.

## 6.6 Summary

In this chapter we have introduced our semantic representations for fragments. The basic idea was to use *underspecification*, i.e. descriptions of logical forms rather than logical forms directly. We have described a particular formalisation of this description idea, taken from (Asher & Lascarides 2003), and then extended it for our purposes. In a nutshell, we added a constraint that restricts in a very general way described formulae to use certain semantic material. For example, for the NP-fragment "Peter." our *underspecified logical form* will describe any base or 'real' logical form which denotes an event in which Peter is in some way involved. It should be clear that such a semantic representation can be constructed in a fully compositional manner: nothing more is required than information about the semantic contribution of the phrase. *How* it can be constructed will be the topic of the next chapter.

# Chapter 7

# A Coherence-Based Approach II: A Grammar of Fragments

In this chapter we show how the representations of the compositional semantics of fragments that we have introduced in the previous chapter can be built on the syntax/semantics-interface. Before we do this, however, we take a step back and discuss in a more abstract way the relative merits of two strategies for describing the syntactic form of fragments: either as phrases that somehow get the same status as sentences, i.e. are allowed to 'stand alone' by the grammar; or as special kinds of sentence-constructions. We will give arguments in favour of the latter strategy, and then turn to implementing it. We first show how MRSs—the structures we introduced in the previous chapter as short-hands for our more detailed $LL$-representations—can be represented and built in feature structure based grammars in general; then in Section 7.3 we develop in detail our HPSG-encoding of the construction type 'fragment'. In Section 7.4 we describe our implementation of the rules and evaluate it under real-world conditions.

## 7.1 What is the syntactic form of fragments?

In this section we discuss the relative merits of two strategies for describing the *syntactic* form of fragments, namely either as phrases that somehow get the same status as sentences, in that they are allowed by the grammar to 'stand alone'; or as special kinds of sentence-constructions. So the question we're trying to answer here is: what *are* fragments, syntactically? What is (270-b), for example? Is it a *sentence*?

(270) a. Who came to the party?
b. Peter.

Clearly, (270-b) conveys a message; this observation was the starting point for our investigation into the phenomenon "fragment". Does that make it a sentence, though? What is a sentence, anyway? This, surprisingly, turns out to be quite a controversial question. Take this definition from a dictionary of linguistics (Matthews 1997, p.337): "Usually conceived [. . . ] as the largest unit of grammar, or the largest unit over which a rule of grammar can operate." (270-b) obviously falls short of this definition; after all, it is just an NP, and they are not the largest units of grammar. The following remark from an introductory syntax textbook (Sag & Wasow 1999, p.72) takes a slightly different angle: "Somewhere along the line, of course, an adequate grammar will need to specify the well-formed 'stand-alone' utterances of the language: the phrases that can be used in isolation to express a complete message are those whose mother is of the category 'S'." Now, (270-b) does express a complete message, but does it do that 'in isolation'? Well, it clearly doesn't, but neither do sentences like those in (271). So, what now?

(271) a. She hates him.
b. He does, too.

What is causing the confusion here is the mixture in these definitions of syntactic notions (sentence) and semantic or pragmatic ones (can stand alone). The question is what is to be seen as primary: if sentences are defined as those expressions that convey messages, then (270-b) *is* a sentence; we can however also choose to give this up as a distinctive feature of sentences, and allow other phrases, like the NP in (270-b), to 'stand alone'.

Let's bring this question down from this rather metaphysical level to a more technical one: what is the right parse for fragments? Is it enough to parse them as the phrases they are, or do we need to do more?[1] Or, in the terminology of (Morgan 1973), as reviewed in Chapter 4, are fragments 'base-generated', or is something else at work? What are the consequences of either decision?

[1]Of course, in the previous chapter we have actually already decided on our answer to this question, by spending much effort on defining constraints that help to formalise a semantics for such fragments that is quite different from that of the phrase itself. So we will only add more arguments here in favour of this decision, and look briefly into a differing analysis.

We first look at the option of allowing all kinds of phrases as ‘stand-alone’ utterances; this approach, as mentioned above, is advocated for example by (Barton 1990).[2] Following this strategy, we can imagine an approach that assigns our fragment (270-b) the logical form (272). The first advantage of such an approach is that we do not have to change any rules on the syntax/semantics interface; this phrase gets the same interpretation as a fragment as it gets as part of a larger phrase.

(272) $\lambda P.P(p)$

However, as already discussed in Section 6.2, one disadvantage of such an approach is that these representations do not encode the information that there is a difference between fragments and phrases, namely that the intended meaning of the former is of a different type than the meaning of the latter. Such an approach must leave the task of recognising this entirely to pragmatics. We further claimed that it complicates the discourse processing, since now fragments can be of all sorts of semantic types, and so the speech-act rules would have to additionally account for that. Finally, pragmatics would then have to reason with logical forms (and all their consequences) themselves; given the need for consistency checks in pragmatic reasoning this would make it uncomputable.

These are semantic and pragmatic arguments against this strategy, but we can also find quite convincing syntactic arguments that favour a syntactic analysis of fragments as (special kinds of) sentences. (273-a) shows a phenomenon that has been called ‘sluicing’ by (Ross 1969): a *wh*-phrase as a complement to a verb.

(273) a. I don’t know where.
b. Where?
c. I don’t know where Peter is.
d. *I don’t know in the park.

We will analyse this construction as involving our $Elab_q$-fragments from Chapter 2,[3,4] and so the ‘where’ in (273-a) is analysed in the same way as that in (273-b). As (273-c) shows, ‘know’ can take a full sentence as a complement, in this case “where Peter is”.[5] It can *not*, however, take a PP as com-

[2]We said above that her arguments in favour of this approach are very (GB-)theory dependent, and so not really relevant for our discussion here.

[3](Ginzburg & Sag 2001) similarly analyse such sluices as embedded fragments.

[4]In fact, Ross’s (1969) original sentence examples, e.g. (i-a), seem to us simply to give a confirmation that a condition on question-asking holds (namely that the questioner does not know the answer), which is required because of implicatures that it might not.

(i) a. There’s a party but I don’t know where.
b. There’s a party.. But, where?

Someone who knows that there is a party is also likely to know where, and the violation of this expectation in (i-a) has to be explicitly signalled with a contrastive discourse particle. This is all that is required, as the reduced variant (i-b) illustrates.

[5]That is an inverted sentence, to be sure, but still it is a sentence in most accounts.

plement, as the ungrammaticality of (273-d) demonstrates. This strongly favours a syntactic analysis of the embedded fragment “where” in (273-a), and hence of (273-b) as well, as a sentence, and not a PP.

A similar argument can be made with short-answers. (274-a) shows a short-answer as a complement of a verb that takes sentential complements, but not NP-complements.[6]

(274) A: Who will do this?

- a. B: John said Peter.
- b. B: John said (that) Peter will do this.

So to summarise, there are good syntactic and semantic reasons to parse fragments as sentences. But how shall we do this? How do we express that there are (syntactic) elements ‘missing’? There are again two options here: we could posit phonologically empty elements that occupy the missing positions, or we could devise special constructions that take phrases to sentence-level. Take fragment (270-b) again, repeated here in dialogue (275).

(275)
- a. Who came to the party?
- b. Peter.

The fragment is resolved in this context to something that can be paraphrased as “Peter came to the party.”—i.e., the fragment fills a subject position. This could lead us to conclude that there is a VP ‘missing’ in (275-b), and to analyse the syntactic structure of (275-b) as shown in (276).

(276) $[_S\ [_{NP}\ \text{Peter}]\ [_{VP}\ \emptyset]\ ]$

However, that would force us in (277-a) to analyse what seems to be the same fragment differently syntactically, in the way illustrated by (277-b)—and again differently for examples where the fragment is resolved to be a complement of an NP for example, or as an adjunct. This is clearly not desirable, and so we decide against this strategy.

(277)
- a. A: Who does Kim like?
  B: Peter.
- b. $[_S\ [_{NP}\emptyset]\ [_{VP}\ [_V\emptyset]\ [_{NP}\ \text{Peter}]\ ]\ ]$

[6]The situation is not as unambiguous here as in (273), though, since one could claim that the short-answer in (274-a) is only quoted (but cf. “John said himself”), or that this is a version of “said” that does take NPs. We will come back to the question of the proper analysis of examples like this one below.

This leaves us with the strategy we will implement in the following section: using *constructions* to realize the pseudo-formal rule $S[frag] \rightarrow (ADV)XP$.[7] But before we come to this, we have to say a few words about how the semantic formalism we use, MRS, can be incorporated into the semantic component of an HPSG.

## 7.2 Background: Representing and Building MRSs in feature structure based grammars

In the last section of the previous chapter we showed how we can view MRSs as a notational variant for our description formulae, and we gave (268) (repeated here as (278)) as an example of such an MRS, representing the meaning of the sentence "every dog barks."

(278)
$$\begin{array}{llll} \langle & h,e, & \{ & h_1 : prpstn(h_2), \\ & & & h_3 : bark(e,x), \\ & & & h_7 : every(x,h_9,h_{10}),\ h_{11} : dog(x) \quad \}, \\ & & \{ & h_2 =_q h_3,\ h_9 =_q h_{11} \qquad \} \quad \rangle \end{array}$$

In this section here, as a preparation for the description of our grammar of fragments, we show how such MRSs can be represented in feature-structure-based grammars.[8]

First, the predicates or *elementary predications* (EPs). To give an example, the EP for 'walk' will be represented by a feature structure as shown in (279), with the relation being represented by the type of the structure, the arguments by features, and the label or 'handle' by a feature as well.[9]

(279)
$$\begin{bmatrix} \textit{walk} & \\ \text{HNDL} & \textit{handle} \\ \text{EVENT} & \textit{event} \\ \text{ARG1} & \textit{ref-ind} \end{bmatrix}$$

In actual instances of such EPs, the function of the variables that are the arguments of the EPs will be filled by structure-sharing; i.e., slots that are filled with the same variable in representations like (278) are simply co-indexed in feature structure representations of MRSs. To give a concrete example, (280) show the MRSs from above as a feature structure. Note how the different components of the MRS-quadruple are represented by different features of a structure of type *mrs*; feature names that probably

[7] A similar rule is used as the syntactic basis in (Ginzburg 1999*b*), and also (Ginzburg & Sag 2001) can be understood as a formalisation of such a rule (although they do not deal with modification).

[8] The presentation here mainly follows (Copestake et al. 1999, Sec.5).

[9] The argument-slots ARG1–ARG3 roughly correspond to thematic roles, but the nomenclature is 'semantically bleached' (Copestake et al. 1999, p.16), to avoid too close an alignment with a specific theory.

need further explanation are SOA for 'state of affairs', BV for 'bound-variable' in quantifiers, where RESTR is the restriction and SCOPE the scope.

(280)
$$\begin{bmatrix} \textit{mrs} & \\ \text{TOP} & \boxed{1} \\ \text{INDEX} & \boxed{4} \\ \text{LZT} & \left\langle \begin{bmatrix}\textit{prpstn} & \\ \text{HNDL} & \boxed{1}\\ \text{SOA} & \boxed{2}\end{bmatrix}, \begin{bmatrix}\textit{bark} & \\ \text{HNDL} & \boxed{3}\\ \text{EVENT} & \boxed{4}\\ \text{ARG1} & \boxed{5}\end{bmatrix}, \begin{bmatrix}\textit{every} & \\ \text{HNDL} & \boxed{7}\\ \text{BV} & \boxed{5}\\ \text{RESTR} & \boxed{9}\\ \text{SCOPE} & \boxed{10}\end{bmatrix}, \begin{bmatrix}\textit{dog} & \\ \text{HNDL} & \boxed{11}\\ \text{INST} & \boxed{5}\end{bmatrix}\right\rangle \\ \text{H-CONS} & \left\langle \begin{bmatrix}\textit{qeq} & \\ \text{SC-ARG} & \boxed{2}\\ \text{OUTSCPD} & \boxed{3}\end{bmatrix}, \begin{bmatrix}\textit{qeq} & \\ \text{SC-ARG} & \boxed{9}\\ \text{OUTSCPD} & \boxed{11}\end{bmatrix}\right\rangle \end{bmatrix}$$

This has shown how complete MRSs are represented by feature structures. Semantic composition, i.e. the process of *building* these representations, is similarly straightforward (Copestake et al. 2001). As a general principle in the grammar we will look at (the ERG), an MRS of a phrase will always contain all EPs and all QEQs/GEQs of its daughters, plus probably some additional scope-constraints (hence, it is monotonic). The connections between the EPs are made by adding the appropriate co-indexations. To give an example, the representation of the NP "every dog" is assembled in the way indicated in Figure 7.1 (the general rule is shown in Figure 7.2). The rule for combining the two arguments—to be neutral about whether the determiner categorises for the noun or the noun for a specifier—co-indexes the indices (here $\boxed{5}$), and thereby takes care that the restriction does indeed restrict the variable that is bound by the quantifier. It is important to note here that the combination rule adds the appropriate scope-constraint between the quantifier and the restriction; we will make extensive use of this ability of rules to add semantic material when we show how fragments are described by our grammar.[10] The fact that the scope of the quantifier is not co-indexed with anything captures the capacity of quantifiers in MRS to 'float' to any scope, as explained in the previous chapter.

We haven't said anything so far about how we represent our fragment constraint *unknown*. It doesn't quite fit in either category: it is not a scopal constraint, and hence H-CONS is not really the appropriate slot, but, given its semantic interpretation in Definition 7 from the previous chapter, it is not an EP either. However, since it makes construction simpler (and since we don't distinguish between LZT and H-CONS in the semantics of MRS anyway), we will put *unknown* in LZT. (281-a) shows the MRS for the

[10]Note that the principle of compositionality is preserved here: one can simply understand the contribution of the construction (here the addition of *qeq*-constraints; in Figure 7.2 this is the feature C-CONT) as coming from a daughter which has only semantics constraints, and which combines with the other daughter(s) in the usual way. This is the approach taken for example in (Copestake et al. 2001).

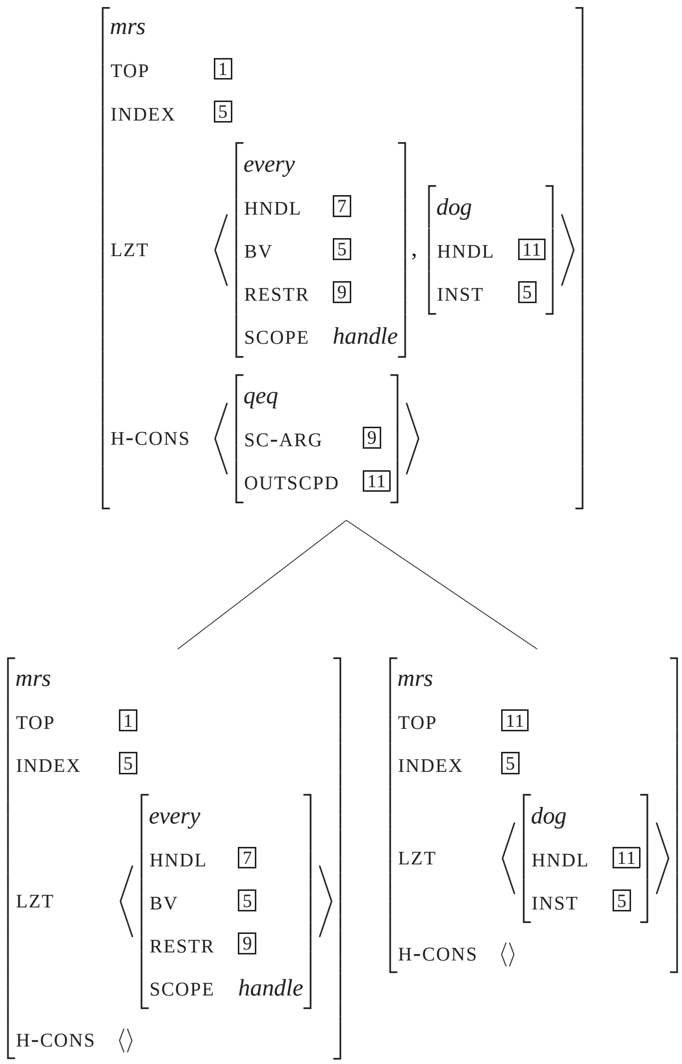



Figure 7.1: Composing the semantics of “every dog”

$$\begin{bmatrix} \textit{headed-binary-phrase} & \\ \text{CONT} & \begin{bmatrix} \text{TOP} & \boxed{1} \\ \text{INDEX} & \boxed{2} \\ \text{LZT} & \boxed{A} \oplus \boxed{B} \oplus \boxed{C} \\ \text{H-CONS} & \boxed{D} \oplus \boxed{E} \oplus \boxed{F} \end{bmatrix} \\ \text{C-CONT} & \begin{bmatrix} \text{TOP} & \boxed{1} \\ \text{INDEX} & \boxed{2} \\ \text{LZT} & \boxed{C} \\ \text{H-CONS} & \boxed{F} \end{bmatrix} \\ \text{HEAD-DTR} \mid \text{CONT} \begin{bmatrix} \text{LZT} & \boxed{A} \\ \text{H-CONS} & \boxed{D} \end{bmatrix} & \\ \text{NON-HEAD-DTR} \mid \text{CONT} \begin{bmatrix} \text{LZT} & \boxed{B} \\ \text{H-CONS} & \boxed{E} \end{bmatrix} & \end{bmatrix}$$

Figure 7.2: Semantic Composition of headed binary phrases

NP-fragment "every dog" in the tuple notation from the previous chapter, (281-b) as a feature structure (description).

(281) a. $$\begin{array}{llll} \langle & h, e, & \{ & h_1 : prpstn(h_2), \\ & & & h_3 : unknown(e,x), \\ & & & h_7 : every(x, h_9, h_{10}),\ h_{11} : dog(x) \quad \}, \\ & & \{ & h_2 =_q h_3,\ h_9 =_q h_{11} \qquad \} \quad \rangle \end{array}$$

b.
$$\begin{bmatrix} \textit{mrs} & \\ \text{INDEX} & \textit{index} \\ \text{LTOP} & \boxed{2} \\ \text{LZT} & \left\langle \begin{array}{l} \begin{bmatrix}\textit{prpstn-rel} & \\ \text{HNDL} & \boxed{2} \\ \text{SOA} & \boxed{3}\end{bmatrix}, \begin{bmatrix}\textit{unknown-rel} & \\ \text{HNDL} & \boxed{4}\\ \text{EVENT} & \boxed{1} \\ \text{ARG} & \boxed{5}\end{bmatrix}, \\ \begin{bmatrix}\textit{every} & \\ \text{HNDL} & \textit{handle}\\ \text{BV} & \boxed{5}\\ \text{RESTR} & \boxed{9}\\ \text{SCOPE} & \boxed{10}\end{bmatrix}, \begin{bmatrix}\textit{dog} & \\ \text{HNDL} & \boxed{11}\\ \text{INST} & \boxed{5}\end{bmatrix} \end{array} \right\rangle \\ \text{H-CONS} & \boxed{C}\left\langle \begin{bmatrix}\textit{geq} & \\ \text{SC-ARG} & \boxed{3}\\ \text{OUTSCPD} & \boxed{4}\end{bmatrix}, \begin{bmatrix}\textit{qeq} & \\ \text{SC-ARG} & \boxed{9}\\ \text{OUTSCPD} & \boxed{11}\end{bmatrix}\right\rangle \end{bmatrix}$$

## 7.3 An HPSG of Fragments

We have two goals in this section. First, we want to encode the

$$S[frag] \rightarrow (ADV)XP$$

rule in our grammar-formalism of choice; that is, we want to license phrases, possibly modified by adverbs, as special kinds of sentences. These sentences are special syntactically, because they lack a matrix VP, but also semantically, and capturing this is our second goal: we want to get the grammar to build the semantic representations for fragments that we introduced in the previous chapter.

We begin with showing an instance of the type of sign we use for NP-fragments, and then we ‘dissect’ this rule into those parts that are common to all (syntactic) types of fragments, and those that are specific to the different types. We will then organise the differences in a multiple-inheritance hierarchy, to achieve a parsimonious representation of our grammar of fragments.

Note that the signs shown here are only meant to give an impression of how the general approach would work in a generic HPSG (with constructions); the actual details of how the rules interface with existing analyses are shown in the section on the implementation below.

### 7.3.1 NP-fragment signs

Figure 7.3 shows, in a tree representation, the sign for the NP-fragment “Peter.” It shows how the NP (we have left out the rule that raises the proper noun to NP-level) is lifted to the level of sentences,

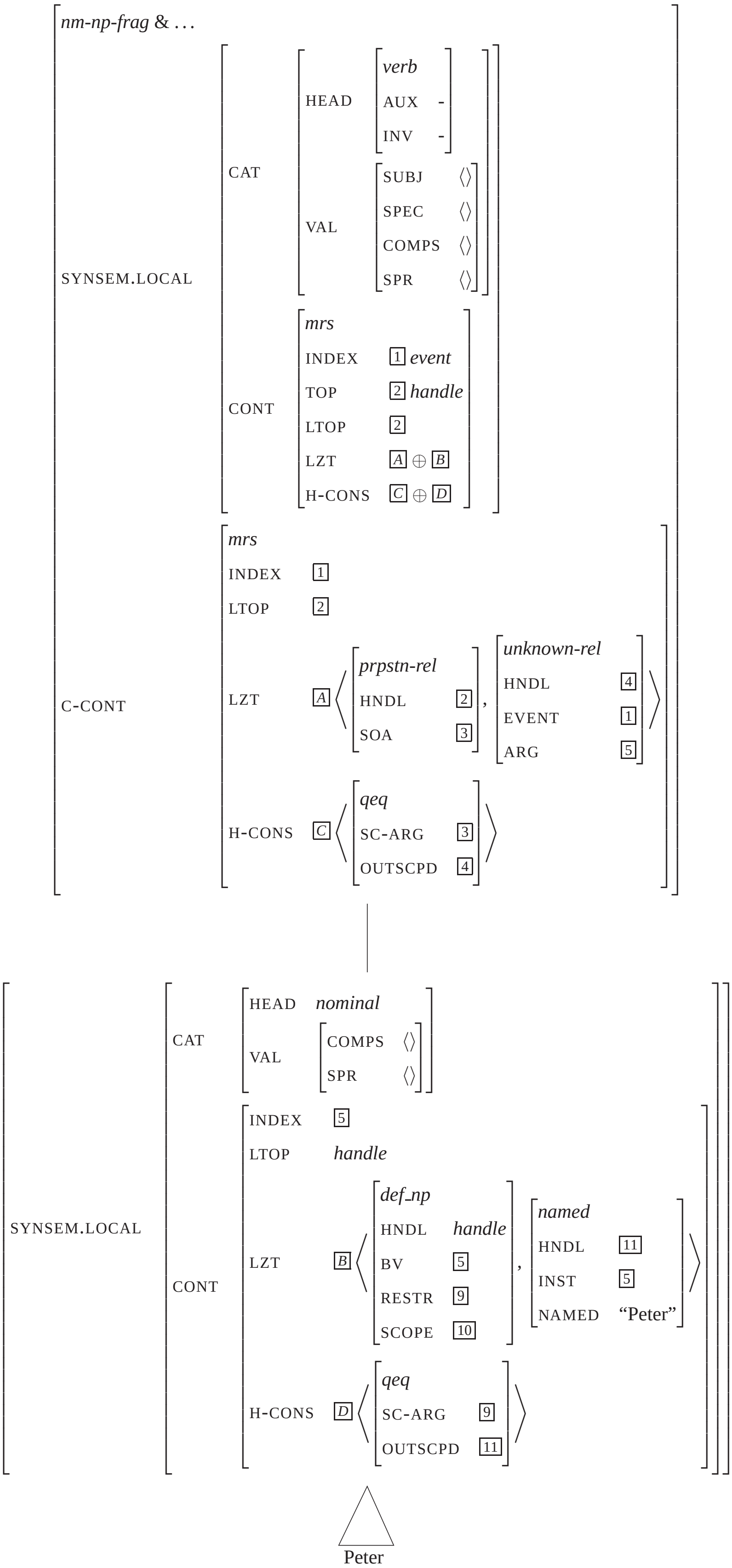



Figure 7.3: The sign for the declarative fragment “Peter”, in rule notation

and how the semantics of that sentence is composed. Let's work our way down from the top in the FS to describe this in detail. The root-sign in this tree has all the syntactic features of a sentence: the value of its SYNSEM.LOCAL.CAT (shown in isolation in (282) below) is of type *verb*, and all valence requirements satisfied.

(282)
$$\begin{bmatrix} \text{SYNSEM.LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{AUX} & - \\ \text{INV} & - \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{SUBJ} & \langle\rangle \\ \text{SPEC} & \langle\rangle \\ \text{COMPS} & \langle\rangle \\ \text{SPR} & \langle\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

It is also semantically like a sentence, in that its top-EP (with the handle $\boxed{2}$) is of type *message* or, more precisely, a *prpstn*. This EP is contributed by the fragment-rule, via the feature C-CONT. This is also how the *unknown*-constraint is introduced into the representation. The connection of this constraint to the semantics of the phrase is made via co-indexation of the argument-slot of *unknown* with the INDEX of the argument phrase (in 7.3 this is $\boxed{5}$). As the reader is invited to verify, the CONT of the fragment-sign is very similar to that shown above in (281-a), which as we said is the FS-representation of the MRS (281-a); this in turn is the MRS-representation of an $LL$-formula.

We now explain how this construction-type is specified in detail. We have seen in the previous section (Figure 7.2) how the semantics of headed phrases is composed out of that of its constituents and possibly a contribution of the phrase, so this part does not have to be stated again specifically for fragments. We also assume that the *generalised head feature principle* (GHFP) from (Ginzburg & Sag 2001) is part of our grammar. We discussed this principle in Chapter 5; it is shown here again in (283). The GHFP states that a mother's SYNSEM will by default be structure-shared with that of the daughter.

(283) *hd-ph:*
$$\begin{bmatrix} \text{SYNSEM} & /\boxed{1} \end{bmatrix} \rightarrow \ldots \mathbf{H}\begin{bmatrix} \text{SYNSEM} & /\boxed{1} \end{bmatrix}$$

So what we have to state as the specific constraints on our type *nm-np-frag* are i) the relations in C-CONT, ii) the way the INDEX of the head daughter is connected to the C-CONT, and iii) the other exceptions to *ghfp* as far as they are not covered by i) and ii). Factors i)-iii) combined are shown in Figure 7.4.

The value for SYNSEM.LOCAL must be specified on the types for the fragments, since it will always be different from that of the head daughter—raising different XPs to sentences after all is the whole point of the rule, and so the default of the GHFP to copy these specifications must be overridden. The value

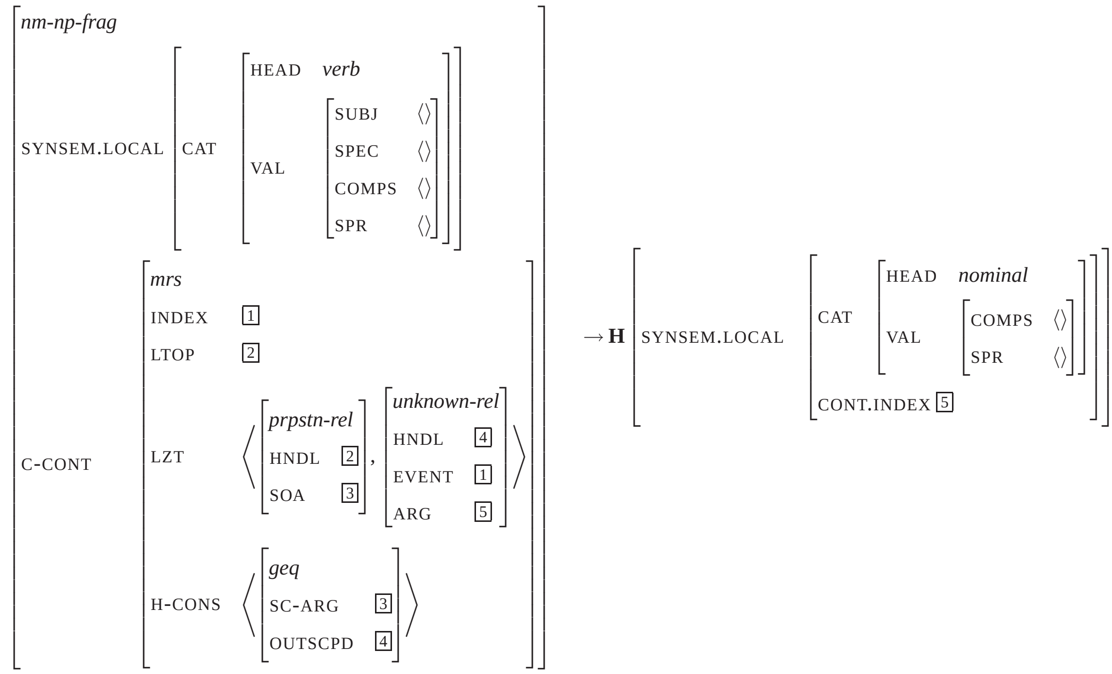
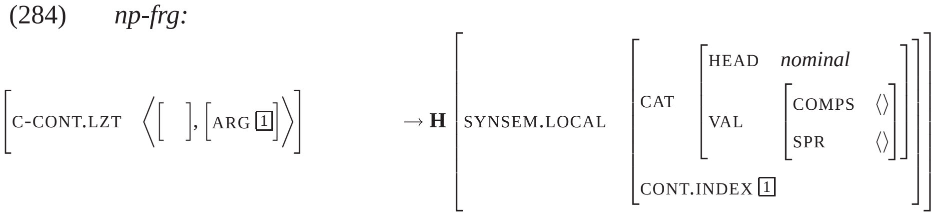



Figure 7.4: The type *nm-np-frag*

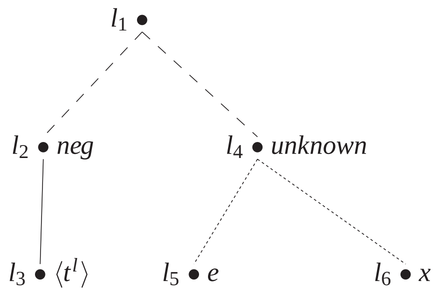
of SYNSEM.LOCAL.CAT will be the same for all types of fragments, and so we can factor this element out to a more general type than for example *np-fragment*. There are more elements of Figure 7.4 that are common to all types of fragments; in fact, the only parts of the constraint from that figure that are specific to NP-fragments are the co-indexation of the INDEX of the head (the NP) with the ARG of the *unknown*-rel, and of course the restriction that the phrase be an NP. So the constraint unique to NP-fragments is simply that shown in (284).

(284) *np-frg:*




We will organise the remaining common constraints on fragments along three dimensions, because they can vary independently. The first variation is whether the fragments are modified by adverbs or not. The example we have seen so far was of a non-modified fragment, and so we show here how a modified NP-fragment differs from that. As a reminder of the representations we want to build for those modified fragments, here first is an excerpt of the constraint-tree for the fragment “not Sandy” (omitting the representation of the proper name).

(285)




So what we want here is to have the adverb-relation as a sister to the *unknown*, for the reasons discussed above in Section 6.4.2 (basically, we want the adverb-relation to be able to scope in between a chain of scope-bearing elements that resolve *unknown*). (286) shows the constraint on this kind of fragment, insofar as it is different from 7.4. It is shown here in a third kind of notation, neither as tree nor as implication/rule but as one sign with additional fields for the daughters (of which only that for the non-head-daughter(s) is relevant here).

(286)

$$\begin{bmatrix} \textit{m-frg} & \\ \text{C-CONT} & \begin{bmatrix} \textit{mrs} & \\ \text{LZT} & \left\langle \begin{bmatrix}\textit{prpstn\_rel} & \\ \text{HNDL} & \boxed{2}\\ \text{SOA} & \boxed{3}\end{bmatrix}, \begin{bmatrix}\textit{unknown\_rel} & \\ \text{HNDL} & \boxed{4}\end{bmatrix}\right\rangle \\ \text{H-CONS} & \left\langle \begin{bmatrix}\textit{geq} & \\ \text{SC-ARG} & \boxed{3}\\ \text{OUTSCPD} & \boxed{4}\end{bmatrix}, \begin{bmatrix}\textit{qeq} & \\ \text{SC-ARG} & \boxed{3}\\ \text{OUTSCPD} & \boxed{5}\end{bmatrix}\right\rangle \end{bmatrix} \\ \text{NON-HEAD-DTRS} & \left\langle \begin{bmatrix} \text{SYNSEM} & \begin{bmatrix}\textit{scopal\_vp\_adv} & \\ \text{LOCAL.CONT.TOP} & \boxed{5}\end{bmatrix}\end{bmatrix}\right\rangle \end{bmatrix}$$

This was the first dimension: whether the fragment is modified or not; it determines what goes on C-CONT.H-CONS. The next element that can vary independently from others is the message type of the fragment. So far we have only looked at declarative fragments, but fragmental questions or requests will only differ in the type of this one relation, and so we can again factor this out and let the rules inherit the specification. To give an example, (287) shows the type *int(errogative)-frag(ment)*.

(287)

$$\begin{bmatrix} \textit{int-frag} \\ \text{C-CONT.LZT} \ \left\langle \begin{bmatrix}\textit{int}\end{bmatrix}, \ldots \right\rangle \end{bmatrix}$$

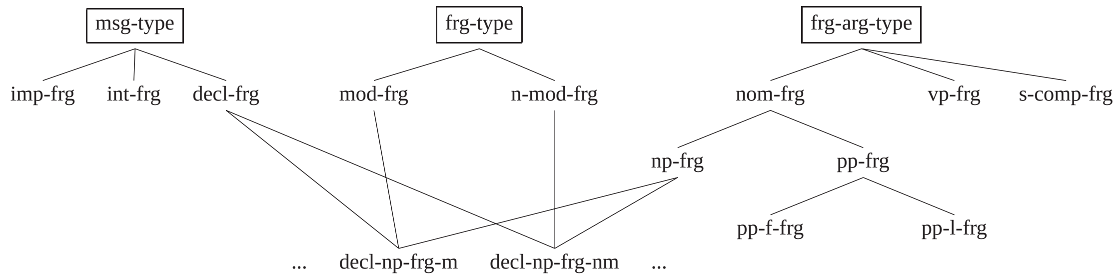



Figure 7.5: An extract of the construction hierarchy for fragments

The rules in this dimension must also make sure that *wh*-phrases can only be *int-frag*s, but since the specifics of how this can be achieved are dependent on how these are analysed we postpone showing the rest of the constraint for interrogative fragments until we discuss the implementation of our rules in the wide-coverage grammar.

The last dimension organises the difference due to the different argument-phrases. We will deal with this in detail in the next section, but before we do this we give here in Figure 7.5 as a summary of this section an excerpt of the inheritance hierarchy, which shows how the type *nm-np-frag* that we used in Figure 7.3 is assembled from *nm-frag* (because it is a fragment that is not modified by an adverb), *decl-frag* (because it is uttered with ‘declarative-intonation), and *np-frag* (because that is the category of the fragment-phrase).

### 7.3.2 Other fragment-types

In this section we will just list the constraints in the dimension *frg-arg-type*; we give them together with examples of the MRSs they build, and so they should be self-explanatory. Note that all these constraints only make assumptions about the head daughter; the types in the dimension ‘modified’ deal with possible non-head-daughters (the adverbs). Moreover, we focus here on the ‘main’-types of fragments, i.e. nominal-fragments and VP-fragments, leaving ADV- and ADJ-fragments, and also conjoined fragments for future work. (For the exact types used in the implementation turn to Appendix A; the signs here only illustrate the ideas.)

(288) The general type constraint for PP-fragments:

$$\begin{bmatrix} \textit{pp-frag} & \\ \text{C-CONT} & \begin{bmatrix} \text{LZT} & \left\langle \left[\ldots\right], \begin{bmatrix} \textit{unknown} & \\ \text{HANDLE} & \boxed{1} \end{bmatrix} \right\rangle \\ \text{H-CONS} & \left\langle \left[\ldots\right], \begin{bmatrix} \textit{geq} & \\ \text{SC-ARG} & \boxed{1} \\ \text{OUTSCPD} & \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix} \\ \text{HEAD-DTR.SYNSEM.LOCAL} & \begin{bmatrix} \text{CONT.TOP} & \boxed{2} \\ \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{prep} \\ \text{VAL} & \begin{bmatrix} \text{COMPS} & \langle\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

(289) a. The type constraint on lexical-PP-fragments:

$$\begin{bmatrix} \textit{pp-l-frag} & \\ \text{HEAD-DTR.SYNSEM.LOCAL.KEYS.KEY} & \textit{independent\_rel} \end{bmatrix}$$

b. "on Sandy" (lexical preposition reading)

$$\begin{array}{llll} \langle & h, e, & \{ & h_0 : prpstn(h_1), \\ & & & h_2 : unknown\_rel(e), \\ & & & h_3 : \_on\_rel(i, i', x), \\ & & & h_6 : def\_np\_rel(x, h_8, h_9),\ h_{10} : named\_rel(x, \text{“Sandy”}) \quad \}, \\ & & \{ & h_1 =_q h_2,\ h_8 =_q h_{10},\ h_2 =_q h_3 \quad \} \quad \rangle \end{array}$$

The feature KEYS.KEY in (289-a) allows one to select propositions; here, '*independent*' or lexical-prepositions are selected. In (290-a) a functional preposition is selected; note that the identification between the argument of *unknown* and that of the preposition has to be made explicitly.

(290) a. The type constraint on functional-PP-fragments:

$$\begin{bmatrix} \textit{pp-f-frag} & \\ \text{C-CONT.LZT} & \left\langle \left[\ldots\right], \begin{bmatrix} \textit{unknown} & \\ \text{ARG} & \boxed{1} \end{bmatrix} \right\rangle \\ \text{HEAD-DTR.SYNSEM.LOCAL.KEYS.KEY} & \begin{bmatrix} \textit{selected\_rel} & \\ \text{ARG3} & \boxed{1} \end{bmatrix} \end{bmatrix}$$

b. "on Sandy" (functional preposition reading)

$$\begin{array}{llll} \langle & h, e, & \{ & h_0 : prpstn(h_1), \\ & & & h_2 : unknown\_rel(e), \\ & & & h_3 : \_on\_rel\_s(i, i', x), \\ & & & h_6 : def\_np\_rel(x, h_8, h_9),\ h_{10} : named\_rel(x, \text{“Sandy”}) \quad \}, \\ & & \{ & h_1 =_q h_2,\ h_8 =_q h_{10},\ h_2 =_q h_3 \quad \} \quad \rangle \end{array}$$

(291) a. The type constraint on VP[*inf*]-fragments:

$$\begin{bmatrix} \textit{vp-inf-frag} & \\ \text{C-CONT.LZT} & \left\langle [\ldots], \begin{bmatrix} \textit{unknown} & \\ \text{ARG} & \boxed{1} \end{bmatrix} \right\rangle \\ \text{HEAD-DTR.SYNSEM.LOCAL} & \begin{bmatrix} \text{CAT.HEAD} & \begin{bmatrix} \textit{comp} & \\ \text{VFORM} & \textit{inf} \end{bmatrix} \\ \text{CONT.TOP} & \boxed{1} \end{bmatrix} \end{bmatrix}$$

b. “to leave”

$$\begin{array}{llll} \langle & h, e, & \{ & h_0 : prpstn(h_1), \\ & & & h_2 : unknown\_rel(e, h_3), \\ & & & h_3 : hypo\_rel(h_4), \\ & & & h_5 : leave\_rel(e', x) \quad \}, \\ & & \{ & h_1 =_q h_2 \; h_4 =_q h_5 \quad \} \quad \rangle \end{array}$$

(292) a. The type constraint for VP[*bse*]-fragments:

$$\begin{bmatrix} \textit{vp-bse-frag} & \\ \text{C-CONT} & \begin{bmatrix} \text{LZT} & \left\langle [\ldots], \begin{bmatrix} \textit{unknown} & \\ \text{ARG} & \boxed{1} \end{bmatrix} \right\rangle \\ \text{H-CONS} & \left\langle [\ldots], \begin{bmatrix} \textit{geq} & \\ \text{SC-ARG} & \boxed{1} \\ \text{OUTSCPD} & \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix} \\ \text{HEAD-DTR.SYNSEM.LOCAL} & \begin{bmatrix} \textit{bse\_verb} & \\ \text{CONT.TOP} & \boxed{2} \\ \text{CAT} & \begin{bmatrix} \text{HEAD.VFORM} & \textit{bse} \\ \text{VAL.COMPS} & \langle\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

b. “leave”

$$\begin{array}{llll} \langle & h, e, & \{ & h_0 : prpstn(h_1), \\ & & & h_2 : unknown\_rel(e, h_3), \\ & & & h_4 : leave\_rel(e', x) \quad \}, \\ & & \{ & h_1 =_q h_2 \; h_3 =_q h_4 \quad \} \quad \rangle \end{array}$$

(293) a. The type constraint on S[*comp*]-fragments:

$$\begin{bmatrix} \textit{vp-inf-frag} & \\ \text{C-CONT.LZT} & \left\langle [\ldots], \begin{bmatrix} \textit{unknown} & \\ \text{ARG} & \boxed{1} \end{bmatrix} \right\rangle \\ \text{HEAD-DTR.SYNSEM.LOCAL} & \begin{bmatrix} \text{CAT.HEAD} & \begin{bmatrix} \textit{comp} & \\ \text{VFORM} & \textit{fin} \end{bmatrix} \\ \text{CONT.TOP} & \boxed{1} \end{bmatrix} \end{bmatrix}$$

b. “that Kim walks”

$$\langle \quad h, e, \quad \{ \quad h_0 : prpstn(h_1), \; h_2 : unknown\_rel(e, h_3), \; h_3 : prpstn\_rel(h_4), \; h_5 : walk\_rel(e', x), \; h_6 : def\_np\_rel(x, h_8, h_9), \; h_{10} : named\_rel(x, \text{“Sandy”}) \quad \}, \quad \{ \quad h_1 =_q h_2 \; h_4 =_q h_5 \; h_8 =_q h_{10} \quad \} \quad \rangle$$

An issue which we leave unresolved in this thesis is what the best analysis is of examples like “I think with a fork”. At first glance, one might assume that what we have here simply is a verb (“think”) taking a sentence as complement, a sentence which happens to be fragmental. However, the detachability of expressions like “I think” could also suggest an analysis of them as parentheses, modifying sentences (in this case a fragmental one). Lastly, an argument could also be made for allowing certain verbs that take sentential complements to alternatively take NPs or PPs rather than Ss, i.e. for ‘lexicalising’ the reduction of the message. All alternatives pose problems. A fragment-embedding analysis has to explain why not in all places where embedded sentences are allowed also fragments are allowed (e.g. “?*Peter is happy that Sandy.”), and moreover why only propositional fragments seem to be allowed (“??Peter wonders if Sandy.”). A sentence/phrase-ambiguity analysis has to explain examples like “Peter thinks not Sandy.”, which show that arguments don’t have to be constituents. A parenthesis analysis would ideally have to explain why certain expressions can form parenthesis and others don’t. We stay on the fence here and simply observe that there are these problems. In the following section, we tentatively offer a rule that allows the embedding of fragments, but we keep in mind the problems mentioned here.

## 7.4 The Implemented Grammar

We have integrated the syntactic analysis of fragments that we have presented in the previous section into an existing wide-coverage grammar of English, the English Resource Grammar (ERG, (Copestake & Flickinger 2000)). This grammar is supported by a grammar development, parsing and generation platform known as the LKB (standing for Linguistic Knowledge Base; see (Copestake 2002)).[11] In this section we will now describe the technical side of the implementation of the grammar, and we will also provide data that allow the impact of these changes to the grammar to be assessed

Before we come to the rules, however, we have to answer a possible objection, a more technically oriented one. Why *extend* a grammar and not devise a separate grammar just for fragments, for example designers of grammar-based speech recognition systems might ask. The answer is that we would gain very little with such a move. Our fragment rules build on the existing rules for phrases like NPs, PPs and VPs, and as (294-b) illustrates, these phrases can indeed be as complex in fragments as they can be in full sentences.

[11] The homepages for the grammar and the development system are `http://lingo.stanford.edu` and `http://www-csli.stanford.edu/~aac/lkb.html`, respectively.

(294) a. When shall we meet again?
b. On the first Wednesday of that week in August when the manager returns, whichever week that is.

Hence the only rules that can be left out of a fragment-grammar are those combining phrases into sentences, and apart from the rules taking phrases to fragmental sentences all other rules in the fragment-grammar would have to be taken from the full-sentence grammar.

Before we can finally go into the details, a few words about the ERG are needed. First of all, it is important to stress again that this grammar is not a toy-grammar; it has been developed with ‘real-world’ applications in mind and for example can currently produce semantic representations “for about 83 per cent of the utterances in a corpus of transcriptions of some 10,000 utterances [from the Verbmobil-Domain described above, D.S.], which vary in length from one word to more than thirty words” (Copestake & Flickinger 2000). Its development required an effort of a total of eleven person-years, and the grammar currently contains over 15,000 lines of code (excluding the lexicon) (Copestake & Flickinger 2000). All this is relevant here, because only within such a comprehensive grammar is it possible to reliably identify side-effects of introducing new constructions and to evaluate the usefulness of such additions in terms of coverage and robustness. This evaluation in turn was only possible through the use of a grammar-profiling tool called `[incr tsdb()]` (Oepen & Flickinger 1998), which records detailed data about parsing performance over a given test suite of sentences, allowing the grammar developer to assess the effects of the changes he has made to the grammar. The results of these tests will be detailed below in section 7.4.2, but first we now turn to the rules themselves.

### 7.4.1 The Rules

Overall, the implemented rules are a fairly straightforward formalisation in the notation used in the LKB of the constraints we have shown in the previous section. We will not go through all the rules in detail here, and only point out where the implementation departs from the analyses from the previous section.[12]

There are two major points of deviation. First, for technical reasons, the ERG does not make use of defaults (although the LKB supports their use), and so we cannot use the generalised head-feature principle from the previous section, which there simplified the presentation of the constructions. For this reason, it was not possible to use in the implementation the feature HEAD-DTR (on which the monotonic head-feature principle will work and from which it would take the value of HEAD for the construction, even when the fragment-phrase is an NP and the resulting fragment-construction should be of sentential type); rather, we had to introduce special features FRAGH-DTR (for *fragment-head-daughter*) and FRAGNH-

[12] The rules are printed in the appendix, with extensive comments. Also, a `diff`-file with which a copy of the ERG can be extended with the fragment rules is available on request from the author.

DTR (*fragment-non-head-daughter*), from which the head of the fragment-construction is available. To give an example, in a fragment like “probably Sandy”, the NP is the fragment-head-daughter and the ADV the non-head daughter.

The other deviation from the analysis of the previous section is that rather than analysing adverbs as sisters of *unknown*, as described above in (285), here we have them outscope that relation. As explained in Section 6.4.2, this means that the ULFs coming out of the grammar do not describe all possible readings. However, this analysis is technically simpler, because it does not require changes to be made to the analysis of adverbs in the grammar, and also, as described in that section, the readings that are lost are relatively marginal.

The organisation of the rules into dimensions is as shown above in Figure 7.5; however, we here have a common supertype **frg** above the types **msg-type**, **frg-type** and **frg-arg-type**, which collects all constraints that are common to all different fragment-constructions. The specification of this type is shown in the following.

```
frg := non_headed_phrase & rule &
  [
    SYNSEM synsem_min &
           [ LOCAL [ CAT s_cat_fin_unspec &
                         [ HEAD verb & [ AUX -,
                                         INV -,
                                         MOD < > ],
                           VAL [ SUBJ <>,
                                 SPR  <>,
                                 COMPS <>,
                                 SPEC <>]
                         ],
                     CONJ cnil,
                     KEYS.MESSAGE <! #msg !> ],
             NON-LOCAL non-local_none
            ],
    C-CONT [ TOP #ctop,
             INDEX #event,
             E-INDEX #event,
             LISZT [ LIST < message & #msg & [ HANDEL #ctop ],
                            unknown_rel & [ EVENT #event ]
                          > ],
             H-CONS [ LIST < geq, ... > ]
           ],
    FRAGH-DTR sign & [SYNSEM.NON-LOCAL.SLASH 0-dlist],
    FRAGNH-DTR sign
  ].
```

To give a feel for the rules as they are implemented, and also to give an idea of how the ERG in general works, let us dissect this rule in the following. We begin with the first line:

```
frg := non_headed_phrase & rule &
```

This says that the type **frg** inherits constraints from **non_headed_phrase** and from **rule**. The former type specifies that the phrase built with this rule (or, equivalently, the sign satisfying these constraints) does not have a head-daughter. We have to do this to block the head-feature principle from applying, for the reasons given above. The other parent-type, **rule**, introduces a list-valued feature for arguments.

The next few lines give the specification of the syntactic face of fragment-signs:

```
[
  SYNSEM synsem_min &
         [ LOCAL [ CAT s_cat_fin_unspec &
                       [ HEAD verb & [ AUX -,
                                       INV -,
                                       MOD < > ],
                         VAL [ SUBJ <>,
                               SPR  <>,
                               COMPS <>,
                               SPEC <>]
                       ],
                   CONJ cnil,
                   KEYS.MESSAGE <! #msg !> ],
          NON-LOCAL non-local_none
         ],
```

In order, these specifications say that fragment signs belong to the category **s_cat_fin_unspec**, i.e. behave syntactically like sentences with a finite verb. The additional specifications constrain this further, by postulating that the 'verb' is not an auxiliary (AUX -) and not inverted (INV -), can't modify anything (MOD is an empty list, ⟨⟩), and has all its valency requirements satisfied (note that this is not redundant, since **s_cat_fin_unspec** does not specify all of this). The specification of `CONJ` as `cnil` blocks these signs from entering into coordination, whereas the value of `KEYS.MESSAGE` makes the message-type of the 'sentence' something that is selectable for other rules. Finally, the value of NON-LOCAL is specified to be empty, which prevents this sign from unifying with anything that wants to fill a gap.

This leaves the semantic contribution of the construction type 'fragment', which is as shown below:

```
C-CONT [ TOP #ctop,
         INDEX #event,
         E-INDEX #event,
         LISZT [ LIST < message & #msg & [ HANDEL #ctop ],
                        unknown_rel & [ EVENT #event ]
                      > ],
         H-CONS [ LIST < geq, ... > ]
       ],
```

This specifies that the top-handle of the construction will be the handle of the message-relation of the fragment (be that *prpstn*, *decl* or *imp*), and that the main event is contributed by the *unknown*-relation. Moreover, this specification says that at least one scope-constraint is contributed by this construction.

Finally, it is this type where the features FRAGH-DTR and FRAGNH-DTR are introduced. We can also already on the mother-type of all fragment-constructions specify that the fragment-phrases are not allowed to contain gaps; this rules out ill-formed fragments like "the father of", which otherwise would be allowed (as an NP containing a gap where the complement of the preposition is).

```
    FRAGH-DTR sign & [SYNSEM.NON-LOCAL.SLASH 0-dlist],
    FRAGNH-DTR sign
  ].
```

The types for the three dimensions now are direct daughters of this general type:

```
frg_md :< frg.
frg_arg :< frg.
frg_mood :< frg.
```

The type **frg_md** (for *fragment-modification*) has two daughters, one for fragments that are modified by an adverb and one for those that aren't. Modification works basically as described above in (286) in the previous section, with some feature names being different. **frg_arg** is a more diverse hierarchy, with sub-types for altogether six different arguments categories (NPs, lexical and functional PPs, VP[*inf*], VP[*bse*] and S[*comp*]; we do not provide analyses for the other types mentioned above); again the contribution of the subtypes is as shown in the previous section, in Section 7.3.2. The type **frg_mood** then finally has only two daughters, one for interrogative and one for declarative fragments. As the reader will have noticed, this leaves out imperative fragments; moreover, we make the restrictive assumption here that only *wh*-phrases are interrogative fragments. Both restrictions are made for the same reason: since obviously no information about verb-inversion for example is available, the only remaining information source for distinguishing for example "Peter." from "Peter?" or "Peter!" would be interpunction (or, in the case of a speech recognition grammar, intonation). The parsing system we use, however, does not (at the time of writing) exploit this information, and so to avoid multiplying parses (by giving three analyses for e.g. "Peter") we have decided for this restrictive approach. This, however, is not a conceptual necessity and could be changed if needed.

All in all, this results in 18 different construction types as the nodes of this hierarchy, which through multiple inheritance can be specified very parsimoniously, as the example for non-modified interrogative NP-fragments shows:

```
np_nm_i_fragment := np_fragment & frag_int & fragment_nm.
```

### 7.4.2 Effects and Side Effects

So far, in this chapter and the previous one, what we have done could be classified as theoretical linguistics: we have given a theory of which kinds of fragment-readings and fragment-signs are *possible*, not of which are *frequent*. But when we move to an actual implementation of an analysis of a phenomenon, we have to balance the cost of extensions of the grammar in terms of increase in ambiguities against the achieved increase in real coverage—and possibly leave out some rules if they turn out to be too costly. This is what we discuss here for the ERG with our modifications for fragments (in the following, $\text{ERG}_{\text{FRAG}}$). We provide three sets of evaluations, all done with `[incr tsdb()]` (Oepen & Flickinger 1998): first, the coverage of the $\text{ERG}_{\text{FRAG}}$ with respect to the phenomenon it is supposed to analyse; second the (side-)effects of the modifications on analyses of full sentences; and lastly the gain in coverage of real-world utterances.

Note that these evaluations, especially those with larger data sets, are only intended to show tendencies or trends in behaviour, since the main indicator used here is simply the number of parses the grammar offers for a given string. Ultimately, a grammar should be evaluated by checking whether the contextually appropriate parse for a particular item is in the set of those given by the grammar. We have done this only for a small set of data, since it is very time consuming; and so our evaluation overall should only be taken as indicating a tendency.

**Coverage of Phenomenon**

To help with the construction of the grammar, early on we devised a very small test-suite of fragment examples for use with `[incr tsdb()]` (printed in Appendix A.2); this allowed us to judge the effects of changes to the grammar with respect to analyses of the phenomenon. For the evaluation of the finished grammar, however, this test-suite of course isn't that useful: in a sense, it constitutes our 'training data', and so for types of fragments we decided to provide with a representation,[13] our grammar reaches the ideal coverage of 100% of the test-items. The results are summarised in Table 7.1.

What is interesting here, however, is the ambiguity the grammar exhibits, as indicated by the numbers in the fifth column of the table. For example, VP-fragments have on average more than two parses, whereas modified fragments have on average more than five parses. To get an idea of these parses, let's go through some examples. First, for the string "kick Sandy". This string gets two parses from the $\text{ERG}_{\text{FRAG}}$ (we forgo showing them here), only one of which is as a fragment. The other one, naturally, is as an imperative sentence. This makes clear that the classification of the items in the test-suite into types of fragments is not exclusive: being a fragment is not a syntactic property, and hence it is possible that there are other parses for a given string that could be a fragment. All we are interested in here is the ambiguity our grammar *adds*, and for this example we fared well (since we only added exactly one reading). A bit more surprising perhaps are some of the parses for "to kick Sandy". (295) first shows the parse that corresponds to the analysis we have given for this string so far: a VP[*inf*] that is projected to S.

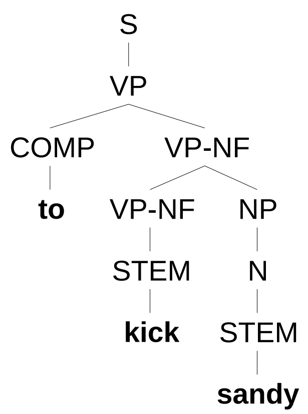
(295)

[13] We should stress again here that there was no substantial reason for not extending our approach to these types; we simply concentrated on the major types. Of course, the impact of providing analyses for the types we left-out would have to be judged again.

| **Phenomenon** | **total items** ♯ | **positive items** ♯ | **word string** ϕ | **lexical items** ϕ | **parser analyses** ϕ | **total results** ♯ | **overall coverage** % |
|---|---|---|---|---|---|---|---|
| NP_frag | 5 | 5 | 2.40 | 6.60 | 2.40 | 5 | 100.0 |
| PP_frag | 3 | 3 | 2.00 | 4.33 | 2.00 | 3 | 100.0 |
| VP_frag | 7 | 3 | 3.00 | 8.67 | 2.67 | 3 | 100.0 |
| ADV_frag | 3 | 3 | 1.00 | 1.67 | 0.00 | 0 | 0.0 |
| ADJ_frag | 2 | 2 | 1.00 | 2.00 | 0.00 | 0 | 0.0 |
| Int_frag | 8 | 8 | 1.37 | 4.50 | 1.50 | 8 | 100.0 |
| Slashed_frag | 6 | 0 | 0.00 | 0.00 | 0.00 | 0 | 0.0 |
| Modified_frag | 3 | 3 | 4.00 | 11.00 | 5.33 | 3 | 100.0 |
| Conj_frag | 5 | 5 | 3.00 | 11.60 | 3.33 | 3 | 60.0 |
| **Total** | **42** | **32** | **2.19** | **6.50** | **2.56** | **25** | **78.1** |

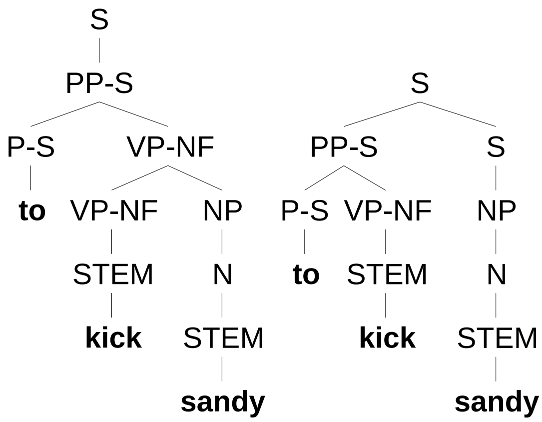
(generated by [incr tsdb()] at 27-nov-2002 (15:38 h))

Table 7.1: Coverage of Fragment-Test-Suite



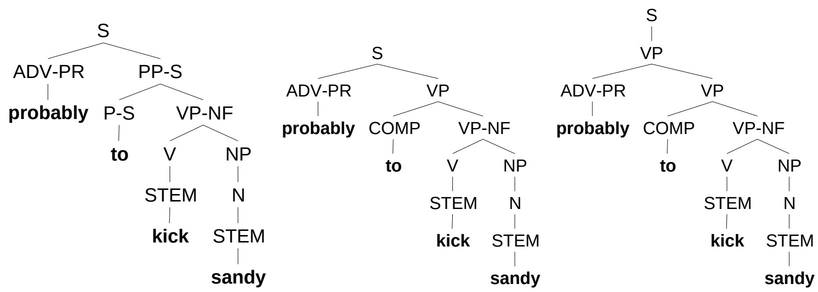

The next two parses, shown in (296), are less obvious. In the first, "to kick Sandy" is interpreted as a PP meaning "in order to kick Sandy"—a parse we didn't think about before, but that seems correct. The second parse interprets the string as consisting of such a PP, modifying a sentence. A full sentence instance of this construction is "to sing, Peter came here"; with a fragment as sentence from which the PP is extraposed this surely sounds rather unplausible, but there is no principled reason for ruling this out. In any case, the discussion of this example nicely demonstrates the advantages of using a wide-scope grammar for implementation, since the interaction with such a construction would otherwise not have been noticed.

(296)




Limitations of space preclude us from discussing many more examples, and so we only show here one more, to explain why modified fragments have so many more parses. The fragment “probably to kick Sandy.” receives six analyses, three of which are like the parses shown before for “to kick Sandy”, with the sentence modified by the sentence-adverb reading of “probably”. The remaining three are shown in (297). Two of these are the result of application of the ‘S → ADV XP’ rule; in the remaining one the ADV is part of the VP, and the resulting VP is projected to S.[14]

(297)




We now come to the side-effects of having rules for fragments on the full-sentences grammar.

### Side-Effects

To test for possible adverse effects on the analyses of full-sentences, we also ran batch-parses of a test-suite of full sentences, namely the CSLI-test-suite which is distributed with `[incr tsdb()]`. This test suite consists of 1348 sentences, of which 961 are marked as syntactically well-formed and 387 as ill-

[14] This makes us expect that the following structure is licensed, which indeed it is: ‘[$_S$ probably [$_{S[frag]}$ probably [$_{VP}$ to [$_{VP}$ probably kick sandy]]]]’. Again, there is no principled syntactic reason to exclude such a parse.

formed. Table 7.2 shows a comparison of the original ERG with four versions of the FRAG-ERG,[15] with respect to the average number of parses per sentence, the percentage of sentences that are marked as well-formed which do get at least one parse (coverage), and the percentage of ill-formed sentences that nevertheless get a parse (overgeneration).

| *Version of Grammar* | *Average # parses* | *% Coverage* | *% Overgeneration* |
|---|---|---|---|
| LinGO ERG, 20/11/02 | 2.86 | 81.4 | 33.9 |
| ERG+frag - - | 3.23 | 82.3 | 36.4 |
| ERG+frag +- | 3.33 | 82.3 | 36.4 |
| ERG+frag -+ | 3.59 | 82.3 | 37.7 |
| ERG+frag ++ | 3.69 | 82.5 | 37.7 |

Table 7.2: Competence comparison of the original ERG with four versions of the fragment-ERG

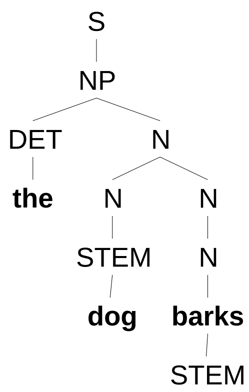
As these data show, the rules for fragments introduce new ambiguity into the full-sentence part of the grammar; only slightly so for the most restricted version of ERG-FRAG (which only adds rules for NP- and PP-fragments and does not allow embedding), up to on average almost one more parse for the version with VP-rules and embedding. We also note a slight increase in coverage (i.e., sentences known to be well-formed that do get parsed) and a more pronounced increase in overgeneration (ill-formed sentences that nevertheless do receive a parse).

But what does this tell us? One has to be careful when interpreting these numbers. It seems a reasonable conclusion to say that the addition of fragment-rules does not lead to an explosion of readings that would render the grammar practically unusable. Without further analysis, however, this does not tell us anything about whether the additional readings (of what are meant to be full sentences) are erroneous or not. The problem of course again is that ‘fragmenthood’ is not a syntactic criterion, and so some strings that can be analysed as sentences can also be analysed as fragments (and hence an increase in ambiguity would be unavoidable); we have just discussed such an example in the previous section (“kick Sandy”). And by the same token, even the increase in coverage and the increase in ‘over-generation’ need not be due to unwanted side-effects, since there could be items that are now parsed quite legitimately as fragments.

Hence, a more detailed analysis is required. To begin with a simple example, (298) shows a parse for the string “the dog barks” which is added by our fragment rules:

[15] The two varying factors are whether VP-fragments are analysed or not and whether embedding of fragments is allowed or not; hence ‘ERG+frag -+’ denotes a version of the fragment grammar where VP-fragments are excluded, but embedding is allowed.

(298)



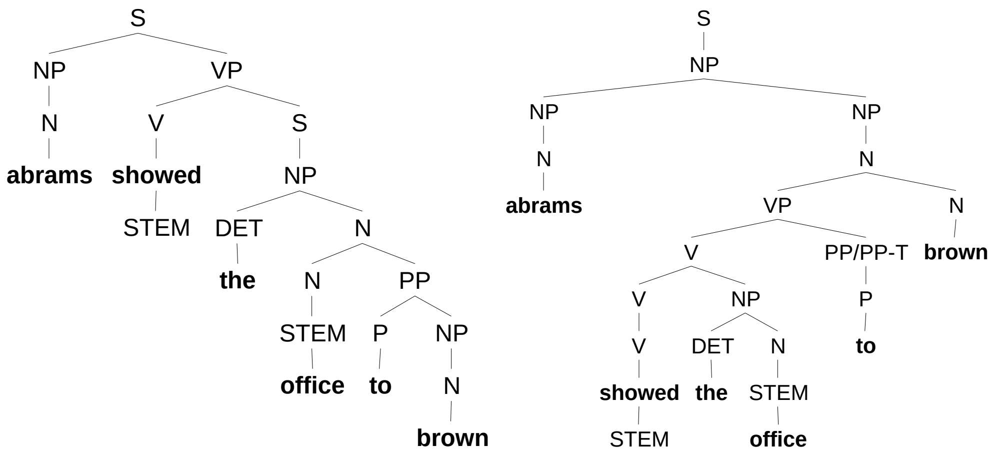

It turns out that the grammar lists ‘bark’ as a noun as well as a verb, and so ‘dog barks’ forms a compound noun which is then part of an NP-fragment. So in this case, the additional parse seems to be legitimate.

A more involved example—which occurs in the test-suite used to get the data quoted above—is the sentence “Abrams showed the office to Brown.” The unmodified ERG gives four readings for this, to which $\text{ERG}_{\text{FRAG}}$ adds six further readings. Five of these additional readings are variations of the one shown on the left in (299); in all of them ‘show’ takes a sentence as complement (as in ‘Peter showed that the Goldbach conjecture is true.’), a sentence which however is a fragment. In the parse shown on the left in (299), this fragment is ‘the office to Brown’ parsed as an NP; there are also parses where ‘the office’ is the direct object of show and ‘to Brown’ as a PP-fragment is the sentential complement. These readings are certainly unwanted, and point to a problem of the version of the fragment grammar that allows embedding, namely that it is too unrestricted. Ideally, fragments should be marked as a special kind of sentence only certain kinds of verbs select for (e.g. ‘think, say, believe’, etc.). To do so, however, would require further analysis of where embedding of fragments is possible, and also would require changes to the sentence grammar, which in this first step we did not want to change. So we conclude here that allowing embedded fragments systematically overgenerates.

| *Version of Grammar* | *tasks* | *time* | *space* | *% red. tasks* | *% red. time* | *%red. space* |
|---|---|---|---|---|---|---|
| LinGO ERG, 20/11/02 | 722 | 0.13 | 4485 | n/a | n/a | n/a |
| ERG+frag - - | 753 | 0.15 | 4761 | -4.3 | -9.2 | -6.2 |
| ERG+frag +- | 761 | 0.14 | 4943 | -5.4 | -4.6 | -10.2 |
| ERG+frag -+ | 770 | 0.14 | 4997 | -6.6 | -6.8 | -11.4 |
| ERG+frag ++ | 781 | 0.15 | 5157 | -8.2 | -13.8 | -15.0 |

Table 7.3: Performance comparison of the original ERG with four versions of the fragment-ERG

(299)




The other parse shown in (299) is perhaps more surprising. In it, the constituent ‘showed the office to Brown’ is parsed as a (rather complex) last name (!), so that the whole string becomes a proper noun NP, which is then projected to S. This most likely is a bug in the ERG, and so shows a case where the analysis of fragments can inform the development of the full-sentence rules.

Before we move on, we have to take a short look at the costs of the additions in terms of performance as well. It is always possible that added rules lead to a significant increase in the processing required, even if they do not add many readings. Table 7.3 shows a comparison of the grammars with respect to several performance measures. It is not necessary to explain those in detail; suffice it to say that there is a noticeable performance cost of using the fragment grammar rules (between four and eight per cent processing time, and six and 15 per cent memory required). This is not surprising, since for example each NP will now have a parse as a sentence, which in full sentences however cannot combine with other elements. Nevertheless, we think the increase in performance cost lies in the range of what keeps the grammar practically useable.

The moral of the story is that the addition of these rules does have some costs in terms of increase in ambiguity and processing time, which—although theoretically defendable—better be outweighed by a gain in coverage of input that can be expected in practical applications. This is what we tested with the next set of data.

### Real-World Effects

Having shown that our grammar rules for fragments do not have lethal side-effects, we should also show that they do indeed have useful effects at all, i.e. that they can deal with 'real-world' data as well. To this end, we analysed some naturally occuring dialogue data, from the Verbmobil corpus of scheduling dialogues we also used in Chapter 2.[16] In particular, we looked at the data from the Verbmobil-CD 6 (in the following, VM6), which consists of 4037 items. Of these, 3709 are marked as well-formed.[17] 1492 (= 37%) items altogether do not receive a single parse from the ERG,[18] and 1239 of these non-parsed items are marked as well-formed (= 33% of all grammatical items). This being a dialogue corpus, it is to be expected that at least some of these items that didn't receive a parse are fragments. Indeed, running $\text{ERG}_{\text{FRAG}}$ over these examples, we could provide at least one parse for 407 (= 32 %) of these items. (These numbers are summarised in Table 7.4.)

| VM | | |
|---|---|---|
| Items corpus VM6: | 4037 | |
| Well-formed items: | 3709 | |
| No parse (all): | 1492 | (= 37% off all). |
| No parse (well-formed): | 1239 | (= 33.4% of WF) |
| Of which parsed by $\text{ERG}_{\text{FRAG}}$: | 407 | (33% of above, 10% of whole) |

Table 7.4: Summary VM6

Again, one must use caution when interpreting these numbers. Just as we had an increase in overgeneration in the previous set of data, it could very well be that these new readings are spurious and not what is wanted. To give a feel for the items that are only parsed by the $\text{ERG}_{\text{FRAG}}$, (300) lists a few randomly selected items. '$\times$' in this list marks items where a parse as fragment is presumably not wanted, and hence where our grammar overgenerates; those that are legitimate fragments are marked with '$\surd$'. Note that we do not mark with this whether a parse as fragment is contextually appropriate.[19]

[16] The data is taken from the resources provided by the redwoods initiative (see description in (Oepen, Flickinger, Toutanova & Manning 2002)). This initiative is building an HPSG-treebank for these dialogue utterances. The first evaluation presented here only uses the test-items in this set, while the second makes use of the fact that for each test item there is information about disambiguation present as well.

[17] The marking for ill-formedness only regards remnants of dysfluencies (e.g., "Let's meet at, erm, on Sun, no, Monday") and not items that could be fragments, and so unlike the marking for the test suite used in the previous section we can safely look only at the well-formed test items.

[18] At the time of writing, the redwoods data is based on a slightly outdated version of the ERG (from June 20th 2002) compared to what we used in the previous sections; to maintain comparability we have added our fragment rules to this version of the ERG.

[19] For example, many of the proper names that are parsed as NP-fragments by our grammar should in their context best receive a different analysis, namely as vocatives, e.g. in "Patty, we should meet next week."

(300)

| | |
|---|---|
| in seminars all day. | √ |
| so we can talk some more about the analysis. | × |
| that, | √ |
| between the fourth and sixth of August. | √ |
| Patty | √ |
| on the, third of February, two to four. | √ |
| ten four good buddy, | × |
| Patty | √ |
| just to talk about some things, | √ |

Since we used the same data as part of the corpus study reported in Chapter 2, here we do actually have numbers about how many of these parses were contextually appropriate. We know from Chapter 2 that there were 369 fragments altogether in the VM6 data. 242 of these were in the set of 407 items only parsed by the $\text{ERG}_{\text{FRAG}}$ parses, i.e. 60% of the items only parsed by the $\text{ERG}_{\text{FRAG}}$ are genuine fragments and are provided with the correct semantic representation, and 65.5% of all fragments are correctly recognised by our grammar.

| VM | | |
|---|---|---|
| Items Corpus VM6: | 4037 | |
| Items only parsed by $\text{ERG}_{\text{FRAG}}$: | 407 | |
| Fragments: | 369 | |
| Of which parsed by $\text{ERG}_{\text{FRAG}}$: | 242 | (= 65.5%). |
| $\text{ERG}_{\text{FRAG}}$ + CONJ-frag: | 305 | (= 82.6%) |

Table 7.5: Summary coverage $\text{ERG}_{\text{FRAG}}$ on VM6

Note that these numbers do *not* say anything about syntactic overgeneration. Fragmenthood simply is not a syntactic property; the 165 items parsed by the $\text{ERG}_{\text{FRAG}}$ which are not fragments *in their respective contexts* might still have legitimate parses as such. However, they do say something about undergeneration: something that is a fragment in context should in any case get a parse, and at the moment we seem to be missing 127 (34%) of the fragments. Where are these items hiding? Among the 832 items that were not parsed even with our extensions to the grammar there were 95 fragments (11.4%); consequently, it seems that 32 fragments received an (erroneous) sentence parse.[20] A look at the items not parsed by either ERG or $\text{ERG}_{\text{FRAG}}$ shows an interesting point. (301) shows a random sample of this set.

[20]The REDWOODS treebank even provides us with further information, namely with preferences for particular parses. This data shows that for 234 items all parses offered by the ERG were rejected; it is likely that our 32 fragments are among them.

(301)

1 and I am free, on Tuesday, in the afternoon.
2 and, actually, I am also free, Wednesday and Thursday in the afternoon.
3 and, Thursday all day.
4 and again I am also still free, afternoons of Monday and Tuesday next week.
5 or would you like to meet, later in the afternoon.
6 great.
7 well, we, get down to the week of the twenty sixth,
8 well I am afraid I am busy, all afternoon that Thursday, so if you move all the way to the fourth of August, I am free in the afternoon there, or the morning of the fifth.
9 and on Monday.
10 alright, so why don't we meet, two o'clock, and why don't I come to your office.
12 or maybe Peter.
13 sounds good to me
14 and, some meetings the rest of the week.

Apart from illustrating that the full-sentence grammar still misses some constructions, this excerpt lays bare a systematic omission in the ERG$_{\text{FRAG}}$: fragments beginning with a conjunction, like items 3, 9 and 12 in (301). In fact, 63 items (7.5%) in that set were of this kind. This means that were we to add rules for this kind of fragment, our coverage would increase to 82.6% of all fragments; a rate that is comparable to what the ERG achieves on full sentences. (The numbers from this section are summarised in Table 7.5.)

We conclude from these three studies that our grammar offers potential to allow real-world applications to tackle the phenomenon of sentential fragments. However, we have also identified a worthwhile area for further work: results could potentially be dramatically improved by covering fragments which start with a conjunction. This however would involve changing the ERG's existing type constraints on constructions of full sentences, in order to avoid a massive proliferation of ambiguity for the parses of full sentences; this we leave to further work.

## 7.5 Summary

In this chapter we have presented our grammar for fragments. Unlike (Ginzburg & Sag 2001),[21] we have offered analyses not only for nominal fragments, but also for VP-fragments and for modified

[21] Who admittedly, as described in Chapter 5, try to do more in their grammar in that they also directly contextually resolve the fragments, a task which we postpone until the next Chapter. Moreover, they offer uniform analyses to other phenomena (like *in-situ*-questions) we do not deal with.

fragments. We have implemented our analysis and evaluated the implementation with a range of test-items, including real-world data. As a result, it seems that our rules offer a significant increase in coverage for such dialogue-data.

# Chapter 8

# A Coherence-Based Approach III: Resolution

It's time now to take stock of what we have done in the last two chapters. We have introduced a logical language for representing the compositional semantics of fragments (Chapter 6), and we have shown how formulae of this language can be built on the syntax/semantics interface (Chapter 7). The element of the story that is still missing is how we can resolve the underspecification in representations of the semantics of fragments, given information about the context in which the fragment was used. This is what SDRT (Asher 1993, Asher & Lascarides 2003) will give us. We will use this theory as our framework in which we make precise what we said in Chapter 2 about how the speech-act performed with a fragment determines its resolution.

In the first section of this chapter we give a very brief overview of this theory and some background concepts, then we introduce a new constraint on interpreting discourse which predicts the different behaviour of *resolution-via-identity*- and *resolution-via-inference*-fragments, and finally we then turn our attention to the details of the resolution of the types of fragments listed in our taxonomy.

## 8.1 Background: SDRT

In this section we give a brief overview of *Segmented Discourse Representation Theory* (SDRT, (Asher 1993, Asher & Lascarides 2003), the theory which we will use to explain how the underspecification introduced by fragments is resolved in dialogue. To give a one-line characterisation of this theory, it can be viewed as “an attempt to enrich dynamic semantics with techniques for encoding the contribution of pragmatics” (*A&L* p.180). In the next section we briefly introduce two concepts or techniques on which the theory builds; in Section 8.1.2 we describe those elements of the theory we use for our application. The description can be relatively short compared to that of $LL$ in Chapter 6, since we will only make a small addition to the theory, for which we don’t need to go too much into the details. In any case, however, we refer the interested reader to the original papers cited here and below in the text.[1]

### 8.1.1 Some Required Background Concepts

#### 8.1.1.1 Dynamic Semantics

The logical formulae we have seen so far—for example the ones we used in various examples to represent the meaning of natural language clauses, but also the language for describing formulae we have introduced in Chapter 6 ($LL$), and the language of typed feature structures used in HPSG—had in common that they are intended to be interpreted *statically*.[2] In a static logic, the central notion is that of satisfaction, i.e. whether a certain structure, possibly together with a certain variable assignment, satisfies a formula. In such a logic the ‘meaning’ of a statement can be defined as the set of models that satisfy it. In a dynamic logic another notion becomes central, namely that of *context change*. In such a logic, formulae can be seen as ‘programs’ that give instructions on how to change the context in which they are evaluated (van Benthem 1996). That is, formulae have the potential to change the structure and assignment function with which they are evaluated; their ‘meaning’ can then be equated with this potential.

This makes such logics particularly useful for modelling the semantics of *discourses*, where sentences can introduce new entities into the ‘universe of discourse’ and there can be anaphoric elements that refer back to these entities. Dynamic semantics captures the idea that the prior sentence in the discourse changes the context in which the current sentence is interpreted, and through doing this it allows one to represent anaphoric dependencies across sentence boundaries. The advantage of this strategy is illustrated by the following mini-discourse.

[1] A succinct introduction to SDRT is (Lascarides & Asher in press), whereas *A&L* provides all the details.
[2] Although we actually only for $LL$ explicitly specified a semantics, we implied that the other languages were interpreted in a ‘standard’ fashion.

(302) a. John does not like a woman.
b. She is his teacher.

The second sentence contains two pronouns which presumably are intended to refer back to entities introduced in the previous sentence. The only way in a static logic to express this would be by re-using variables and extending the scope of the quantifiers that bind them so that the formula representing the second sentence is included. This leads to an overly-complex method for constructing such logical forms for discourse. In a dynamic logic, on the other hand, the first sentence is taken to introduce two entities into the discourse, which can be picked up by the second sentence. In other words, the sentences are 'chained' together; the output context of the first is the input context of the second.

The study of constraints on anaphoric potential is the main contribution of dynamic semantics as introduced by (Kamp 1981, Heim 1982, Kamp & Reyle 1993, Groenendijk & Stokhof 1991). So-called constraints on *accessibility*, which are sensitive to the structure of the representation, can explain why for (302) the only reading of the first sentence that is coherent in this context is the one in which there is a certain woman that John doesn't like (as opposed to the reading which amounts to 'John doesn't like any woman')—otherwise the second sentence would not have an antecedent for 'she'.

We do not have to go into technical details here; from now on we will simply assume that the formulae we use for the representation of natural language clauses get such a dynamic interpretation. We can do this without making any changes to what we've done so far; we will now simply assume that MRSs describe (base-language) formulae that are interpreted dynamically. This illustrates nicely an advantage of using semantic underspecification, namely that the description and base language can have different semantics, and that the choice of semantic interpretation adopted for the base-language need not affect the construction of logical form in the grammar.

Even though it is possible to equip predicate logic with a dynamic semantics (cf. (Groenendijk & Stokhof 1991)), we will later see representations inspired by the notation introduced by *Discourse Representation Theory* (DRT) (Kamp & Reyle 1993), and so we give here one example of such a *Discourse Representation Structure* (DRS).[3] (303) shows the DRS for the reading of (302) discussed above. In this configuration, the variables appearing in the top field of the (outer) box are interpreted as being existentially quantified, and anaphoric links are symbolised by equating variables. The atomic formulae are essentially *tests* on the input context (namely whether they support the claim made, here for example whether it is true that *x named John*), while adding a discourse referent transforms that context (by making available a new referent for later anaphoric expressions).

[3] As the name suggests, SDRT was originally an extension to DRT; in later incarnations SDRT has become more general and now only assumes some form of dynamic logic as its basic building blocks.

(303)

| *x y z k* |
|---|
| *named(x, John)*<br>*woman(y)*<br>¬ [ | *like(x,y)* ]<br>*teacher(z)*<br>*of(z,k)*<br>*z= y*<br>*k = x* |

Once one moves from the semantics of clauses to that of discourses consisting of clauses, the question of how such representations are constructed becomes important. The representations of the clauses are constructed by the grammar, but how are they combined into semantic representations of discourses? How are the anaphoric relations resolved? It is here where the main addition of SDRT lies. Whereas in DRT new clauses are first translated into DRSs which are then simply merged, in SDRT this relation of new information to the context is much more complicated, due to the introduction of rhetorical relations (cf. Chapter 2). To compute this connection, SDRT takes into account pragmatic information. Since, as argued at many places before in this thesis, pragmatic information is inherently defeasible and must be revisable in light of new evidence, a special kind of logic is needed to model this. So before we turn to SDRT we will briefly have to say something about this kind of logic.

#### 8.1.1.2 Non-Monotonic Inference

According to ‘standard’ notions of inference, adding information does not affect the validity of an inference. That is, the following holds for $\vdash$:

$$\text{If } \Gamma \vdash \phi \text{ then } \Gamma, \psi \vdash \phi$$

Hence, inference is *monotonic*. Obviously, non-monotonic logics do not have this property. But why would anyone want to give it up? The classic motivating example for the use of such non-monotonic inference is the “Tweety-example”: if you hear that Tweety is a bird, you normally infer that Tweety can fly. On learning that Tweety in fact is a penguin, however, you will retract this inference, and infer that Tweety cannot fly. Hence, in this case, the conditional above does *not* hold; adding new information makes you retract an inference.

As this very simple example illustrates, such a notion of inference is very useful for reasoning with incomplete information, where there is the chance that subsequent information refutes earlier conclusions. Pragmatic information is a prime example of information of this kind; for example, if it concerns the intentions of interlocutors, about which we can never be sure. We again won't have to go into technical details of how such inference patterns can be formalised—for a textbook introduction to non-monotonic reasoning see (Brewka, Dix & Konolidge 1997), for the particular kind used in SDRT see (Asher 1995); an automated theorem prover for this logic is described in (Schlangen & Lascarides 2002*a*). We only point out here that such logics must rely on consistency checks, so as to guarantee that the set of conclusions is always consistent with the facts. In the Tweety-example the conclusion that Tweety can fly becomes inconsistent once the information about Tweety's species is added (provided of course that the fact that penguins don't fly is recorded as well). From now on, we will simply assume that we have available an operator '$>$' that stands for 'normally implies' (so '$\phi > \psi$' means "if $\phi$ then normally $\psi$", c.f. (Asher & Morreau 1991)), and an inference relation '$\mid\!\sim$' that evaluates it.

## 8.1.2 SDRT

In this section we give the promised overview of the elements of SDRT that are relevant for us. We begin with explaining how discourses are represented in this theory; then we look at how (descriptions) of such representations are built in an incremental fashion. Finally, we describe how the pragmatically preferred reading of a discourse is identified, and how this all works together to resolve underspecification.

### 8.1.2.1 Representation

We have already motivated the need for rhetorical relations in representations of discourses (see Chapter 2). What we have deliberately left open there is what we think is related with such relations. Here, SDRT re-uses the concept of labelling content described above in Chapter 6, by labelling the representations of clauses and using these labels as arguments to the rhetorical relations. This makes it possible to uniquely identify bits of information in the context, even when they express the same proposition.[4] In that sense, these labels give a handle on individual utterances of propositions (or questions, or requests), i.e. individual *speech acts*, and hence are also called *speech act discourse referents*.

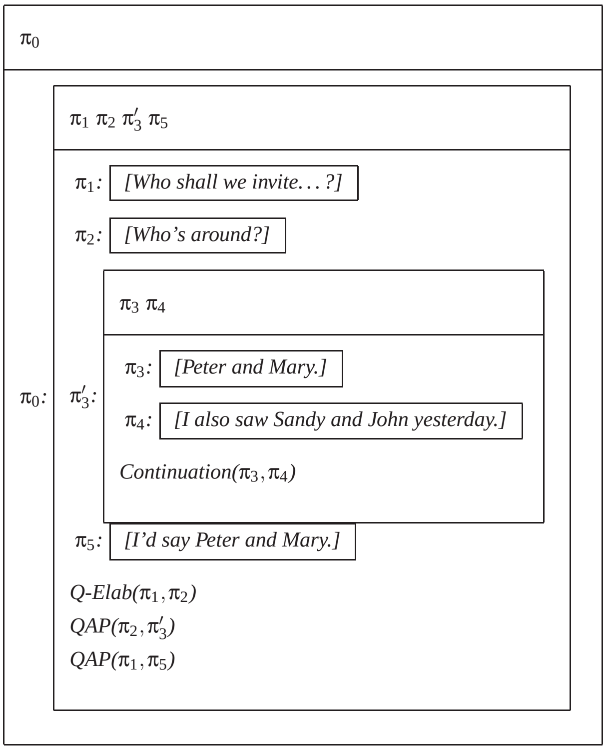
To make this concrete, let's consider an example. (304) shows a short dialogue; we have assigned each clause a speech act referent of the form $\pi_n$.

| | | | |
|---|---|---|---|
| (304) | $\pi_1$ | A: | Who shall we invite to the dinner party? |
| | $\pi_2$ | B: | Well, who's around? |
| | $\pi_3$ | A: | Hm, Peter and Mary. |
| | $\pi_4$ | | I also saw Sandy and John yesterday. |
| | $\pi_5$ | B: | So, I'd say Mary and John. |

[4] This avoids certain philosophical problems concerning the individuation of propositions; however it means that SDRT is committed to *representationalism*, i.e. the discourse representations are not dispensable as in Montague-semantics. For further discussion see *A&L*.

In the DRT inspired notation alluded to above, this dialogue can be represented as follows.[5] Note how $\pi_3$ and $\pi_4$ form a segment (labelled $\pi'_3$) that is the argument to the relation *QAP*, i.e. the content of this whole segment is an answer to $\pi_2$. We assume that the labels $\pi_1, \ldots, \pi_5$ label dynamic representations for the clauses; this is symbolised by the clauses in square brackets.

(305)




It is important to stress here that the labels $\pi_n$ are part of the discourse representation 'base language', i.e. unlike in $LL$ they are not a means to an end (underspecification); as arguments of the relations they are an integral component of the representations for discourses in SDRT. *A&L* also define an description language for SDRSs, where labels can also be used to represent 'missing' information. This language is needed for the definition of $update$ (construction of SDRSs), but before we come to this, we have to mention some additional consequences of introducing rhetorical relations into the representations, and we also have to explain how structures like those in (305) are interpreted.

First we note that the rhetorical relations introduce additional structure. The only internal structure of DRSs is given by the embedding of sub-DRSs in other DRSs (for example by the negation operator in (303) above). As we have said, this structure has a bearing on what the candidates are for antecedents of pronouns. SDRSs have more structure, namely that built by the rhetorical relations (additionally to that built by logical operators and connectors), and this structure has a similar influence on the anaphoric potential of pronouns, and also on where new information can attach (which we will deal with in the

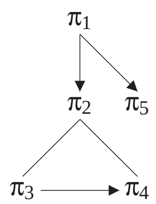
[5]Note that this is only a notational variant for the language of discourse representation introduced in *A&L*; however, this suffices here and we will not formally introduce either language.

next subsection). For example, in (306) below (repeated from Chapter 1) the second clause seems to block a connection of the fragment to the first clause, whereas in (305) above the material intervening between $\pi_1$ and $\pi_5$ does *not* block the latter from connecting to the former.

(306) Paul: Peter gave Mary a letter. Then she gave him a present.
Sandy: #From Joe? (= Did Peter give Mary a letter from Joe?)

SDRT models this via a distinction between subordinating relations (of which *Q-Elab* and *QAP* are instances) and coordinating relations (e.g., *Continuation*). The hierarchical structure built by this distinction can be graphically presented for (305) by the following graph, where subordination is represented by downwards connections and coordination by sidewards connection. The notion of *availability* determines that only labels on the 'right frontier' of this graph are available for rhetorical connection, i.e. only labels that are not the first argument of a coordinating relation (except when *Parallel* or *Contrast* hold, where this right-frontier constraint breaks down). We will return to this in the next subsection.

(307)




As we have said above, the content of the individual clauses in the discourse is represented by dynamic formulae (for example DRSs) in SDRT, and we have said in the previous section a little about how those are interpreted. But how is the contribution of the relation evaluated? Let's look at a part of (305), the relation $Continuation(\pi_3, \pi_4)$. The meaning of this is defined in SDRT as the 'chain' of context changes brought about by interpreting in order the content labelled by the first argument, the content labelled by the second argument, and the consequences of the relation. Formally, this looks as follows, for a *veridical* relation *R*.[6] ($K_{\pi_n}$ stands for the formula labelled by $\pi_n$.)

(308) $(f)[\![R(\pi_1,\pi_2)]\!]_M(g)$ iff
$$(f)[\![K_{\pi_1} \wedge K_{\pi_2} \wedge \phi_{R(\pi_1,\pi_2)}]\!]_M(g)$$

The $\wedge$ in (308) is dynamic, which means that the output context of the first conjunct is the input context for the second. $\phi_{R(\pi_1,\pi_2)}$ stands for the semantic consequences of the relation, parametrised by $\pi_1$ and

[6] *Veridical* relations are those that entail that the content of both of their arguments is true. Relations relating questions to other material for example can't be veridical, since questions do not have a truth-value. *Alternation* (representing the logical connective $\vee$) is not veridical, as one expects given the truth function of that connective.

The definition is slightly simplified compared to *A&L*. The logic they use is intensional, i.e. has a world parameter, which can be changed by certain relations involving imperatives or corrections; for expository reasons, we leave out this parameter.

$\pi_2$. These $\phi$s are the location where the semantic contribution of the relation comes in. For example *Continuation*, as defined in Chapter 2, demands that there is a (non-trivial) common topic between its arguments. Formally, this is expressed by (309) (where $\sqcap$ is the operation of topic-forming; and $\alpha$ and $\beta$ are meta-variables over labels, i.e. for our example they would be replaced by $\pi_3$ and $\pi_4$).

(309) `Consequences of Continuation`

$$\phi_{R(\alpha,\beta)} \rightarrow \neg\Box(K_\alpha \sqcap K_\beta)$$

In words, (308) says that $\phi_{R(\alpha,\beta)}$ is true only if the topic which summarises $K_\alpha$ and $K_\beta$ is a contingent proposition. As this illustrates, knowing (for whatever reason) that the relation *Continuation* holds between two discourse segments constrains these segments; they are not free anymore to talk about non-related topics, for example. Observe how this differs from the semantics of predicates like 'love'—the latter are tests on the context, as we mentioned above, while $[\![Continuation(\pi_1, \pi_2)]\!]$ might transform the context. In that sense, the rhetorical relations in SDRT 'fill in' bits of content that are only implied but not said. As we will see, such meaning postulates can also resolve explicit underspecification, namely if they are only consistent with a certain way of resolving it. This is one way of getting rid of fragment underspecification in SDRT. There is another one, but before we say more about this, we have to explain how such SDRSs are constructed in the theory.

#### 8.1.2.2 Construction

In the previous section we said that construction of DRSs for multi-sentence discourses is an easy task, given the DRS for the context and that for the new clause: the representations are simply merged.[7] SDRT adds rhetorical relations to the representations, and hence has to do much more when integrating new information (or equivalently, when updating the context with new information): it has to decide where and with what relation to attach the new information.[8] Inferring such relations is necessary: after the whole discourse has been processed (or to move this more towards dialogue again, at specified points like turn transition points), for any clause that has been processed there has to be at least one connection to another clause. This is how the notion of *discourse coherence* is spelled out in this theory, as discussed in Chapters 1 and 2.

We have said above that the question where new information can attach to is decided by what is *available* in the context, that is by the structure of the discourse so far. The next question is with which relation(s) the new information should be connected to the available sites. To infer these relations, often more information is needed than what the grammar for example has at its disposal. For instance, that $\pi_3$ and $\pi_4$ in (305) above should be connected via *Continuation* is a consequence of their content, not their

[7]The situation is a bit more complex, because sentences with embedded clauses like 'if $X$, then $Y$' lead to different, more structured updates, but in general the process of updating the representation is one of merging.

[8]Again we simplify here; another decision that must be made is *when* to make a connection, and in which order to connect, but we ignore that here.

syntax. We have also already said that such inferences must be defeasible, i.e. it has to be reasoning of the form ‘as long as there is no evidence to the contrary, infer this’. For instance, in example (310), the third clause gives information which makes clear that the temporal order of the described events corresponds to the order of clauses, and the second clause does not offer an explanation of the first.

(310) Max fell.
John kicked him.
John always kicks people who are lying on the ground.

So the logic in which we infer the *R*s has to be non-monotonic. As mentioned above, such logics rely on consistency checks, and this leads to a problem: if we need to know about the content of the utterances that are to be related, and if this content is represented in a language that is at least first order (as that of DRSs is), then checking the consistency of conclusions will not be decidable. SDRT offers an interesting way out of this dilemma, by separating the logic of information content and that of information *packaging*, in which the relations are inferred. The latter logic is only granted restricted access to the content (and to other information sources, as we will see presently), and hence can be kept decidable.

The separation between the logic of content (which interprets SDRSs) and the logic of construction of representations is done in a way similar to how $LL$ is separated from its base logic: the construction or glue logic works on *descriptions* of the content, and hence can have a different semantics.[9] We won’t go into the details of the description logic here, apart from saying that it is a conservative extension of $LL$. In this language all aspects of discourse structure can be underspecified, from the representations of the content of the clauses to the values of rhetorical relations. For the latter *A&L* write $?(\alpha,\beta,\lambda)$, which can be read as “β attaches to α with a relation I don’t know yet, whose label however is λ”.

We have noted in Chapter 6 the methodological similarity between the constraint-based grammar formalism HPSG and the underspecification formalism $LL$, which basically offers a way of specifying constraints on BLFs. To this SDRT now adds a constraint-based approach to the construction of discourse structures. The update of the context with new information is in this theory defined as a collection of all constraints that can be non-monotonically inferred from the available information. More formally (but still slightly simplified), the result of updating a (description of the) context $\text{ULF}_c$ with a (description of) new information $\text{ULF}_n$ is defined as:

(311) $$update(\text{ULF}_c, \text{ULF}_n) = \bigcup\{\psi \mid ?(\alpha,\beta,\lambda), \text{ULF}_c, \text{ULF}_n \mid\!\sim \psi\,\}$$

for all pairs $\langle\alpha,\lambda\rangle$, where α is available in $\text{ULF}_c$ for the label of the new information β and λ is a possible

[9] But note that the logic of construction is *not* the logic of description. The latter logic is a static and monotonic logic, relating descriptions of SDRSs to SDRSs, whereas the former is a non-monotonic logic that relates descriptions to descriptions.

label for the relation. $\mid\!\sim$ is a non-monotonic inference relation as explained in the previous section. Note that *update* as defined above does not make decisions about what β attaches to; it is the result of set union over all possible attachment points.

Since $\mid\!\sim$ embeds $\vdash$, the resulting ULF is at least as specific as $\text{ULF}_c$ and $\text{ULF}_n$, i.e. contains all information that is in these descriptions. However, what we are interested in is possible additional information, for example about rhetorical relations. This information is inferred via axioms in the glue-logic, which are of the form (312).

(312) $(?(\alpha,\beta,\lambda) \wedge \mathit{Info}(\alpha,\beta)) > R(\alpha,\beta,\lambda)$

This rule schema expresses that if we want to connect β to α, and certain stuff about α and β is true, then if there is no evidence to the contrary we can infer *R*. But what is this *Info* about α and β that is needed to infer relations? Recall that the glue-logic only works on descriptions, i.e. only 'knows' about the form of the content. As we have said above, in the case of *Continuation*s for example it is the content that determines whether there is a common topic or not. In other cases, information about cognitive states of the dialogue participants might be needed, or domain or world knowledge, or, more linguistically, information about the lexical semantics of elements of the utterance. But can this information be made available without compromising the computability of the glue logic? The trick *A&L* use here is to define computable translation functions that transfer information from these sources into the glue logic. Again, we will not go into details, and simply assume that descriptions of information from these sources will be available. Later we will add syntax to these sources available to the glue-logic.

For concreteness, (313) shows the rule for inferring *Explanation*. In words, this rule says that if β is a proposition ($\beta : |$) and there is evidence that β causes α ($\mathit{cause}_D$), then infer this relation. This might seem like evading the problem, since now the burden is on the predicate $\mathit{cause}_D$; however, this encapsulation allows very fine-grained rules for what counts as such evidence.

(313) `Rule for inferring` *Expl*

$(?(\alpha,\beta,\lambda) \wedge \mathit{cause}_D(\beta,\alpha) \wedge \beta : |) > \mathit{Expl}(\alpha,\beta,\lambda)$

For example, (314) shows how information from lexical semantics (about event-types) can be used to infer a cause-relation. Roughly, the rule expresses that if β mentions an internal state (like 'being angry') and α an intentional action (like 'shouting'), then β is a cause for α. Note that even though (314) is a monotonic rule, together with (313) this does not mean that in such a case *Expl* is inferred monotonically; conflicting evidence can still block the application of (313). For more details and some examples, see below the section on this relation.

(314) $([Intentional\text{-}action(e_\alpha, x)](\alpha) \wedge$
$[Internal\text{-}state(e_\beta, x)](\alpha)) \quad \rightarrow cause_D(\beta, \alpha)$

As one last example for such rules, (315), the axiom for inferring *Q-Elab*, shows that such detailed information is not always required to infer a relation. As *A&L* show, this rule can be derived in a logic of cognitive modelling from axioms about cooperativity and goals.

(315) `Inferring` *`Q-Elab`*
$(?(\alpha, \beta, \lambda) \wedge \beta :?) > Q\text{-}Elab(\alpha, \beta, \lambda)$

So although the (dynamic) semantics of $Q\text{-}Elab(\alpha, \beta)$ (in the base-language) is defined in terms of beliefs and intentions, one can infer that it is part of the LF of the discourse purely on the basis of sentence moods. In words, `Inferring` *`Q-Elab`* means that the default rhetorical role of a question is to seek information that will help achieve some prior goal (underlying the prior utterance).

Have we now defined the pragmatically preferred update, which is what we claimed SDRT does? Not quite. Recall that the result of *update* as defined above in (311) is still a description. It is a collection of constraints, containing all possible inferences with all possible pairs of labels. The pragmatically preferred SDRS will be somewhere in the set, but picking it out requires an additional constraint, which we will come to now.

#### 8.1.2.3 Picking out the pragmatically preferred structure

As we said above, the description that is the result of *update* (understood as set of SDRSs) contains all possible ways of connecting the new information to the old. So one further task SDRT has to tackle is to distinguish among the members of this set, i.e. to give reasons to prefer one structure over another. And there is more to do. The set also contains all possible ways of *continuing* the discourse as well, since, as with $LL$ in Chapter 6, the description is satisfied by models that contain nodes which aren't referenced in the description. Or in other words, the model can contain more information, as long as it satisfies the description. Hence, what we also need is some form of a minimality constraint that restricts in a principled way the size of the model.

*A&L* tackle these tasks in a principle called *Maximise Discourse Coherence* (MDC). This principle is applied when 'extracting' a structure that satisfies the description, and so strictly speaking it is not part of the construction of the discourse description. We will not give the formal definition here, but only list the informal statement of the principle (after (Lascarides & Asher in press)). It is formulated in the form of a comparison between two *descriptions*, and hence defines a partial order on them, the maximum of which will be a description of the pragmatically preferred reading.[10] Note that this does allow that

[10]Note that this partial order does not necessarily have only one maximum, i.e. it is possible that more than one reading is

there is still underspecification even in the preferred reading; it might just be the case that the discourse doesn't afford a full disambiguation.

Here is the principle, consisting of the ordering and the interpretation principle proper:

1. The coherence ordering:
   (a) All else being equal, prefer the structure that minimises the number of labels, *proviso* certain structural constraints are satisfied.
   (b) All else being equal, the more rhetorical connections there are between two items in discourse, the more coherent the interpretation.
   (c) All else being equal, the more underspecification is resolved, the higher the coherence.
   (d) All else being equal, an interpretation which maximises the quality of its rhetorical relations (more on this below) is more coherent than one which doesn't.
2. *Maximise Discourse Coherence*: Interpret the discourse so that coherence is maximised. That is, the set of SDRSs that represent the interpretation must be in the update and ranked higher according the ordering principles in 1 than any other interpretation in update.

It is time now to bring the discussion back to what we actually want to use SDRT for, namely the resolution of fragment-underspecification.

#### 8.1.2.4 Underspecification

All the ingredients needed to resolve underspecification are already present: the combination of *update* and MDC will remove underspecification as far as possible. How?

There are two basic ways resolution of underspecification happens in SDRT, corresponding to two possible ways of how information can flow. The first is taken in those cases where a rhetorical relation is inferred based on certain evidence (maybe cue words, or, as with *Q-Elab*, sentence modes). In these cases the semantic consequences of the relation might not already hold, but since they are (in a 'shallow' version) part of the glue-logic, they will also be inferred and added to the description. This can bring about that in the resulting description the underspecification is resolved, if only one way of resolving it is compatible with the additional constraints. Here, the information flows from inferring relations to adding content which might resolve underspecification.

The other way of resolving underspecification relies on MDC, and on the fact that at any point the description *update* describes an infinite number of SDRSs, including those where underspecification is resolved in the way intended by the speaker, and including those where there are rhetorical relations that weren't explicitly inferred during *update*. As it were, we 'only' have to pick the right SDRS with

pragmatically preferred, in which case the discourse is ambiguous.

the right resolution. This is where MDC comes in. Recall that this principle prefers a maximal number of rhetorical relations, but also wants the relations to be of maximal quality. We have postponed above saying how this quality is determined. Some relations are inherently scalar, like for example *Narration*: the better the common topic the better the *Narration*. However, all rhetorical relations are at least minimally scalar, in that they all 'prefer' to be inferred based on evidence. That is, given the infinite set of described SDRSs, an SDRS $K$ in which underspecification is resolved in such a way that a relation $R$ can be inferred (according to the glue-logic) will be preferred over an SDRS from that set in which $R$ is also present, but underspecification was resolved in a way that does not support a (glue-logic) inference to $R$. Hence, here information flows from resolving underspecification to inferring relations.

Note that this is a declarative statement of the solution, which happily deals in infinite sets. When trying to proceduralise such an approach, one will of course have to think hard about how to make this selection efficient. But even regarding the declarative statement we have to answer a possible objection: is this still computable? We have said above that keeping discourse structure computation computable is one of the goals SDRT sets itself, and now we define a partial order on infinite sets. The answer is, unfortunately it is not always computable anymore. If in the descriptions of the clauses only scope underspecification and anaphora occur, then it is computable, since there generating models is decidable.[11] If we restrict ourselves to *resolution-via-identity* fragments, then it also is still computable, since there is also only a finite number of possibilities, given a certain antecedent $\alpha$. However, once we turn to the basically unrestricted *resolution-via-inference* cases, we probably lose the property of computability. We think this is not disastrous, however. If we look at our familiar example (316), it seems that even for human interpreters, there is a certain element of uncertainty about the correct interpretation of such utterances.

(316) A: Why did Peter leave?
B: Exams.

All we know for certain from (316) is that the reason for Peter leaving involves exams, but it could be that Peter is the one supervising the exams, or taking the exams, or maybe his children have to take exams and this somehow affects him. Clearly, the more additional information we have about Peter, the more the possible interpretations are constrained. We will below briefly discuss some information sources that support an inference for a causal relationship. (But note that how this might guide the resolution process is not worked out in detailed in SDRT yet.) In any case there is a degree of 'open-endedness' to this kind of reasoning; something which starkly contrasts to the *resolution-via-identity* cases we have seen. We will not follow up the discussion here and simply note that this is a problematic area where implementations of the theory will have to use heuristics to yield results.

This concludes our description of SDRT. In the following section, we will add one minor element to the theory, to deal with the syntactic constraints on *resolution-via-identity*-fragments we have noted earlier

[11] Cf. *A&L*.

in the thesis. Once this is done, we are in theory done, since with that addition in place, SDRT as it is will do the job of resolving fragments for us. However, for the sake of concreteness, we return in Section 8.3 to all the rhetorical relations and give their formal definition, either as given in *A&L*, or, if we introduced the relation, we formalise what we gave as informal definition in Chapter 2.

## 8.2 Computing Intended Meanings: The Basic Story

To make concrete what we said in the previous section, and also to describe how the additional information source 'syntax' will be used, we will in this section go through an example. We choose a speech act which can be realized with both *resolution-via-identity* and *resolution-via-inference* fragments, namely QAP (*Question-Answer Pair*), to show how a further constraint allows us to deal in a general way with the syntactic 'parallelism' problem rather than having to build subclasses of speech acts according to whether they exhibit this phenomenon or not.

### 8.2.1 *resolution-via-identity*

The following example shows short-answers to an argument-question, for which we observed above that there is a certain syntactic influence from question to answer. This is illustrated by the inappropriateness of (317-b), which even though its content could be analysed as being identical to that of (317-a) (if parsed as functional preposition, and if functional prepositions are analysed as having no content), is not a coherent answer.

(317) A: Who came to the party?

a. Sandy.

b. #On Sandy.

We will show in this section how (317-a) is resolved to the intended content, and how (317-b) is recognised as being incoherent.

The relation that holds between (317-a) and the question is *QAP*, as described in Chapter 2. The (simplified) formal definition of its semantics shown in (318) is a straightforward formalisation of the informal one given previously, stating that $\beta$ must be true, and must be a direct answer to $\alpha$ (this is what *Answer* encapsulates).[12] The operator ˆtakes formulae to their intension.

[12] There are different ways of spelling out what a direct answer is; this issue, however, is orthogonal to what we want to explain here.

For simplicity, we omit the world indices here. Note also that this relation is *right*-veridical, since it only entails that the second argument is true. (The first, being a question, cannot be true.) For the full definition, see *A&L* Chapter 7.

(318) `Semantics of QAP`
$(f)[\![QAP(\alpha,\beta)]\!](g)$ iff $(f)[\![K_\beta]\!](g)$ and $(f)[\![Answer(\mathcal{K}_\alpha,\mathcal{K}_\beta)]\!](g)$

This semantics is as expected, but how do we infer whether the relation holds, and based on what evidence? For example, a condition on direct answers is that they are *true*. Deciding whether this is the case must clearly be based on the content of question and answer, and knowledge about the world. The glue-logic does not have access to this. Instead, SDRT makes use of certain *structural* similarities between questions and (natural language realizations of) their direct answers that are to be expected given the semantics of answer-hood. The intension of an interrogative is a partitioning of the logical space into mutually exclusive and jointly exhaustive sets of possible worlds, where these sets represent the different ways in which the variable bound by the question operator can be filled. This gives us an idea about how we can recognise, given the semantic *structures* of question and reply, whether the reply is a *potential* answer to the question: if the reply has a semantic structure that is like that of the question, except that the *wh*-"slot" is filled, then potentially it is an answer. We can only say that such a reply is *potentially* an answer, because all we can derive from this structural similarity is that the reply will denote something that is in the intension of the question.

But that is fine. Similar to deriving the rule for *Q-Elab* given above, one can reason that given assumptions about cooperativity and sincerity of the interlocutors, the utterance is believed to be true, thereby making its illocutionary purpose (by default) to provide an answer. (Note that this reasoning is short-circuited in SDRT, in the sense that no reasoning about intentions has to take place every time; it is 'hard-coded' in the rules.) The upshot of this is that information about the structure of question and answer is enough to make a default inference towards this relation. If *QAP* is inferred on the basis of this default axiom, then the answer being true will also be inferred via `Semantics of QAP`.

In a formalisation of this constraint, it has to be made concrete what it means to say "the reply has a semantic structure like that of the question, except that the *wh*-'slot' is filled". For this, *A&L* make use of the oft-observed fact that it is normally signalled in natural language answers which element of the answer fills the questioned 'slot', namely through the placing of focus in the answer.[13] Information about the focus position and the resulting focus/background-partitioning can be assumed to be present in the LF (and hence, via the appropriate translation function, in the glue-language as well). Thus, the constraint can be formalised using the structure of question and reply as shown in (319-a). SDRT uses this constraint as a normally sufficient condition to infer *QAP*, as shown in (319-b).[14]

(319) a. `QAP Constraint`
$QAP\text{-}sat(\alpha,\beta)$ iff:

[13] This phenomenon of question–answer-focus congruence, as it is often called, has been observed as early as (Paul 1880), for a very brief overview of the literature see for example (Rooth 1996).

[14] Both rules are taken from *A&L* p.415; we have restricted (319-b) to declarative βs (this is what the conjunct $\beta : |$ expresses) to distinguish this relation from $QAP_q$, which will be discussed later. See also Chapter 2.

There is a mapping $\vartheta$ of the focus and background partition of $\beta$ into a partition of $\alpha$ such that:

1. $\vartheta(\mathit{focus}(\beta))$ is a variable bound by an operator in $\alpha$ that's introduced by a *wh*-element if there is one—otherwise, $\vartheta(\mathit{focus}(\beta)) = \emptyset$; and
2. $\vartheta(\mathit{background}(\beta)) \leftrightarrow \mathit{background}(\beta)$

b. QAP
$(?(\alpha,\beta,\lambda) \wedge \mathit{QAP\text{-}sat}(\alpha,\beta,\lambda) \wedge \beta : |) > \mathit{QAP}(\alpha,\beta,\lambda)$

The constraint (319-a) is of a form that we will encounter again for other relations: a certain structural similarity—realised as a mapping between the formulae—is required, which is guided by information about focus/background-partitioning here; and additionally certain semantic requirements have to be fulfilled by the elements that are linked by this mapping, which here are that the backgrounds must be semantically equivalent.[15]

As we said above, the definitions of the rhetorical relations do not have to be changed in any way to deal with fragments. Short answers carry enough information already to apply *QAP-sat*, since the fragment phrase always at least contains the focus, and parts of the background are 'missing'.[16]

If we assume that this information about focus/background structure is transferred to the descriptions, then the rules given in the previous section already resolve this kind of fragment, in the following way. The compositional semantics of the fragment (317-a) describes (among infinitely many other readings) the content that is paraphrased by "Sandy came to the party". Hence, there is an SDRS in the set of described structures after updating the context with (317-a) which contains this reading. Given the focus/background-structure (inherited from the fragment), *QAP-sat* is satisfied, and hence a structure containing this resolution and a relation *QAP* is preferred over one that doesn't. So, when MDC is

[15] It has to be *equivalence* rather than identity to allow full sentence answers like the one in the following example.

(i) A: Who is talking to [the woman in black]$_1$?
B: [Peter]$_F$ is talking to [Mary]$_1$.

This is a 'shallow' form of equivalence, however, which is computed at the level of the glue logic and not the logic of SDRSs.

[16](Ginzburg 1999*b*) seems to suggest that the fragment *is* the focus, but that seems too strong. There are other constraints on fragments that must be satisfied, for example that the fragment must be a phrase. This rules out (i-c) in the example below (after (Büring 1999)), even though it is exactly the focus part of the full-sentence answer.

(i) a. A: What kind of caftans did the pop stars wear?
b. B: All the pop stars wore [DARK]$_F$ caftans.
c. B: ??Dark.
d. B: [Dark]$_F$ ones. / [Dark]$_F$ caftans.

Sometimes it is optional to repeat parts of the background:

(ii) A: What kinds of dark clothing did the pop stars wear?
B: Dark [CAFTANS]$_F$. / [CAFTANS]$_F$.

Also, certain modifiers are allowed in the fragment that aren't part of the focus, e.g. in "Who came to the party?—Only Sandy".
So, a more accurate formulation is to say that the fragment *contains* the focus.

applied on *update*, the correct resolution is picked.

#### 8.2.1.1 The syntactic constraint

As the final element of our story, we now have to explain the syntactic constraints on *res-via-id*-fragments. Above we gave the functional-PP-fragment "on Sandy" as an example of a fragment that is incoherent because of its *form* and not its content; we also repeat here one of the (German) examples we gave earlier. The explanandum in both cases is how slighltly different form can make such a difference.

(320) a. A: Wem$_{dat}$ hast Du geschmeichelt?
*A: Who did you flatter?*
b. B: [Dem Mann]$_{dat}$.
*B: The man.*
c. B: #[Den Mann]$_{acc}$.
*B: The man.*

There are two possible approaches to this question one could take in our framework. One could put additional, *syntactic* constraints on the antecedents of the rules for inferring certain speech acts, as we did in (Schlangen 2002, Schlangen & Lascarides 2002*b*); constraints that demand categorial congruence between $\alpha$ and $\beta$. This, however, is not very elegant, because in effect it re-introduces through the backdoor a sub-classification of certain speech act-types (e.g., in syntactically parallel and syntactically non-parallel short answers), which we wanted to avoid in Chapter 2 on grounds that the classification in *res-via-id* and *res-via-inf* seems independent from the fragment-(speech-act-)type. Making such sub-classification of fragment types misses generalisations, for example that both *QAP* and *Elaboration* can be either. Moreover, such a classification would be motivated by reasons other than the truth-conditional impact of the relation, which as we said should be the only criterion in SDRT for introducing new relations. So we will not follow this approach here, but rather show how these apparent syntactic constraints on certain speech acts can be made a direct consequence of how the fragment is resolved. The strategy can be described as follows: if the semantic constraints of a certain relation force a resolution that is *semantically* very close to an antecedent ('close' in a sense that we have to make precise), then we also demand a certain syntactic 'closeness'—what we have called syntactic parallelism above. The 'modus tollens' of this constraint is that if this syntactic parallelism is violated, then the semantically 'close' resolution cannot be the intended one.

To achieve this, the only modifications of SDRT we have to make are (a) to introduce a new structural rhetorical relation, which is a generalization of *Parallel*; and (b) to put an additional, syntactic constraint on this new relation. This relation—we will call it *G-Parallel* (for "generalized *Parallel*")—is similar to *Parallel* and *Contrast* in that it is inferred if its arguments are semantically very 'similar' or are 'contrasting' (we come to the differences between these notions in a minute),[17] which meets the first part of the description of our approach given above. Unlike *Parallel* and *Contrast*, however, we will allow *G-Parallel* to connect arguments of all message types, not just propositions (this is the generalization alluded to in the name of the relation). Consequently, *G-Parallel* cannot be veridical; moreover, the structural similarity must be defined such that it can relate questions and propositions.

What we then want to achieve is that semantically (g-)parallel resolutions which don't also fulfill the syntactic conditions are ruled out. This is important, since we don't just want to block an inference to *G-Parallel* if the syntactic constraint (which will be further specified in a minute) is violated, we really want to throw away the whole SDRS with this resolution from the set of described discourse structures. Or in other words, what we want to express is something like "if you could infer *G-Parallel* if it wasn't for a violation of the syntactic constraint, then throw away this SDRS".

To this end, we put rule (321) as a general constraint on *update*; again we stress that it is *not* a glue-logic axiom (where it would just block *Parallel*). The predicate *syn-constr* will be defined in a minute.

(321) $\textit{G-Parallel}(\alpha,\beta,\lambda) \wedge \neg \textit{syn-constr}(\alpha,\beta) \rightarrow \bot$

We can now explain how this rules out certain fragments as incoherent. Let's take the example from the begining of this section, repeated here as (322).

(322) A: Who came to the party?
B: #On Sandy.

If we assume that the fragment has the same semantic representation (in the base language) as an NP, then an inference to *QAP* is validated (as described in the previous section for "Sandy"); moreover, we assume that an inference to *G-Parallel* is validated. However, as we will define the syntactic constraint below, it is *not* satsified here, and this, via (321), has the consequence that the whole SDRS is 'removed' from the set (by adapting the description accordingly). This means that a *resolution-via-identity* is not available for this fragment in this context, and, if there is no other relation that fires, the discourse is considered incoherent.

The final element that is still missing is stipulating this syntactic constraint. Since we tie it to *G-Parallel*, we have to be careful that we are not too restrictive. Of course syntactic congruence is not required in general for all pairs where we want to infer *Parallel* (note that $\textit{Parallel}(\alpha,\beta) \rightarrow \textit{G-Parallel}(\alpha,\beta)$, since the latter simply is a generalized variant of the former). For example, (323) is a perfectly fine parallel

[17] This semantic similarity or contrast demanded by these relations partially refers to a similarity of the semantic representations, and so in a sense is still syntactic similarity (of LFs). However, this is distinct from the (surface) syntactic "similarity" which we want to express here, and so we will continue to use the term 'semantic similarity' for this kind of (LF-)syntactic closeness.

pair, even though not all elements are of the same syntactic category.

(323) John made Joe eat a cake and
John also forced Joe to eat a sandwich

What we want to achieve is that the syntactic constraint *syn-constr* ensures that if *a fragment* is resolved to be a complement of some predicate—note that adjuncts are not subject to this constraint—then there is congruence of syntactic category between the fragment phrase and the element it is (semantically) parallel to. Thus, we first of all need information about syntactic categories to be available in the glue logic. We will assume that there is a translation function that transfers information from the grammar about the subcategorizations of all argument-taking elements (e.g., verbs and nouns). This information will be present on the semantic labels of these predicates, so for example if $l_3 : bark(x)$ is part of the description we will also have access to a predicate $l_3 : arg\text{-}st(NP[nom])$. Note that this information allows us to align syntactic categories with semantic variables, in this case $x$ with $NP[nom]$. Moreover, we assume that information about the category of the fragment-phrase is present. With this in place, we can formulate *syn-constr* as follows.

(324) `Syntactic Constraint on G-Parallel`
$syn\text{-}constr(\alpha, \beta) \leftrightarrow$
a) There is a partial isomorphism between the DRS-structure of $K_\alpha$ and that of $K_\beta$, and
b) no argument of a predicate in $K_\beta$ has a syntactic category-specification different from what the *arg-st*-specification of that predicate demands.

Note the careful wording of clause (b): what (Ginzburg 1999*b*) calls syntactic *parallelism* is here only indirectly a parallelism, namely via the parallelism of the *predicates* in $\alpha$ and $\beta$. The constraint (324) only demands that *if there are specifications*, they are met *within each clause*. First of all, we assume that only fragment-phrases carry syntactic specifications, and so only for (resolutions resulting from) fragment underspecifcation is this relevant. Secondly, even if we assumed that for all arguments the categorial specification is transferred into the glue-logic, the constraint would still be satisfied, since congruence is only tested between the predicate which takes as argument the semantic index of the content of the phrase and that phrase. Thirdly, this formalisation explains without any additional assumptions why there are the same syntactic constraints on specifications of optional arguments, for which there is no parallel element in $\alpha$.

#### 8.2.1.2 Discussion of *syn-constr*

Let us now briefly elaborate on the syntactic constraint, and discuss some more problematic examples.

The first question we have to answer is which information exactly we assume is being transferred from the grammar to the discourse module. It should be clear that not all syntactic information is required; for example, agreement information like person and number clearly does not have to congrue between fragment and antecedent: whereas “who” in the following example is specified as singular, the fragment is not.

(325) A: Who is coming to dinner?
B: My parents.

We assume that only two kinds of information are relevant (at least for English): information about case, and about verb-form. Technically, we assume that the translation function converts these feature values into atomic terms, which we represent as follows: $\text{NP}_{acc}$, VP[*bse*], etc.

We now show how constraint (324) can explain our German example, repeated here as (326):

(326) a. A: $\text{Wem}_{dat}$ hast Du geschmeichelt?
*A: Who did you flatter?*
b. B: $[\text{Dem Mann}]_{dat}$.
*B: The man.*

First, as described above, we assume that together with the semantic information we can access information about subcat-requirements, which is connected to the semantic information via reuse of labels. So if for example the relation *flatter* is labelled with $l_5$, we also find something like $l_5 : \langle NP_{nom}, NP_{dat} \rangle$. Similarly for the fragment, where we have information about the category. Now, this information is present on the description level—only there do we have labels—but we also assume that it carries over to the base language level. A (rather crude and very much simplified) base language representation of (326-a) is shown in (327-a), with the syntactic information about subcat-requirements shown alligned. A representation corresponding to a plausible resolution of the fragment (326-b) is shown in (327-b), again adorned with syntactic information. This resolution is described by our underspecified semantic representation for the fragment—along with infinitely many others. However, only this resolution triggers an inference to *G-Parallel*. To satisfy *syn-constr*, we first check whether we have to copy over information about subcat-requirements (which in this case we have to, for *flatter*, since this information is only present on (327-a)), and we then test whether the requirements are in fact satisfied (here they are, since the $NP_{dat}$ coming from the fragment is compatible with $\langle NP_{nom}, NP_{dat} \rangle$).

(327) a. $\langle NP_{nom}, NP_{dat} \rangle$
$?\lambda x.praise(peter, x)$
b. $praise(peter, student)$

$\langle NP_{nom}, NP_{dat} \rangle \; NP_{dat}$

It should be clear from this example how the constraint rules out "[Den Mann]$_{acc}$" as a coherent short answer to (326-a), even though this fragment has the same semantic representation as (326-b).

We should stress that this mechanism works for all kinds of syntactic elements that take complements. This can explain the optionality of prepositions in certain constructions, which we illustrated with the following example:

(328) A: Who can we rely on? { B: On Sandy.<br>B: Sandy. }

This optionality directly falls out of our approach: both fragments (the PP and the NP) have the desired resolution (which can be paraphrased as "we can rely on Sandy") as part of their set of described formulae, and the PP satisfies the syntactic requirements of "rely", whereas the NP satisfies that of "on", and so the syntactic constraint is satisfied in both cases.[18]

We now discuss some problems of the approach. First, VP-questions as in (329).

(329) A: What did he do?
B: Sing a song.
B′: # To sing a song.
B′: Nothing.

This example shows two things: first, there does seem to be a syntactic constraint on the fragment (at least if we assume that there is no *semantic* difference between $VP[bse]$ and $VP[inf]$), and second there is an ambiguity between VP and NP here. Our approach forces us to assume that there is a verb "do" that takes both $VP[bse]$ complements and a very restricted class of NPs (negative quantifiers, basically, e.g. "not much", "not a lot", etc.). Secondly, we have to assume that the corresponding predicate is present in the resolution: for example for (329) we need something like $do(sing(he, a\_song))$, both to trigger the parallelism and to satisfy *syn-constr*. This solution is not entirely satisfying, but we leave improvements to further work.

Another potential problem is posed by questions that ask for determiners, as in the following example:[19]

[18] Note that for this to work we have to make use of a peculiarity of the ERG, namely that it does represent even semantically empty elements like the "on" in this example in the descriptions (where we are still free to let them describe a tautology, i.e. to make them ineffective on the base language level), since for the syntactic constraint to work we need a label for "on".

[19] Note that the superficially similar example (i) has to be analysed differently, namely as an elided NP. This is suggested by the fact that the full NP "Peter's book" sounds odd as a reply in (330), but perfectly fine in (i).

(i) A: Which book did you read?
B: Peter's.

(330) A: Whose book is this?
B: Peter's.

The question now is whether we think that *syn-constr* is at work here—in which case we would have to assume that Ns select their DETs rather than the other way round—or not. We tend to the latter solution, since the required constraints on the fragments seem to be expressable on the semantic level: the *wh*-word "whose" introduces a *possesive_rel*, and for the fragment to be parallel it has to be a possessive as well. This clearly is semantic information, and so no additional syntactic information is required to capture this constraint.

#### 8.2.1.3 The Puzzles revisited

We now have all the elements in place to explain the puzzling data concerning *res-via-id*-fragments that we reviewed in Chapters 4 and 5.

First, the data concerning what (Morgan 1973) calls 'complementizer choice'. This is directly explained by *syn-constr* as specified above. We have seen an example of a PP-fragment above in (322); the starred fragments in (331) below are ruled out as incoherent by this constraint in the same way.

(331) a. (i) A: What does John want?
B: To come over after dinner.
B′: *Come over after dinner.

(ii) A: What did John help you do?
B: *To wash my car.
B′: Wash my car.

b. A: What does John think?
B′: That Tricia has given birth to a 7-pound chin.
B″: *Tricia's having given birth to a 7-pound chin.
B‴: *For Tricia to have given birth to a 7-pound chin.

In Chapter 5 we noted that the syntactic 'parallelism' sometimes seems to be flexible, in that it under certain conditions allows elements to be dropped. We showed this with examples we repeat in the following.

(332) a. A: Who can we rely on? { B: On Sandy.
B: Sandy. }

b. A: Who shall we give the money to? { B: To Kim. / B: Kim.

c. A: What did the message appear on? { B: On the screen. / B: The screen.

As mentioned in the previous section, we can explain this in our approach by pointing out that since prepositions are argument-taking elements as well, there will also be syntactic-specifications for them available to *syn-constr*. (This is independent from whether the preposition is represented in the logical form or not.) Hence, we assume that the PP-fragment is licensed because it fills the specification of the verb, whereas the NP-fragment is licensed because it fills that of the preposition.[20]

We have also observed earlier that there has to be a parallelism of scope resolution between question and answer; we illustrated this with (333), where the question and the answer must be interpreted with the same ordering of quantifiers. This requirement follows automatically from our use of *Parallel*, which is maximal only if scope is resolved in a parallel way.

(333) A: Who gave a book to every student?
B: Peter.

Other problematic data from these chapters is captured by the grammar in our approach; we review this data here because only now do we have all components of the theory together. First, we have observed that sometimes there are divergences between what the verb subcategorises for in full sentences and what is allowed in 'freer' environments like fragments. We repeat the respective example here as (334).

(334) A: What does John believe?
B: That Optimality Theory is great.
B′: #Optimality Theory is great.
B″: John believes that Optimality Theory is great. /
John believes Optimality Theory is great

We can explain this pattern without having to appeal to *syn-constr*. In our theory, only B in (334) will receive an underspecified representation (that is, one containing *unknown*), and only it can resolve to a direct answer of A. So even though *syn-constr* would allow an uncomplementised sentence, this is not coherent for semantic reasons.

Further, in the discussion of the GBA we have reviewed examples like (335) below, where there seems to be a categorial mismatch between fragment-category and what the verb licenses.

[20] Note that this flexibility must be restricted somehow to capture the stricter requirements for example in German. Incidentally, this observation could form the starting point for an interesting investigation in the differences between English and German propositions, which however we have to leave to future work.

(335) A: Concerning the weather, what can we rely on?
B: That it will rain.
B′: *We can rely on that it will rain.

This is a problem that we cannot solve here, and we simply point out again that it occurs in other constructions as well.[21]

These examples showed how our theory, consisting of a grammar and a compositional semantics of fragments and constraints on their use can explain data that is problematic for earlier approaches. However, some problems remain. The confusing pattern regarding English pronouns (as shown in (336)) cannot be straightforwardly explained with our set-up (the GBA has the same problem): the verb demands NP[*non-acc*], but the only well-formed fragments in this context are NP[*acc*] ones.

(336) A: Who cooked this meal?
B: # I/ Me/ # She/ Her/ etc.
B′: I / *Me/ She /*Her /etc. cooked this meal.

At the moment, we cannot offer more than the rather *ad-hoc* solution of assuming lexical ambiguity here by having special lexical items for fragmental pronouns.

After this brief return to earlier data we can now show how the theory we use deals with the kind of data neglected by both Morgan's and Ginzburg's approach.

### 8.2.2 *resolution-via-inference*

What we have seen in the previous section were examples of *resolution-via-identity*, where a part of the intended meaning of the fragment is identified with some element in the context. Since short-answers can be both *res-via-id* and *res-via-inf*, we have to show either how *QAP-sat* helps in resolving *res-via-inf*-short answers like our familiar 'exams'-example (repeated in (337) below) as well, or if it doesn't, what does.

(337) a. A: Why did Peter leave so early?
b. B: Exams.

Let's assume for a minute that the meaning of (337-a) is represented by the LF in (338-a) (which is inspired by the analysis of the ERG, where 'why' is represented as 'for which reason' in the MRS). A resolution of the fragment to something like (338-b) would satisfy *QAP-sat* (i.e., this would be a *res-via-id*, where the fragment phrase content is 'plugged into' the questioned slot).

[21] Cf. the discussion above in Section 5.5.1.3.

(338) a. $?\lambda e'.leave(e,p) \wedge reason(e') \wedge for(e,e')$
b. $*leave(e,p) \wedge reason(e') \wedge for(e,e') \wedge exams(e')$

However, (338-b) is not well-formed. The variable bound by the question-operator in (338-a) is typed to be an event (because reasons are events), but exams aren't events (even though e.g. *sitting* them is one).[22] This type-clash can be detected on the level of descriptions, and so constraint (319-a) is not satisfiable. This is good, because it explains without additional stipulations why a *res-via-id*-strategy can't deal with examples like (337). However, it leaves us without an explanation of how (337-b) is actually resolved. To give such an explanation, we have to use the additional expressivity that SDRT affords us. First, we adapt the representation of question (337-a). Relations between events are encapsulated in SDRT in the semantics of the rhetorical relations. The relation between the two events expressed in (338-a) by '$for(e,e') \wedge reason(e')$' can in this framework be more conveniently expressed as *Explanation*$(\alpha,\beta)$ (where $e_\alpha = e$, and $e_\beta = e'$). We have said in Chapter 2 that, because of the compositional semantics of this kind of question, the question word '*why*' can be seen as a monotonic clue for a relation *Explanation*$_q(\alpha,\beta)$, which entails that all answers to the question $\beta$ are explanations of $\alpha$, and so we assume that this relation is present in any SDRS updated with (337-a). Moreover, a more thorough analysis of questions like (337-a) reveals that they presuppose that the event which they want explained actually happened. So a better representation for (337-a) in our framework would be one where there is a presupposition "Peter left early" (let's label this $\alpha'$) and a question "why is this so?" (let's label this $\alpha$), where *Explanation*$_q(\alpha',\alpha)$ holds. The semantics of *Explanation*$_q$ now can be transferred into the glue-logic as shown in (339).

(339) $$?(\alpha,\beta,\lambda) \wedge Explanation_q(\alpha',\alpha,\lambda') \wedge Explanation(\alpha',\beta,\lambda'') \rightarrow QAP(\alpha,\beta,\lambda)$$

Applied to our example (337), this means that if we a) try to attach (337-b) to the part of (337-a) we have called $\alpha$ above, which itself, as explained above, is connected via *Explanation*$_q$ to the presupposed bit $\alpha'$, and b) can prove that (337-b) explains $\alpha'$, then c) the new information (337-b) connects to the question via *QAP*. In this chain of inferences, the constraints on *Explanation* now are satisfied by an appropriate resolution of the fragment in this case, just like *QAP-sat* was satisfied by (317-a), and this explains why the fragment is coherent. Satisfying the constraints on this relation might require world-

[22] In a more thorough semantic analysis, 'exams' could be seen as denoting an underspecified event; but in any case, its semantic type is not the same as that of an event denoted by a clause.

Morevore, it seems that the most important factor in determining the acceptability of such fragments is how salient an event involving the NP is. It is difficult to find examples which evoke events with only very little context, which might explain the varying naturalness of the examples in (i). However, it is not difficult to think of contexts for each of these fragments which make them sound natural.

(i) A: Why did he leave?
B: Exams. / ?The nagging wife. / ??The wife. / The cat.

knowledge in some cases; in any case it requires more knowledge than was used in *QAP-sat*, and so this explains why we called this an instance of *resolution-via-inference*. We postpone explaining in detail the constraints on this relation *Explanation* until the discussion in the next section.

This concludes the description of our additions to SDRT. Our strategy can be summarised as follows: if a fragment is resolved in such a way that is semantically and structurally very close to its antecedent (i.e., if a generalised version of *Parallel* called *G-Parallel* can be inferred), then we demand that a certain (surface) syntactic constraint is satisfied as well. This general constraint explains why all *resolution-via-identity* fragments exhibit syntactic-cohesion phenomena, and also how there can be *res-via-id* and *res-via-inf* instances of the same speech act. With this in place, we now briefly mention how this interacts with other sources of underspecification, and then take a closer look at the classes introduced in Chapter 2.

### 8.2.3 Interaction with Other Sources of Underspecification

It is worthwile to repeat here that our model of the resolution of fragments is based on accummulating and solving constraints. For example, knowing that a given fragment is an answer constrains its resolutions in a certain way. This may or may not have an influence on other sources of underspecification, but this influence is independent from that on fragments. For example, in (340) below, the scope-ordering of the answer is not constrained, i.e. we allow both a reading where ‘a book’ outscopes ‘every student’, and one where the order is reversed.

(340) A: What does every student like?
B: A book.

This similarly holds for *resolution-via-inference* fragments. In the following modification of our standard example the approach constrains the fragment to be resolved to readings in which the reason for every student’s leaving involves in exams some way, but it leaves completely open whether this is the same set of exams for all students or not.

(341) A: Why did every student leave?
B: Exams.

Our machinery does put some restrictions on scope-orderings, however. In the following example, the readings of the fragment that mirror the ordering of the scopal elements in the question will always be preferred (since only they make *G-Parallel* inferrable, and a greater number of relations is always preferred). Any decision on scope-order in the question will hence also influence that on the fragment—this seems to be in accordance with intuitions for (342).

(342) A: What did Peter say every boy gave to a girl?
B: A book.

## 8.3 The Relations Revisited

As we said in the previous section, the constraint on *update* concerning *G-Parallel* is all we have to add to SDRT to make it handle our problem; the theory as defined in *A&L* will then provide us with the intended resolutions and make the right predictions concerning the licensing of fragments in context. However, for the sake of concreteness, we will briefly go through the definitions of the semantics of the speech acts listed in Chapter 2 and also give axioms for inferring them. Moreover, not all of the speech acts we listed there are defined in *A&L*, and so for some we have to provide new definitions here. Note that none of these new relations are specific to fragments.

In the presentation below we concentrate on two main information sources for inferring relations (and consequently, for resolving fragments), namely structural information as in the constraint on *QAP* we've shown above, and information about lexical semantics. As we will show, following *A&L* and others, this lexical semantic information is particularly useful for inferring *Elaboration* and *Explanation*, because it allows one in certain cases to restrict the required inference quite considerably. For some fragments, however, more general knowledge, for example about the domain in which the dialogue is set, or even about the world in general, or about plans for achieving goals is required. We will only show here how SDRT interfaces with these knowledge sources; doing the requisite reasoning in these areas is a problem in its own right. As we will show, one of the advantages of using SDRT is that it allows us to restrict the need for this kind of reasoning as much as possible, and to encapsulate the linguistic task of building a logical form for a dialogue (and hence the sub-task of resolving fragments as well). As a consequence of this encapsulation, the level of detail given will vary for the different fragment types, depending on the kind of information needed to work with them. For example, we will deal with speech act types related to plans only cursorily here, whereas we will go into some more detail for *Elaboration*s.[23]

### 8.3.1 The Types in Detail

The structure of the following subsections is as follows. We repeat the informal definition of the semantics from Chapter 2, formalize it, and then give a few illustrating examples of fragments instantiating this kind of fragment. The (fragment-) speech act types are described here in the same order as in Chapter 2.

[23] But note that the implementation described in the next chapter does deal with plan-related relations and so we will say something there about how plans and communicative intentions interact with the resolution of fragments. We can do this in the implemenation because there the domain of the dialogue is sufficiently restricted that the goals can be fixed.

#### 8.3.1.1 Question-Answer-Pair

We have already looked at this speech act type in some detail in the previous section, and so we only summarize this discussion here. First, the (simplified) semantics of the relation, as given informally in Chapter 2 and formally in Section 8.2 above.[24]

(343) a. `Semantics of` *QAP*`, informally`
$\beta$ provides a direct answer to $\alpha$.

b. `Semantics of` *QAP*`, formally`
$(f)[\![QAP(\alpha,\beta)]\!](g)$ iff $(f)[\![K_\beta]\!](g)$ and $(f)[\![Answer(\tilde{K}_\alpha,\tilde{K}_\beta)]\!](f)$

We have seen above how SDRT makes use of structural relations between questions and possible answers when inferring whether this relation holds between an element of the context and the new information; we formalized this in rule (319), repeated here as (344).

(344) `Rules for inferring` *QAP*

a. $QAP\text{-}sat(\alpha,\beta)$ if: There is a mapping $\vartheta$ of the focus and background partition of $\beta$ into a partition of $\alpha$ such that:

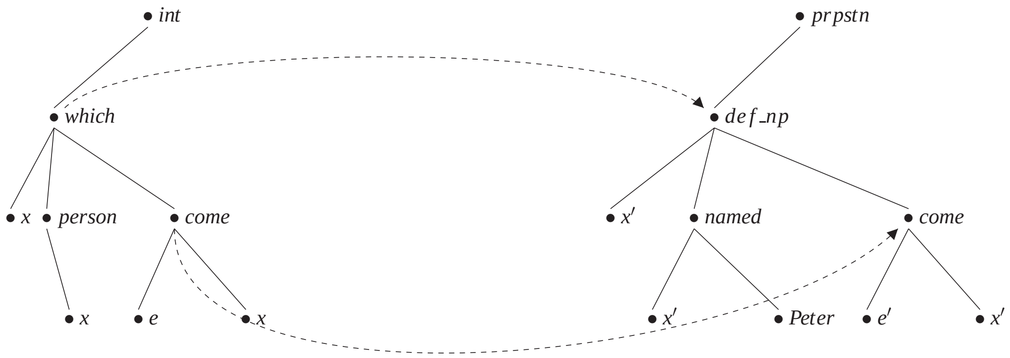
1. $\vartheta(focus(\beta))$ is a variable bound by an operator in $\alpha$ that's introduced by a *wh*-element if there is one—otherwise, $\vartheta(focus(\beta)) = \emptyset$; and
2. $\vartheta(background(\beta)) \leftrightarrow background(\beta)$

b. $(?(\alpha,\beta,\lambda) \wedge \beta : | \wedge QAP\text{-}sat(\alpha,\beta,\lambda)) > QAP(\alpha,\beta,\lambda)$

Example (345) shows two BLFs for which *QAP-sat* is satisfied, with the mapping as indicated by the arrows. (We omit the contribution of "the party" from the example.) The BLF shown on the right is a possible resolution of "Peter.".

[24] Again we omit the world indices. See *A&L* for the full definition.

(345) “Who came to the party?” — “Peter came to the party.”




We should note that there is another way of formalising this constraint. For all *resolution-via-identity* fragments, we could define resolution on ULFs, for example using the puzzle-parallelism-constraint of CLLS (Egg et al. 2001), which are used there to handle VPE. However, that would entail that we had to ‘know’ beforehand whether something is *res-via-id* or not. Our declarative statement does not require this information, and so allows the more general treatment of the syntactic constraint detailed in the previous section. As we will see in the next section, however, when it comes to practical implementations, a less general but more efficient method is probably required.

There are two things we do not formalize here. As in the previous chapters, we have concentrated on short answers to single *wh*-questions. However, our approach should in principle be extendable to short-answers to multiple *wh*-questions. Answers like “$[\text{Peter}]_F$ $[\text{with Sandy}]_F$” have two foci, and so we need to extend the mapping $\vartheta$ so that it maps both of them to *wh*-elements. The *syn-constr* on *Parallel* will then take care of mapping the fragment elements onto the intended *wh*-elements.

The other thing we did not deal with are questions like “what did the girls eat?”, under a reading where answers like “the SMALL girl ate BLACK beans, and the BIG girl ate RED beans” are expected. (I.e., questions where question-operator is outscoped: ‘for each girl *x*, what did *x* eat?’.) We suggest that the difference to ‘normal’ cases of *QAP* lies in the different semantics of the question here, but we will not further investigate this problem.

#### 8.3.1.2 Question-Answer-Pair, $_q$-Version

As before, we first give the semantics of this relation (which is not in *A&L*).

(346) a. `Semantics of` $QAP_q$`, informally`
Positive answers to y/n-question β provide a direct answer to α, negative answers a partial answer.

b. `Semantics of` $QAP_q$`, formally`
$(f)[\![QAP_q(\alpha,\beta)]\!](g)$ iff
$yn\_ques(\beta)\wedge$
$(\forall p(Answer(\mathcal{K}_\beta, p) \rightarrow (pos\_answ(\mathcal{K}_\beta, p) \wedge Answer(\mathcal{K}_\alpha, p)) \vee$
$(neg\_answ(\mathcal{K}_\beta, p) \wedge PAnswer(\mathcal{K}_\alpha, p))))$

In this definition *yn_ques* is a property that is true of yes-no-questions, and *pos_answ* and *neg_ans* are relations that hold between y/n-questions and their positive and negative answers, respectively; *PAnswer* encapsulates the concept of partial-answers, i.e. answers that only rule out some potential answers, but do not give "positive" information. An example of such a partial answer is shown in the following:

(347) A: Who did you talk to? Peter?
B: No.

The answer "no" in this example resolves to something that can be paraphrased as "I didn't talk to Peter"; this is a partial answer to A's first question insofar as it takes away a proposition from the answer set (namely "I talked to Peter."), but doesn't narrow this set down to only one partition.

We have said already when we introduced this type in Chapter 2 that it is very similar to "normal" short-answers in that the fragmental question would count as an answer if it were uttered in a declarative intonation. This intuition is validated by the semantics of the relation. The structural similarity between question and answer, which we used as one criterion for inferring *QAP*, is carried over in (346-b) from α to β by the requirement that all answers to β are also answers to α. And so we are justified in simply re-using the constraint *QAP-sat* in a rule for inferring this speech act type. The only difference to (344-b) above is that here we have to restrict the sentence mood of β to 'question'. This rule will then resolve certain $QAP_q$-fragments in the same way as described above.

(348) `Rules for inferring` $QAP_q$
$(?(\alpha,\beta,\lambda) \wedge \beta :? \wedge QAP\text{-}sat(\alpha,\beta,\lambda)) > QAP_q(\alpha,\beta,\lambda)$

Note that we can also find *res-via-inf* fragments of this type, as in example (349). Again we assume that this fragment is resolved via the additional constraints on *Explanation*, and postpone the discussion of this example to the section on *Explanation* / $Explanation_q$.

(349) A: Why did he leave so early? Exams?

#### 8.3.1.3 Elaborations of propositions, and requests for them

In this section we deal with pairs of utterances (or discourse segments) where the first is asserted and the second is either a proposition that elaborates the first one, or an interrogative that requests such an elaboration. We begin with what we called $Elab_{pp}$ in Chapter 2. This section is comparatively long, since it introduces a new information source besides the structural information that was used in *QAP-sat*.

**Elab**$_{pp}$

This is probably the most familiar kind of elaboration, where simply a proposition is elaborated by another one. (350) shows two examples of fragments of this type.[25]

(350) a. A: Peter enjoyed the lovely meal.
A very nice salmon fillet.

b. A: Peter damaged a garment.
A shirt*(, to be precise)*.

We will return to these examples in a minute, but first give the semantics of this relation (as defined in *A&L*); the informal part is repeated from the taxonomy-chapter.

(351) a. `Semantics of` $Elab_{pp}$`, informally`
$\beta$ elaborates on some aspect of the indicative $\alpha$, e.g. by giving details about a subevent of the event described in $\alpha$, or by providing more information about participants involved in the event.

b. `(Partial) Semantics of` $Elab_{pp}$`, formally`
$\phi_{Elab_{pp}(\alpha,\beta)} \rightarrow Part\text{-}of(e_\beta, e_\alpha)$

We can see how the informal definition (351-a) captures the relation between the utterances in the examples in (350). In the first one, the fragment intuitively is resolved to something like "Peter ate a very nice salmon fillet." This event of eating is understood in this context as a sub-event of the 'larger' event of enjoying a meal. In a more fine-grained analysis, we can also assume that the salmon fillet is a subtype of meal, and so we find this subtype relation also on the level of elements of the utterances.

[25] After *A&L* and (Danlos 1999), respectively.

In (350-b), this relation between elements of the utterances is even more obvious: a shirt is a type of garment.

The meaning postulate in (351-b) formalises these two aspects of elaborations: *Part-of* ensures that the main event of the elaboration is temporally a part of the main event of the elaborated proposition, while the second conjunct ensures that there is a semantic connection as well, i.e. that what $\beta$ talks about normally is a (sub-)part of what $\alpha$ talks about (e.g., from the information that someone ate a nice salmon fillet normally follows that this person had a nice meal). Both constraints express the condition on elaborations that they involve a 'zooming in' of some sort, temporally and semantically. We will now show how this necessary condition on the speech act can be turned into a sufficient one which as a side-effect explains how fragmental elaborations like those in (350) are resolved.

The main feature of elaborations is that a subtype-relation must hold between the elaboration and the elaborated proposition. If there is evidence in the discourse that such a relation holds, we can infer this speech act. *A&L* gloss this with a predicate called $subtype_D$, which holds if there is such evidence in the discourse, whatever that may be. Hence, the rule for inferring this speech act can be given as shown in (352).[26]

(352) `Rule for inferring` $Elab_{pp}$

$$(?(\alpha,\beta,\lambda) \wedge subtype_D(\beta,\alpha) \wedge \alpha : | \wedge \beta : |) > Elab_{pp}(\alpha,\beta,\lambda)$$

Of course, this rule just shifts the burden of proof to the rules for inferring $subtype_D$. There are several ways evidence that such a relation holds can be found. Often, checking for such a sub-type relation must rely on fairly open-ended reasoning with world knowledge; this is notoriously difficult to formalise, and also computationally quite expensive. As we said in the introduction to this section, we cannot say much about this here, and only show how such reasoning interfaces with the framework we use. However, as *A&L* show, there is a good source for computationally "cheap" information about subtypes, which sometimes is enough to provide evidence for $subtype_D$: the lexicon. In many cases we can assume that we have available lexical semantic knowledge about type-relations between predicates, and this can help infer certain relations.[27] We will not review here the evidence for having such relatively detailed knowledge about semantic classes in the lexicon,[28] but rather just summarise what SDRT offers us that can help to resolve fragments of this kind.

The SDRT-rule for inferring $subtype_D$ shown in (353) makes use of semantic type information that is transferred from the lexicon into the glue-logic in the form of the relation $\sqsubseteq$, where $y \sqsubseteq x$ means that $y$ is a subtype of $x$. The predicate $\theta_i$ in this rule represents the thematic role the argument plays, and so in

[26] *A&L* do not describe elaborations of interrogatives, and so only have what we call $Elab_{pp}$ and $Elab_{pq}$ (for which they don't give a formal definition). To distinguish the variants we have to explicitly state the sentence mode in our rules here.

[27] See also the discussion in (Asher & Lascarides 1995).

[28] The view of the lexicon as containing highly structured information goes back to (Pustejovsky 1991); the observations in (Levin 1993) *inter alia* provide further support for this view.

words the rule expresses that if we have an element of α and one of β so that these elements stand in the lexical-semantic-subtype relation and fill the same thematic role in their respective utterances, and also the main events of these utterances stand in such a subtype relation, then we can infer $subtype_D$.

(353) `Rules for inferring` $subtype_D$, `I`
$(\theta_i(x,\alpha) \wedge \theta_i(y,\beta) \wedge y \sqsubseteq x \wedge e_\beta \sqsubseteq e_\alpha) \rightarrow subtype_D(\beta,\alpha)$

This rule already explains how fragments like (350-a), which is repeated here as (354), are resolved, or, to be more precise, how (only) a certain resolution allows an inference to this relation.

(354) A: Peter enjoyed the lovely meal.
A very nice salmon fillet.

First, as *A&L* make plausible, there is lexical semantic information that the semantic index of the representation corresponding to "meal" is typed as **food**. Evidence for this is that in a construction as in the one in this example ("x *enjoyed* the meal"), the most prominent interpretation is that the event of *eating* the meal was enjoyed. We will make more use of this observation, but first we note that we can also assume that there is information from lexical semantics that the representation of the NP containing "salmon fillet" is of sub-type *food* as well, to be precise, it is the edible portion of salmon which in turn is a subtype of *food*: this part of the compound noun is built via a productive lexical rule that changes terms denoting kinds of animals into terms that denote the flesh of that animal, prepared for the purpose of eating. And so one conjunct of (353), namely $y \sqsubseteq x$, is validated with these elements being $x$ and $y$. We know the thematic role of $x$ in α—it is the *patient* of the enjoying—and so to infer this relation, we need a resolution of β where $y$ is also a patient. The last conjunct, $e_\beta \sqsubseteq e_\alpha$, helps to specify the value for the underspecified event of β. This event has to satisfy two constraints: it must be sortally compatible with $y$, which as we now know is its patient, and is typed as **food**; it also must be, via rule (353), a subtype of "eating", which, as we said above is what "enjoy the meal" by default resolves to. This narrows down the readings that allow an inference to this relation.[29] In effect, then, this constraint is not so unlike the ones we have seen so far. It also demands that there is a mapping between at least some elements of α and β, which here is mediated by thematic roles and $\sqsubseteq$.

We can also take rule (353) as a starting point for devising a rule that can resolve the other example we gave, (350-b), repeated here as (355).

(355) A: Peter damaged a garment.
A shirt*(, to be precise)*.

[29] Using this hierarchy has the welcome effect that we do not always have to resolve such fragments to predicates that have lexical counterparts; we can assume that fragments are only resolved to a level of generic predicates like *act-on*, or here a generic *eat* (which is a supertype of *devour*, *gobble down*, etc.).

For this example it presumably is not plausible to assume that the subtype relation between "garment" and "shirt" is lexicalised. However, we can assume that world- or domain-knowledge ontologies are also accessible to the glue-logic, for example encapsulated in a relation $\sqsubseteq_{WK}$. If such information is available, then the fragment in (355) can be resolved in a similar way to the example we discussed above. The respective rule is given below in (356).[30]

(356) `Rules for inferring` $subtype_D$`, II`

$$(\theta_i(x,\alpha) \wedge \theta_i(y,\beta) \wedge y \sqsubseteq_{WK} x \wedge e_\beta \sqsubseteq_{WK} e_\alpha) \rightarrow subtype_D(\beta,\alpha)$$

These rules can also already explain elaborations where optional arguments are filled, as in the following example.

(357) A: I made a purchase.
A new computer.

Following *A&L*, we assume that the representation of 'purchase' contains an argument slot for the object of the purchase.[31] In A's utterance in this example, this slot is not filled and so in the representation of the sentence it is just bound off with an existential quantifier. The fragment now is resolved simply as a filling this slot, resulting in something like "I made a purchase of a new computer." For this to work with the rules for inferring $subtype_D$, we have to either assume that *purchase* specifies a semantic type for this argument, of which the semantic index of the NP 'a new computer' is a subtype; or else we assume that if there is no explicit type specification then the type is $\top$, which is the root of the type hierarchy, to which everything is a subtype.

The following example shows that type–specifications on arguments can also be just defaults. The noun 'drinker' carries the implication 'drinker of alcohol', which however can be overridden (though with a slightly humorous effect):

[30] Note that the relation $\sqsubseteq$ 'ties' its arguments together referentially in the sense that they are not free to refer to 'unconnected' entities or events anymore. In example (355) this goes as far requiring that $x$ and $y$, and also $e_\alpha$ and $e_\beta$ co-refer: the garment that is stained *is* the shirt mentioned in the fragment, and the event of staining the shirt *is* the event of staining the garment (and not just a (proper) subpart of it). (Danlos 1999) introduces for this a new relation *Particularisation*, which explicitly states that the event-variables in its arguments and other entities co-refer. She claims that this is necessary to explain how, contrary to standard interpretation in dynamic semantics, indefinites like 'a garment' and 'a shirt' can be forced to co-refer rather than introduce new information. In our set-up, however, this is just a limiting case of our relation $\sqsubseteq$ and still covered by this definition, and so we do not need a relation different from *Elaboration* for this. Moreover, co-reference even of indefinites as a result of semantic consequences of relations is nothing special in SDRT; for instance, the answer in the following example also has to 'co-refer' to the variable in the question bound by the question operator, and hence also co-refer to the demonstrative in the question.

(i) A: Who was that?
B: A man from the inland revenue.

[31] Having such optional arguments is a property of all nouns resulting from nominalisations of verbs.

(358) A: Peter is a heavy drinker.
Green tea, mostly.

More open-ended reasoning with world-knowledge is needed to resolve the following fragments, although the general line of reasoning is the same. However, here we cannot find elements of $\alpha$ and $\beta$ that are in a subtype-relation, but rather we have to infer a sub-*event* of $e_\alpha$ which can takes the fragment-phrase as argument (e.g., "watch a film", "eat good food", etc.).[32]

(359) a. Max had a lovely evening yesterday.
A nice film, lots of good food, wine, dancing.

b. I went to the cinema yesterday.
Spiderman.

The rules we have seen so far cannot deal with examples like (360) below, because they are restricted to mappings between *complements* (which fill thematic roles). Hence, for these fragments we need an additional rule.

(360) a. A: I talked to Peter.
On the phone.
b. A: John kissed Sandy.
In front of the Empire State Building.

If we look at the intended resolutions of these fragments—e.g., for (360-a) this would be "I talked to Peter on the phone"—the relation between this and the elaborated utterance is clear: the fragment just 'adds' a modification. If such a relation is present between two utterances, we can take this as evidence that the one containing the modification (i.e., the one giving more restricted information) is a subtype of the other, as formalised in (361) below. Note that this constraint demands that all variables (except those introduced in the modifier) co-refer.

(361) `Rules for inferring` $subtype_D$`, III`
If $\beta = \alpha + modifier$ then $subtype_D$

We have now seen the major types of rules for inferring this relation, and we can come to some more problematic examples, where the subtype relation is context dependent and has to be built "on the fly". A prominent type of this is the 'referential zooming in' in the following example.[33]

[32] In (359-a) we have actually have a sequence of elaborations, possibly with an internal narrative structure, but we gloss over this here.

[33] We ignore here the presuppositions of definites (like proper names); in a fully formulated approach clarifications of proper

(362) A: I talked to Peter.
Peter Miller.

The ‘subtype’ relation here is one between sets of possible referents, where a specification via a full name provides more information than one via just a first name. This explains why the following is no coherent elaboration, and even sounds odd.[34]

(363) A: I talked to Peter Miller.
?#Peter.

Such information could be ‘hard-coded’ in the glue-logic in a constraint that states that ‘first names’ are always a subtype of ‘full names’ (the form of which can be recognised in the glue logic); this however is not a very general solution, since the following cases seem similar but do not involve proper names.

(364) a. A: He talked to Peter.
One of his students.

b. A: Peter talked to [the woman in red]$_1$ all evening.
He was really trying to chat [Maria]$_1$ up.

We close the discussion of $subtype_D$ with a look at a final constraint we have to impose on elaborations, namely that elaborated elements must actually refer to some entity. This explains why A’s first utterance in (365-a) must be read *de re*, and why (365-b) is odd.

(365) a. A: John seeks a unicorn.
Amalthea*(, to be precise)*.

b. A: John talked to no one.
?#Peter*(, to be precise)*.

The rules for inferring $subtype_D$ we have given above are the basis for inferring *Elab*; they capture the core of the elaboration aspect. The rules for inferring the other kinds of *Elab* are like (352), except that they impose different requirements on the sentence modes, as indicated by the subscripts. The meaning postulates of these relations, however, are quite different, and will be discussed in some detail in the following.

---

names would be $Elab_q$s of the presupposed content and not directly of α.

[34]We will encounter similar examples when we discuss what often is called ‘clarification-question’.

**Elab**$_{pq}$

This class contains questions where all answers are elaborations of their antecedent; an example is shown in (366). The semantics of the relation is defined in (367); this definition contains that for $Elab_{pp}$ given in (351-b) above, connected to $\beta$ via *Answer*. In other words, $\beta$ is a question where every answer is an elaboration of $\alpha$. Note that this relation is left-veridical (i.e., $\alpha$ is asserted).

(366) A: I talked to Peter.
B: Peter Miller?

(367) a. `Semantics of` $Elab_{pq}$`, informally`
Any answer to $\beta$ elaborates on some aspect of the indicative $\alpha$, e.g. by giving details about a sub-event of the event described in $\alpha$, or by providing more information about participants involved in the event.

b. `Semantics of` $Elab_{pq}$`, formally`
$f[\![Elab_{pq}]\!]g$ iff $f[\![K_\alpha \wedge \phi_{Elab_{pq}(\alpha,\beta)}]\!]g$
$\phi_{Elab_{pq}(\alpha,\beta)} \rightarrow$
$\forall p(Answer(K_\beta, p) \rightarrow Part\text{-}of(e_p, e_\alpha)$

This relation is inferred using the $subtype_D$ predicate from the previous section; only the restrictions on sentence mode are different.

(368) `Rule for inferring` $Elab_{pq}$
$(?(\alpha,\beta,\lambda) \wedge subtype_D(\beta,\alpha) \wedge \alpha : | \wedge \beta :?) \rightarrow Elab_{pq}(\alpha,\beta,\lambda)$

Note that we do not have to introduce a further sub-classification to get the difference between B and B′ below in (369) (one is a *wh*-ques, the other a *y/n* one); both kinds are covered by (367).

(369) A: I read an interesting manuscript.
B: Which one?
B′: My memoirs?

#### 8.3.1.4 Elaborations of interrogatives, and requests for them

We now come to elaborations of questions. As discussed when we introduced these types in Chapter 2, their semantics must make reference to intentions.

**Elab**$_{qp}$

First, elaborations/clarifications of one's own questions:

(370) A: Did Peter call?
Peter Miller.

(371) a. `Semantics of` $Elab_{qp}$`, informally`
β elaborates on the *intention* behind asking α, e.g. by giving details about a sub-event of the main-event of α, or by providing more information about participants involved in the event.

b. `Semantics of` $Elab_{qp}$`, formally`
$Elab_{qp}(\alpha,\beta)$ iff
$\forall p'\exists p[(Answer(\check{}K_\alpha, p) \wedge Answer(\check{}?K_\beta, p')) \rightarrow (\check{}p' \rightarrow \check{}p)]$

In words: something is an elaboration of a question if all answers to the polar-question formed with this proposition (that's what $?K_\beta$ means) entail an answer to the original question.

We infer this relation (and hence resolve fragments of this type) via $subtype_D$, as before:[35]

(372) `Rule for inferring` $Elab_{qp}$
$(?(\alpha,\beta,\lambda) \wedge subtype_D(\beta,\alpha) \wedge$
$\alpha :? \wedge \beta : | \wedge Agent(\alpha) = Agent(\beta)) \rightarrow Elab_{qp}(\alpha,\beta,\lambda)$

If we resolve the second fragment in our example (370) above to "Did Peter Miller call?", we can infer $Elab_{qp}$, because all (true) answers to 'Did Peter Miller call?' entail (true) answers to 'Did Peter call?' in a given world *w*, provided that 'Peter' and 'Peter Miller' co-refer in that world. (In fact, the answers will be equivalent.)

**Elab**$_{qq}$

Finally, clarifications of questions.

(373) (α) A: Did Peter call?
(β) B: Peter Miller?

Here is the constraint on $Elab_{qq}$ (where $B_{Ag_\alpha}\phi$ means 'the speaker of α believes ϕ'):

[35] $Agent(\alpha) = Agent(\beta)$, which means that α and β must be produced by the same speaker, might be a bit too strict, since other speakers can elaborate a question if they are 'collaboratively' asking a question, but we gloss over this.

(374) a. `Semantics of` $Elab_{qq}$, `informally`

Any answer to β elaborates on the *intention* behind asking α, e.g. by giving details about a sub-event of the main-event of α, or by providing more information about participants involved in the event.

b. `Semantics of` $Elab_{qq}$, `formally`

$Elab_{qq}(\alpha,\beta)$ iff

(a) $\forall p, p'(Answer(K_\alpha, p) \wedge Answer(K_\beta, p') \rightarrow (\check{p}' \rightarrow \check{p}))$

(b) $\exists q[q = ?\forall p'(Answer(K_\beta, p') \rightarrow \textsc{sarg}(\alpha, B_{Ag_\alpha} p'))] \quad \wedge$

(c) $\forall p[Answer(K_\beta, p) \rightarrow Answer(q, p)]$

In words: $Elab_{qq}$ holds between α and β, if β is an elaborating question in the sense above, i.e. all answers to it elaborate answers to the original question, and also all answers to β answer the question "is knowing an answer to β a speech-act-related-goal of the original question α?". This gets at the aspect of such elaborations that they 'change' the original intention, or, to put it differently, that they commit the speaker of α to now want to know answers to the more specific question β. The new elements we have to introduce here are the modal operator $B$ for beliefs, which is parametrized by the bearer of the belief, and the relation SARG which connects an utterance (label) with a conventionalized goal.

Reasoning about intentions is not required when *inferring* the speech act type (and hence when resolving the fragments), and so the rule looks as follows:

(375) `Rule for inferring` $Elab_{qp}$

$(?(\alpha,\beta,\lambda) \wedge subtype_D(\beta,\alpha) \wedge$
$\alpha :? \wedge \beta :? \wedge Agent(\alpha) \neq Agent(\beta)) \rightarrow Elab_{pp}(\alpha,\beta,\lambda)$

#### 8.3.1.5 Correction

In Chapter 2 we introduced two kinds of correction, one for corrections of indicatives and one for corrections of interrogatives. We deal with the first kind first. (376) shows the formal definition of its semantics, as given in *A&L*.

(376) a. `Semantics of` $Corr_1$, `informally`

β corrects (an element of) α.

b. `Semantics of` $Corr_1$, `formally`

$[\![Corr_1(\alpha,\beta)]\!]$ iff

1. there is a bijection ζ from the focus/background structure of $K_\beta$ onto the logical forms of subclausal constituents of $K_\alpha$, such that:

$K_\alpha \mid\!\sim Apply[\zeta(Focus(K_\beta)), Bg(K_\beta)]$; and

$Apply[\zeta(Focus(K_\beta)), Bg(K_\beta)] \mid\!\sim K_\alpha$.

Also,

2. $K_\beta$ is inconsistent with $K_\alpha$.

The first clause of (376-b) is of a similar form as *QAP-sat* above: it requires that there is a mapping between the arguments of the relation (where again that mapping is guided by the focus/background structure of β), and it imposes a semantic constraint on the parallel elements. This semantic constraint demands that if we 'replace' the focus of β—which is the main correcting element—with its parallel element from α, we get something that is defeasibly equivalent to α. *A&L* give the following dialogue as an illustration:[36]

(377) A: They gave Peter the new computer.
B: No, $[\text{John}]_F$ got the computer.

To satisfy (376-b), the bijection must map the element of β corresponding to "John" to "Peter", so that the result of substituting the focus of β by its parallel element in α, "Peter got the computer", is defeasibly equivalent to "they gave Peter the new computer". This constraint can also explain the referential identity that is required, as shown in (378) (which is similar to (i) in note 15 above).

(378) A: Peter talked to $[\text{the woman in red}]_1$.
B: No, $[\text{John}]_F$ talked to $[\text{Mary}]_1$.

Although this is not explicitly mentioned in *A&L*, this constraint can also handle 'sloppy-corrections', i.e. sloppy-pronoun readings of corrections. (Gardent et al. 1996) gives the following example of such a correction. (The sloppy reading is easier to get if one assumes a context in which only one man can like his wife.)

(379) A: $\text{Peter}_1$ likes $\text{his}_1$ wife.
B: No, $\text{Sam}_2$ likes $\text{his}_{1/2}$ wife.

Similarly, a sloppy reading (as well as a strict one) is available for the following correction containing VP-ellipsis.

(380) A: $\text{Peter}_1$ talked to $\text{his}_1$ mother.
B: No, $[\text{the teacher}]_2$ did *(talk to* $\textit{his}_{1/2}$ *mother)*.

[36]Cited from (van Leusen 1994).
Note that we analyse "no" here as a cue-word for this relation and not as an anaphor.

Examples like this can also be constructed with fragmental corrections, as shown in (381).[37]

(381) A: John introduced Mary to her future husband.
B: No, $[\text{Sandy}]_F$.

To capture this reading, we have to allow an ambiguity in the specification of the background (note the similarity to the Higher-Order-Unification approach to VPE described in Chapter 3) as to whether all instances of the focus are abstracted over or not. For our example (381) this means that we have to allow both $\lambda x.introduce(john,x,y) \wedge wife\_of(y,x)$ and $\lambda x.introduce(john,x,y) \wedge wife\_of(y,mary)$ as background.

The second clause of (376-b) adds the requirement that the correction has to be inconsistent with the correctum, i.e. that not both $\alpha$ and $\beta$ can be true at the same time. This means that even if $\neg\alpha$ is not entailed by $\beta$ it is implicated; this explains why the following dialogue carries the strong implicature that John doesn't speak German, even though speaking German should be logically compatible with speaking French.

(382) A: John speaks German.
B: No, French.

The meaning postulate (376-b) can again be transferred into the glue-logic, in a similar way to that for *QAP*. We gloss this glue-logic version as *Corr-sat*, and so the glue-logic rule for inferring $Corr_1$ looks as shown in (383).

(383) `Rules for inferring` $Corr_1$
$(?(\alpha,\beta,\lambda) \wedge \alpha : | \wedge Corr\text{-}sat(\alpha,\beta,\lambda)) > Corr_1(\alpha,\beta,\lambda)$

We now come to the variant of correction where the correctum is an interrogative, as in the following example.

(384) A: Did you talk to Peter?
Erm, no, to Paul, I mean.

[37] Interestingly, it seems that if VPE is available, fragments are dispreferred, as the relative awkwardness of (i-a) and the preference for interpreting (i-b) as correcting the object illustrates.

(i) a. A: Peter talked to his mother.
B: ?No, Sandy. / No, Sandy did.
b. A: Peter loves Sandy.
B: No, Paul.

The dialogue effect of such a correction is clear: the speaker of the original question corrects the content of that question (and hence the proposition he or she wants to know). We have expressed this in Chapter 2 in the following informal definition:

(385) `Semantics of` $Corr_2$`, informally`
$\beta$ corrects (an element of) the intended content of the interrogative $\alpha$.

This forms the starting point for our formalisation of the semantics of this speech act type. The following meaning postulate simply constructs the appropriate question by applying the focus of $\beta$ to the background of $\alpha$; for example in (384) the focus of $\beta$ is *paul*, and the background of $\alpha$ is something like $\lambda x.?talk\_to(you,x)$, and hence the resulting question can be paraphrased as "Did you talk to Paul?". The revised goal behind the original question then is to know answers to the corrected question. In that sense, $Corr_2$ has an element of a 'meta-talk' speech act, because it is 'about' an earlier utterance (similar to the elaborations of interrogatives above).

(386) `Semantics of` $Corr_2$`, formally`
$[\![Corr_2(\alpha,\beta)]\!]$ iff $\forall p(Answer(\check{}(?Apply(Focus(\beta),Bg(\alpha))),p) \rightarrow Answer(\alpha,p))$

To infer this relation (and hence, to resolve fragments of this kind), we simply re-use the constraint *Corr-sat* from (383) and only adapt the requirements on the sentence mode of $\alpha$.

(387) `Rule for inferring` $Corr_2$
$(?(\alpha,\beta,\lambda) \wedge \alpha :? \wedge Corr\text{-}sat(\alpha,\beta,\lambda)) > Corr_2(\alpha,\beta,\lambda)$

We should stress that these are not the only effects of corrections: they can also correct assumptions about discourse structure, i.e. about what the illocutionary function of previous utterances was. To give an idea of how this could happen, recall that all except discourse-initial utterances will be rhetorically connected to previous utterances. Veridical relations entail that both of their arguments are true, while corrections crucially imply that the corrected one isn't. Hence, not both can hold, and discourse structure has to be revised. This however does not influence how fragments are resolved, and so we do not go into the technical details of how this is formalised.

#### 8.3.1.6 Parallel

As described in the previous section, (a generalised version of) this relation is the central element in our explanation of the difference between *resolution-via-identity* and *resolution-via-inference* fragments. We have only informally described its semantics above, and so we give it here (as defined in *A&L*).

(388) `Semantics of` ***`Parallel`***

*Parallel*$(\alpha, \beta) \rightarrow$

a) $K_\alpha$ and $K_\beta$ have similar *semantic structures*. That is, there is a partial isomorphism between the DRS-structure of $K_\alpha$ and that of $K_\beta$. All else being equal, the closer the mapping is to an isomorphism, the better the *Parallel* relation.

b) There must be a common theme between $K_\alpha$ and $K_\beta$. This is computed on the basis of the above partial isomorphism. The more informative the common theme, the better the *Parallel* relation.

Clause a) of this definition ensures that the elements of the relata are structurally parallel (similar to what *QAP-sat* and *Corr-sat* specify) while clause b) makes sure that there also is a semantic relation. This latter requirement ensures that (389-a) is a 'better' parallelism than (389-b).

(389) a. Peter loves Mary, and
she adores him.
b. A dog barks, and
Peter runs.

(Kehler 2002), following (Hobbs 1990), formalises the semantic requirement on parallelism as follows (we have already briefly mentioned this in Chapter 3). If $\alpha = p_0(a_1, \ldots, a_n)$ and $\beta = p_1(b_1, \ldots, b_n)$, then for these two propositions to be parallel there has to be a common property $p$ of $p_0$ and $p_1$, and common properties $q_i$ for all argument pairs $a_i, b_i$.[38] In (389-a), this common property $p$ would be a relatively natural 'to feel affection', and the $q_i$s would be 'humans mentioned in the discourse'. (389-b), on the other hand, can intuitively only produce the rather uninformative common property 'activities', and 'animate entities'.

Neither *A&L* nor (Kehler 2002) spell out how exactly such common or contrasting properties are computed.[39] As mentioned in Chapter 3, there are several approaches to this problem, e.g. (Asher 1993), (Prüst et al. 1994), (Grover et al. 1995), (Hobbs & Kehler 1997) and (Gardent 1999), to name but a few. These approaches all have in common that they rely in some way on knowledge about lexical or semantic entailments, possibly as given by ontologies (e.g. '*love*' and '*hate*' are both predicates of emotional attitudes, and are antonyms). We will not review this work (and its problems) here—we have mentioned in passing some of the approaches above in Chapter 3—, and just assume in the following that there is some way in our theory of computing such properties. In any case, what's more important

[38] A limiting case for this relation would be a repetition of a clause, since there both the structure and the content would be identical. However, such a repetition would presumably be dispreferred by other pragmatic constraints, and so all instances of *Parallel* can be expected to have at least one element where the common property isn't 'identity' (as in e.g. $q_1(a_1), q_1(b_1), q_1 = \lambda x(x = a_1)$).

[39] *A&L* refer to (Asher, Hardt & Busquets 2001) for details, but the method that is described there simply computes structural parallelism, relying on syntactic operations on DRSs that do not have the power to compute commonalities between properties; their method would not predict a useful common theme for (389-a).

for us now is how these meaning postulates can be turned into sufficient conditions for inferring these structural speech acts.

Clause a) of (388) can be transferred relatively straightforwardly into the glue-logic, since the requirement for an isomorphic mapping can also be defined on descriptions. The operation of computing a common theme, however, is more difficult to define in that logic. Here it depends on how much domain specific knowledge about ontologies is transferred via $\vdash_{tr}$ and hence accessible to this logic. We summarize the requirements in the following glue-logic rule, glossing the structural and semantic requirements as *Par-sat*, as we did with *QAP-sat*:

(390) `Rule for inferring` *`Parallel`*
$(?(\alpha,\beta,\lambda) \wedge \mathit{Par\text{-}sat}(\alpha,\beta,\lambda)) > \mathit{Parallel}(\alpha,\beta,\lambda)$

As the following examples show, this relation can also directly be inferred for fragments, on the basis of cue phrases (e.g. "too"). Note that typically in these examples there will be other relations present as well.[40]

(391) a. A: Peter introduced Sandy to her future husband.
Mary, too.
b. A: Peter likes Mary.
(And) Sandy, too.

#### 8.3.1.7 Contrast

The relation *Contrast* in SDRT connects sequences of utterances that have a contrasting theme. Two examples given in *A&L* are (392-a) and (392-b), where the former is of the subsort of *Contrast* called *formal contrast* and the latter is an example of a *violated expectation*.

(392) a. John speaks French. Bill speaks German.
b. John loves sport. But he hates football.

*Contrast* has a semantics that is very similar to that of *Parallel*. In fact, clause a) of (388) can be directly transferred, and the only difference is that in clause b) we require a *contrasting* theme. Because of this similarity, we forgo repeating this here and only repeat the informal definition from Chapter 2.

[40] It again seems to be the case that where there is VPE available, fragments seem to be a dispreferred way of conveying content, as illustrated by the following example.

(i) A: Peter likes Mary.
?#Paul, too. (= Paul likes Mary).

(393) `Semantics of Contr, informally`
$\alpha$ and $\beta$ have a *contrasting theme*.

The main context in which we can encounter fragmental contrasts is as follow-ups to negative answers to polar-questions. In Chapter 2, we gave an example which we repeat here as (394); the contrast in this example is *formal* and not *violated expectation*.

(394) A: Were they in tents?
B: No, caravans.
[BNC GYS 72041879]

The first element of our analysis of B's utterance here is the representation of the particle 'no'. In SDRT, this can be treated as lexically specifying the relation *QAP*, with an underspecified attachment site (since the grammar doesn't know which question is being answered).[41] We also have to express the element of negation, and so we will treat "no" as updating a context with the following formula: $QAP(?'_{\pi}, \pi, \lambda) \wedge \neg prop(?'_{\pi})$. In this formula *prop* stands for a function that returns the propositional part of an utterance, which for a y/n-question is the proposition inside of the question operator. Hence, in (394) above 'no' resolves to the representation of "they weren't in tents". It is this content to which the fragment "caravans" forms a contrast, and not A's question. The requirement for a contrasting theme is illustrated by (394) in the following way. If we resolve "caravans" to "they were in caravans", the contrasting elements of this and the negative answer (resolved to "they weren't in tents") are the negation and the 'positive operator' in "are in" and "weren't in" and the entities 'tent' and 'caravan' contrast.

Interestingly, it seems that replacing 'no' in these examples with its resolved proposition makes a difference in coherence. For instance, replacing B's utterance in (394) with either of the utterances in (395) makes the discourse odd. We currently have no good explanation for this.

(395) B′: They weren't in tents. ??#Caravans.
B″: They weren't in tents. ?#But Caravans.

Note that there is a superficial similarity between contrasts and corrections, as e.g. in the following.

(396) a. A: Peter called.
B: No, Paul.
b. A: Did Peter call?
B: No, Paul.

[41] Note that this means that the word "no" is lexically ambiguous according to our analysis, with one reading being the monotonic cue for *Correction* and the other reading as a discourse particle discussed here.

However, even though B's utterances look the same, they have very different semantic effects. First of all, they connect to different utterances in their context: the correction in (396-a) connects to A's utterance, while the contrastive fragment in (396-b) connects to the anaphorically expressed proposition $[\![$Peter did not call$]\!]$. Secondly, the correction entails that the assertion to which it connects is false, i.e., it disputes the truth value of a previous utterance. This has the repercussion for discourse structure that we discussed above. The fragmental contrast does not dispute anything but rather offers more information.

As observed by (Ginzburg 1999*b*) and discussed in Chapter 4, there also seem to be certain constraints on the focus/background partitioning of $\alpha$. We showed this in that chapter with examples we repeat below.

(397) a. A: Did [John and Bill]$_F$ leave this morning?
B: #No, Harry (= No, John and Harry left this morning.)

b. A: Can you help me with my [homework]$_F$?
B: No, but with your [carpentry]$_F$.

If we assume that the anaphora "no" receives the same partitioning as its antecedent, then we can already explain the different licensing pattern in (397): the partitioning required to get the indicated reading in (397-a) is impossible, whereas in (397-b) it is fine. Unlike (Ginzburg 1999*b*), we can also explain why the contrast cue-word in (397-b) is licensed (and indeed required).

#### 8.3.1.8 Continuation

We begin with continuations of indicatives. We gave in Chapter 2 the following example.

(398) A: I'm free on Monday.
And on Wednesday afternoon.

We have already shown the formal definition of the semantics of this relation as (309) above in Section 8.1.2, but we repeat it here as (399), together with the informal definition from Chapter 2.[42] As mentioned above, the operator $\sqcap$ in the rule stands for an operation of topic-forming.

(399) a. `Semantics of` *`Contn`*`, informally`
$\beta$ continues a topic of $\alpha$.

b. `Semantics of` *`Contn`*`, formally`
$\phi_{Cont(\alpha,\beta)} \rightarrow \neg\Box(K_\alpha \sqcap K_\beta))$

[42]This relation is in *A&L*, although it is not formalised there.

We restrict ourselves here to continuations where only one element is 'exchanged' in $\beta$ compared to $\alpha$, and so we can define a structural constraint for inferring this relation, as follows. Note that for this to work we have to assume that the conjunction in (398) is not part of the focus. Further, we need information about 'thematic relatedness', which could be formalised similar to the subtype information needed for *Elaboration*.

(400) a. `Constraint on` *`Cont`*
*Cont-sat*$(\alpha,\beta)$ if:
There is a mapping $\vartheta$ of the focus and background partition of $\beta$ into a partition of $\alpha$ such that:
1. $\vartheta(focus(\beta))$ is a variable denoting an entity that is thematically close to $focus(\beta)$; and
2. $\vartheta(background(\beta)) \leftrightarrow background(\beta)$

b. `Rule for inferring` *`Cont`*
$(?(\alpha,\beta,\lambda) \wedge \textit{Cont-sat}(\alpha,\beta,\lambda) \wedge \alpha : | \wedge \beta : |) > QCont(\alpha,\beta,\lambda)$

We have also defined in the taxonomy chapter a version of this speech act type connecting questions; its semantics is given by the following rules.[43]

(401) a. `Semantics of` *`QContn`*`, informally`
$\beta$ continues a topic of $\alpha$.

b. `Semantics of` *`QContn`*`, formally`
$$\phi_{Cont(\alpha,\beta)} \rightarrow$$
$$\forall p \forall p' (Answer(\mathcal{K}_\alpha, p) \wedge Answer(\mathcal{K}_\beta, p') \rightarrow \neg\Box(p \sqcap p'))$$

The thematic connection between the question is expressed in this meaning postulate via their answers; in words, the MP demands that for every answer $p$ of $\alpha$ there is an answer $p'$ of $\beta$ such that there is a contingent commmon theme of $p$ and $p'$ (i.e., an informative, not necessarily true one). To give an example, the fragmental question in (402) resolves something like "what is the meeting place of CS360?". Answers to this question have the theme 'properties of CS360' in common with those to A's first question.

(402) A: What is the meeting time of CS360?
B: 7.00 P.M. on Monday night.
A: The meeting place? / CS362?

[43]This speech act type is our addition to SDRT.

The requirement for an informative common topic explains why in (403) (3) is odd as a continuation of (1)-(2), whereas (4) is fine.

(403) (1) A: What is John's telephone number?
(2) B: 650 4415.
(3) A: ?#His hair colour?
(4) A: His email address?

Note that this difference is just a matter of preferences in MDC; the fragmental question (3) in the example above is still interpretable as "what is his hair colour?". In cases where all else is equal, interpretations which result in a better thematic relation are preferred. For example, in (404) the, theoretically possible, resolution "Who gave Mary to Sandy?" is greatly dispreferred compared to the resolution "Who gave Mary John's telephone number?".

(404) (1) A: Who gave Sandy John's telephone number?
(2) B: . . .
(3) A: Mary?

The rules for inferring this relation are just like those for *Cont*, making use of the same structural constraint *Cont-sat* and only restricting the arguments differently; we forgo repeating it here.

Note that as predicted by our rules of computing the background, we can find strict/sloppy-ambiguities.

(405) A: Does $\text{Bill}_i$ love his wife?
B: . . .
A: John? (= Does John love Bill's wife *or* Does John love Johns wife.)

Moreover, the constraint leaves room for ambiguity. In (406), the fragmental question can resolve both to "What did Sarah give Sandy?" and to "What did Peter give Sarah", with both readings ranked equally by MDC (via *Cont-sat*).

(406) A: What did Peter give Sandy?
B: . . .
A: Sarah?

#### 8.3.1.9 Q-Alt

The speech-act *Q-Alt* was defined in Chapter 2 as shown in (408), and was introduced with the example repeated below as (408).

(407) `Semantics of Q-Alt, informally`
$\beta$ resolves to an alternative-question involving an element of $\alpha$.

(408) A: Are you free on Monday?
Or on Tuesday?

The difficulty here is to get the combination of elements into one question right, and to make the resulting question available for attachment for example via *QAP* by its answers. We simplify here and (semi-)formalise the relation as modifying the intention behind $\beta$, where the function *combine* combines $\alpha$ and $\beta$ in the intended way.

(409) `Semantics of Q-Alt, formally`
$\phi_{QAlt(\alpha,\beta)} \rightarrow \text{SARG}(\beta, combine(\alpha,\beta))$

In the rule for inferring this relation we make use of the closeness of this relation to that of *Q-Cont* defined above, and use *Cont-sat* to identify the elements out of which the combined question is built. To avoid inferring both relations for example in (408), we make the rule for this relation more specific than that for *Q-Alt*, by demanding the presence of a cue-word.

(410) `Rule for inferring Q-Alt`
$(?(\alpha,\beta,\lambda) \wedge Cont\text{-}sat(\alpha,\beta,\lambda)) \wedge \alpha : ? \wedge \beta : ? \wedge or(\beta) \rightarrow Q\text{-}Alt(\alpha,\beta,\lambda)$

#### 8.3.1.10 Explanation

*A&L* give the semantics of the relation *Explanation* as follows.

(411) a. `Semantics of Expl, informally`
$\beta$ explains $e_\alpha$.

b. `Semantics of Expl, formally`

$$\phi_{Expl(\alpha,\beta)} \rightarrow \begin{array}{l} cause(e_\beta, e_\alpha) \wedge \\ \neg(e_\alpha \prec e_\beta) \wedge \\ (event(e_\beta) \rightarrow e_\beta \prec e_\alpha) \end{array}$$

In this formalization the predicate *cause* in the logic of content is true if its first argument is the cause of the second;[44] the other conjuncts specify the temporal constraints on this speech act type, namely that the eventuality that is being explained must not have happened earlier than the explaining one—if the latter is an event (and not a state), it in fact must have happened first.

This semantics is used to resolve fragments in the by now familiar way. Similar to *Elaboration*, contexts in which there is discourse evidence for a causal relation validate a special glue-logic predicate, $cause_D$. The rule for inferring the relation looks as follows.

(412) `Rule for inferring` *`Expl`*
$(?(\alpha,\beta,\lambda) \wedge cause_D(\beta,\alpha) \wedge \beta : | \rightarrow Expl(\alpha,\beta,\lambda)$

The information required to infer whether $cause_D$ holds might again be relatively costly world- or domain-knowledge. However, there are again cases where information from lexical semantics can help; we will briefly discuss one example where this might be the case.

Consider the fragmental question in (413). (We discuss *Expl* and $Expl_q$ at once.)

(413) A: The baby is crying!
B: Hunger?

Now, *A&L* assume that there is lexical semantic knowledge available about the kind of eventuality denoted by a verb, for example whether it denotes an intential action or an internal state. With this information, a rule like the following (which we have shown above already) can be formulated.

(414) $$([\mathit{Intentional\text{-}action}(e_\alpha,x)](\alpha) \wedge [\mathit{Internal\text{-}state}(e_\beta,x)](\beta)) \rightarrow cause_D(\beta,\alpha)$$

In words, this rule says that if there is a slot for an agent of an intentional action, and one for the possessor of an internal state, and they can be unified, then there is evidence that the possessor of the internal state is the agent of the intentional action. Or in other words, if you're hungry you might cry.

Note that even if we can't assume that all information is lexicalised, information about the difference between intentional and non-intentional actions seems to guide the inference in cases like those below in (415). The non-volitional action in (415-b) seems to be 'more in need' of an explanation than the volitional one in (415-a).

[44] There are different ways of spelling this out, using material implication or strict implication, or some common sense notion (if *p* then normally *q*). This need not concern us here, however.

(415) a. A: Max sat down on the pitch.
B: ??A foul?
b. A: Max fell down on the pitch.
B: A foul?

Incidentally, our compositional semantics of this relation also explains how ‘because of NP’ constructions are resolved, as in (416).

(416) A: Why did Peter leave so early?
B: Because of (the) exams.

We view “because of” as a monotone cue for *Explanation*, and the NP as an argument to an underspecified proposition similar to our *unknown*-constraints from Chapter 6; however, we will not further go into details here.

Just to avoid a possible misunderstanding, note that it is not the case that *res-via-inf*-explanation-fragments are ‘elided’ *because-of-NP*; both constructions have different distributions, as shown in (417). Whereas in ‘because-of-NP’ constructions all NPs seem to be allowed, in explanation fragments, as observed above, more event-y NPs are preferred.[45] Perhaps more surprisingly, as (417-b) and (417-c) show, it is not the case that explanation-fragments can always be substituted with ‘because-of-NP’ constructions. We currently have no explanation for this observation.

(417) a. (i) Peter left. Because of his wife.
(ii) Peter left. ?His wife.
b. (i) I was happy. Two weeks of holidays.
(ii) I was happy. ?Because of two weeks of holidays.
c. (i) I’m sorry, I can’t make it. Too much work.
(ii) I’m sorry, I can’t make it. ?Because of too much work.

#### 8.3.1.11 Explanation$_q$

We have already given an example of this speech-act type in the previous section, so we just provide the definitions here. Inference to this speech-act relies on the same predicate $cause_D$ as in the previous section.

[45]This should not be surprising, since the ‘because of’ construction gives a monotonic cue for inferring *Explanation*, whereas in the NP-fragment cases the possibility that the fragment offers an explanation is only one that has to be entertained. To make it easier for the hearer to understand the speaker’s speech act, a salient event should be chosen.

(418) a. `Semantics of` $\mathit{Expl}_q$`, informally`

Any answer to β explains $e_\alpha$.

b. `Semantics of` $\mathit{Expl}_q$`, formally`

$$\phi_{Expl_q(\alpha,\beta)} \rightarrow \forall p(Answer(\mathcal{K}_\beta, p) \rightarrow cause(e_p, e_\alpha) \wedge \neg(e_\alpha \prec e_p) \wedge (event(e_p) \rightarrow e_p \prec e_\alpha))$$

(419) `Rule for inferring` $\mathit{Expl}_q$

$$(?(\alpha,\beta,\lambda) \wedge cause_D(\beta,\alpha) \wedge \beta :? \rightarrow Expl_q(\alpha,\beta,\lambda)$$

#### 8.3.1.12 Result

This is the ‘inverse’ of *Explanation*, where α explains β, but apart from that the meaning postulate and the rule for inferring the relation are very similar, so we forgo showing them here.

#### 8.3.1.13 Plan-related relations

We forgo formalising the family of plan-related relations here (i.e. *Plan-Elab*, *Plan-Correction*, *Q-Elab*, *Ack*, $Ack_q$), since doing so would require the introduction of quite a few additional concepts related to plans and the semantics of actions.[46] We will give simplified versions of the relations in the next chapter.

This concludes our catalogue of rhetorical relations, and we now summarise what we have done in this chapter.

## 8.4 Summary

In this section we have shown how a pre-existing theory of discourse-semantics can use our semantics and syntax of fragments to compute intended meanings in context. We have extended this theory with a syntactic-constraint that can explain the observations by (Morgan 1973, Morgan 1989).

We should be clear about the consequences of this change. Like (Ginzburg & Sag 2001), we are proposing a change to the traditional syntax / semantics / pragmatics interface; our ‘pragmatic’ module now needs access to syntactic information. However, we note that this is independently motivated (if we construe (Kehler 2002) as proposing something similar). Further, our change is still more conservative than Ginzburg & Sag’s (2001), who need a bi-directional interface where grammar and pragmatics aren’t clearly separated (with all the resulting problems discussed in Chapter 5).

[46] Cf. *A&L* for details about these relations.

With this our theory of the interpretation of fragments is complete. However, to demonstrate the practical utility of the theory, we will discuss an experimental implementation of it in the next chapter.

# Chapter 9

# RUDI: An Implementation of the CBA

In this section we describe an experimental implementation of (aspects of) the theory of fragment-interpretation that we developed over the last three chapters. This implementation works in a very restricted domain, to make it possible to concentrate on the main issue of proceduralising the theory. The system shows how the resolution of fragments can be integrated with the resolution through discourse information of other kinds of underspecification.

## 9.1 Introduction

RUDI (which stands for Resolving Underspecification using Discourse Information) is a computer program that, building on the output of the ERG with our modification for fragments, computes certain aspects of discourse structure for dialogues in the domain of scheduling appointments, together with the semantic consequences of that discourse structure which serve to resolve certain semantic underspecifications which were generated by the grammar.[1,2]

To give an idea of how exactly RUDI enriches the output of the ERG, we now briefly discuss an excerpt from a typical dialogue in this domain, as shown in (420). (The $h_n$ in parentheses are the *labels* of the utterances; the $t_n$ are the main temporal referents introduced in the respective utterance, e.g. $t_4$ is the semantic index of "next week").

(420) $(h_1, t_1)$ A: What is a good time for you in the next couple weeks?
$(h_2, t_2)$ B: After 2pm on Monday...
$(h_3, t_3)$ ...and I'm also free on Wednesday afternoon.
$(h_4, t_4)$ A: Actually, I now see next week isn't that great.

There are basically two things RUDI computes from the output of the ERG: a logical form for the dialogue in terms of a *discourse structure*, and (parts of) the *model* that satisfies that structure. The discourse structure that is computed for (420) is shown in an (S)DRT-style notation in Figure 9.1.[3] In general, the discourse structure computed by RUDI consists of:

- the segmentation of the utterances into larger discourse units (e.g., in (420) utterances $h_2$ and $h_3$ are grouped together);
- rhetorical relations connecting these segments (e.g., $h_3$ in (420) is a continuation of $h_2$, and both together provide an indirect answer (i.e. are connected via *IQAP*) to question $h_1$); and
- resolutions of (some) underspecification in the logical forms: firstly, and most importantly here,

[1]The version of RUDI described here is based on previously published collaborative work: the first version of the system was described in (Schlangen, Lascarides & Copestake 2001), which was then substantially revised and extended for (Schlangen & Lascarides 2002*b*). (These papers are available from `http://www.cogsci.ed.ac.uk/∼das`.) For the present chapter, we again revised and extended the system, to cover more speech acts and also to reflect more closely the structure of the underlying theory.
The system also contains some code written by others: a package for handling feature-structures in PROLOG, by Michael Covington (Covington 1994), and, as part of the non-monotonic theorem prover CETP (Schlangen & Lascarides 2002*a*), code written by Patrick Blackburn and Johan Bos (Blackburn & Bos 2000) for calling external theorem provers from PROLOG.

[2]At the moment RUDI is still more a proof of concept and hence there is no publicly available release (but the code is available on request from the author); however, substantial work is being done on the system as part of the Edinburgh-Stanford-Link funded project ROSIE (*RObust Semantic InterprEtation*), which will result in a released version.

[3]We have omitted in this representation the logical forms coming out of the ERG (these are elided in square brackets); that for $h_2$ is paraphrased, since RUDI resolves the underspecified logical form the ERG provides. This presentation is also simplified compared to what is actually computed by RUDI in two other respects. First, we only show one rhetorical relation between any two utterances whereas in reality RUDI may infer more than one relation between them (plurality of discourse relations being permitted in SDRT, as described above in Chapter 8). Secondly, we only show here bridging relations (cf. references below) for what we called 'main temporal referents', while the 'afternoon' in $h_3$ (for example) has to be bridged as well, namely to the 'Wednesday' in the same utterance. RUDI captures these bridging relations too, but we omit them here for simplicity.

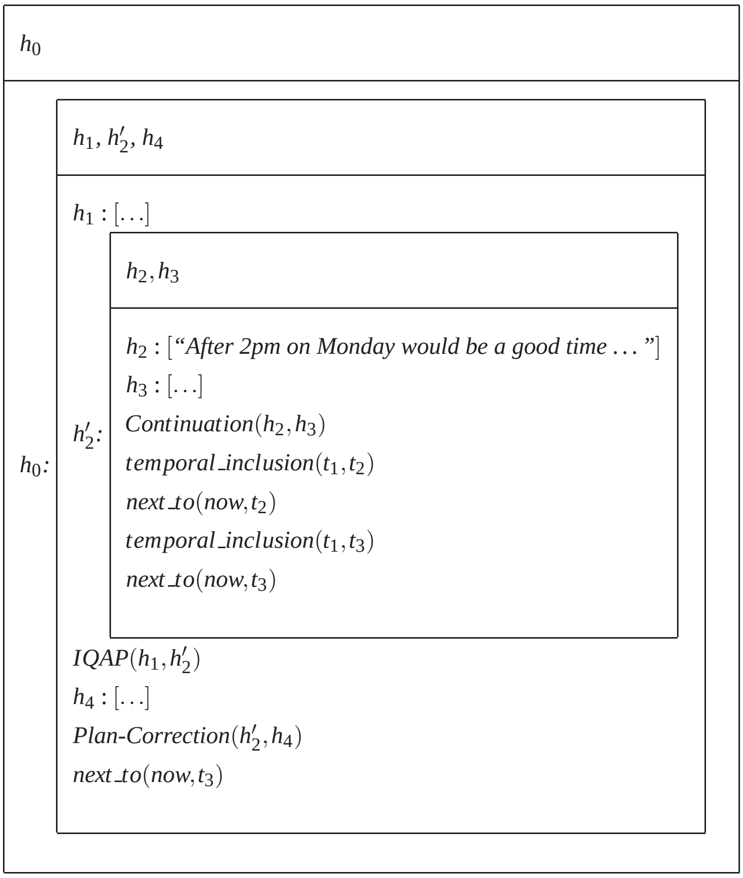



Figure 9.1: Discourse structure of the dialogue in (420).

that arising from the use of fragments (e.g., $h_2$ in (420) is resolved to something paraphraseable by "After 2pm on next Monday is a good time..."); the other kind of underspecification is that arising from the need to "bridge" (cf. (Clark 1977) and, in the context of SDRT, (Asher & Lascarides 1998*a*)) definites to the context (e.g., "Wednesday afternoon" in $h_3$ is resolved to be the next Wednesday afternoon after the time of utterance). The resolution tasks are interdependent: sometimes information about bridging relations is required to recognise the speech act that was performed, which in turn determines how the fragment is resolved.

The parts of the model that RUDI computes are denotations of temporal referents—defined as intervals on the calendar—that make the discourse structure true, given knowledge about when the conversation took place. RUDI also computes the purpose behind each utterance, insofar as the overall goal of finding a time to meet is concerned. So supposing the exchange (420) took place at 14:00 on December 10th 2002 (and that "couple" means exactly two), then the denotations are shown in (421):

(421) $t_1$ = interval from Dec. 16th to Dec. 29th;
$t_2$ = interval from 2pm to 11:59pm on Dec. 16th;
$t_3$ = Dec. 18th from 12:01pm to 6pm;
$t_4$ = interval from Dec. 16th to 22nd

Further, the goal of $h_4$ would be to find a time in the next two weeks that is not in the week from the 16th to the 22nd, i.e. to find a time in the second of those two weeks. Computing these properties of the models is necessary, because on occasion they affect logical form *construction*; in particular, the values of the rhetorical connections. For example, the *denotation* of *next week* (and in particular, whether it includes a time which satisfies the description *the 16th* or not), affects the rhetorical connection between the utterances in (422) and hence also its implicatures:

(422) a. A: Let's meet next week.
b. B: I'm away until the 16th.

If next week contains the 16th, then the discourse relation connecting the utterances entails that the purpose of (422-b) is to *elaborate* a way of achieving the goal behind (422-a); if not, then a different discourse relation which connects the utterances conveys the fact that (422-b) *rejects* the goal behind (422-a).

We chose the domain of scheduling dialogues for several reasons, the more practical of which is that we had access to a corpus of realistic dialogues from that area (the VM/REDWOODS data described in Sections 2.2.4 and 7.4.2) and to a grammar/parser that was capable of producing logical forms for them (the ERG/LKB combination). Less contingent reasons are that this is a domain where the requisite knowledge for discourse processing is relatively easy to model (knowledge about possible denotations of calendar terms and about reasoning with intervals) and, as we will see, possibilities for resolution of certain underspecifications are conventionally relatively restricted. Moreover, the domain-plan (in the sense of (Allen & Litman 1990)) in this domain can be modelled in a relatively simple way, by assuming that the joint goal of finding a time to meet at must be met by "zeroing in" on a time, i.e. by agreeing on successively smaller time intervals. Lastly, assuming that all utterances address this overall discourse goal, the goal behind individual utterances can easily be computed.[4] For example, we can assume that the goal of an utterance like "I am free on Monday" is to meet on the Monday denoted by the definite description. In spite of the simplicity of the domain, however, the corpus exhibits a wide variety of speech act types and anaphoric expressions such as fragments.

In the remainder of this chapter, we first describe the overall architecture of the system and how it relates to the general theory we laid out in the previous chapters (i.e., to what extent the system implements

[4]This non-digression assumption is of course unfounded in the general case, but can be justified in our simple restricted domain.

SDRT and our additional rules for fragments); in Section 9.3 we then show which fragmental speech acts RUDI is capable of handling and go through examples for each of them.

## 9.2 Architecture of the system

When interpreting a dialogue, RUDI proceeds incrementally, processing one utterance at a time, progressively updating an *information state* (IS) by accumulating constraints and trying to satisfy them. The type declaration for this information state is shown on the left-hand side in Figure 9.2; we will turn to the update process below when we discuss Figure 9.3.

$$\begin{bmatrix} \text{CONTEXT} & \begin{bmatrix} \text{LING} & \begin{bmatrix} \text{COND} & \textit{set} \\ \text{SAS} & \textit{set} \end{bmatrix} \\ \text{COG} & \begin{bmatrix} \text{SARGS} & \textit{set} \\ \text{TDLRS} & \textit{set} \end{bmatrix} \\ \text{HK} & \begin{bmatrix} \text{LAST-UTT} & \textit{label} \end{bmatrix} \end{bmatrix} \\ \text{CUR-UTT} & \begin{bmatrix} \text{LING} & \begin{bmatrix} \text{COND} & \textit{set} \\ \text{ANAPH} & \textit{set} \\ \text{B-RELS} & \textit{set} \\ \text{SAS} & \textit{set} \end{bmatrix} \\ \text{COG} & \begin{bmatrix} \text{SARG} & \textit{tdr} \\ \text{TDLRS} & \textit{set} \end{bmatrix} \\ \text{HK} & \begin{bmatrix} \text{ATTACH} & \textit{fifo} \end{bmatrix} \end{bmatrix} \end{bmatrix}
\qquad
\begin{bmatrix} \text{INDEX} & \text{xb2} \\ \text{UNIQUE} & + \\ \text{CONNECTED} & + \\ \text{LENGTH} & \begin{bmatrix} \text{UNIT} & \text{day} \\ \text{COUNT} & 1 \end{bmatrix} \\ \text{NAMED_INT} & \begin{bmatrix} \text{TYPE} & \text{dofw} \\ \text{NAME} & \text{Fri} \end{bmatrix} \\ \text{SE_LIST} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{START} & [2003, 01, 3, 00, 00] \\ \text{END} & [2003, 01, 3, 23, 59] \end{bmatrix} \\ \text{TAIL} & [\,] \end{bmatrix} \end{bmatrix}$$

Figure 9.2: RUDI's information state (left) and a TDL-representation (right)

In this IS, the field CONTEXT holds all relevant information about the discourse context, while CUR-UTT represents the current utterance with which the context is to be updated. The complex value for both fields consists of a linguistic part (LING) and a cognitive part (COG). The linguistic part consists of a set of EPs (i.e. an MRS-style representation) in COND, and a list of the inferred speech act types in SAs.[5] CUR-UTT additionally has fields to keep track of the anaphora and possible resolutions. The cognitive part represents information about the intended model of the structure (the intended denotations) and about goals, as explained above; we call this *cognitive* information because it indicates how the dialogue participants must understand the dialogue. This information is represented in the form of TDLRs, which are representations in a domain specific language, the temporal domain language (TDL). We encapsulate all knowledge about calendars, durations and intervals in this language; all domain specific reasoning takes places on these structures. Fig. 9.2 shows as an example a TDL representation for Friday 3rd January 2003. The start- and end-points of the interval are specified in a list-structure, so as to allow representation of non-connected intervals. The feature HK finally holds "housekeeping" information

[5]These two kinds of information are separated for technical reasons; taken together these fields can be understood as descriptions of SDRSs as introduced in the previous chapter.

needed during the update process, namely the handle of the last utterance in CONTEXT and about the attachment assumptions in CUR-UTT.

RUDI separates the discourse interpretation task into several modules, thus reflecting the modular nature of SDRT. Figure 9.3 shows a schematic representation of the information and control flow through the modules. As this figure already indicates, RUDI's algorithm is to a large extent a straightforward proceduralisation (and simplification) of SDRT;[6] we will point out in the following step-by-step description where it differs from the theory. We will also highlight where domain specific information is encoded and where the general discourse theory is implemented.

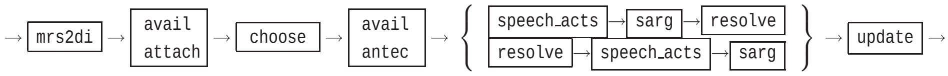



Figure 9.3: The algorithm

The input to the system consists of the MRS of the current utterance and a representation of the context (which is empty for the first utterance), combined in one IS as described above. In the first module, `mrs2di`, the MRS coming from the parser is postprocessed in two ways.[7] Firstly, underspecified bridging relations are added for all definite descriptions. For instance, if the definite in "let's meet on Monday" were represented as $monday(x)$, then this module would add something like $R(z,x) \wedge R =? \wedge z =?$ to the representation, to indicate that this temporal referent must be bridged to the context somehow.[8] To anticipate the further discussion a bit, a possible value for this underspecification would be $R = next, z = now$, constraining the Monday to be the next Monday from now.

As a second postprocessing task, and making use of our simplifying assumptions about the domain, `mrs2di` approximates domain specific knowledge of which events permit meeting at time $t$ and which don't, adding predicates that indicate whether the main time mentioned in the current utterance was a good time or a bad time. Hence we abstract over information which is irrelevant to the task at hand, such as, for example, whether the utterance was about going to the dentist or going on vacation; they both generate $bad\text{-}time(t)$. These postprocessing rules basically use the matrix verb and certain constructions to decide for *good-time* or *bad-time*; for fragments, this information of course is not available, and so there we leave this decision underspecified. This underspecification will be resolved once we decide for an attachment site, as we will soon see.

This kind of postprocessing rule simply encapsulates knowledge of actions in the domain. Others are derived logically "off line" (i.e. manually) in SDRT: for example, in this domain, SDRT validates the

[6] Although PROLOG, the computer language in which RUDI is realized, is often called a "declarative" language, programs written in this language have to contain non-declarative elements to achieve an efficient search through a potentially large space of possibilities; hence the declarative formulation of SDRT that was described in the previous chapter had to be proceduralised somewhat here. Also, in the interest of efficiency, some elements had to be approximated.

[7] Actually, we had to manually edit the MRSs in some cases, namely for fragmental questions where, as described in Chapter 7, only *wh*-fragments are parsed as interrogatives, since the grammar (at the time of writing) does not take interpunction into account.

[8] In the system these predicates are of course labelled, since this works on the level of descriptions (namely, MRSs). Indeed, $R =?$ and $z =?$ themselves aren't base-language expressions anyway.

inference that asking a question about a time *t* implicates that it's a *good-time*$(t)$ for the speaker to meet. The reasoning goes as follows. By default, a question attaches as *Q-Elab*. The semantics of this relation, namely that the question helps achieve a SARG of a prior utterance, is only met, given our additional assumption, if the utterance serves as a suggestion of a good time. This reasoning is 'hard-wired' into the post-processing rules, and thus we 'short-circuit' some SDRT inferences in the postprocessing of MRSs.

Following this domain specific preparation phase, the next three modules straightforwardly implement domain independent components of SDRT: in `avail attach` the constraint on available attachment sites (cf. Section 8.1.2); the preference for low attachment in `choose`, where an attachment site is picked; and the accessibility constraint in `avail antec`, where the potential antecedents for anaphoric material are determined. (As a slight domain specific deviation form the theory, `choose` also disambiguates the underspecified *good-time*/*bad-time* predicate for fragments, following the rule that a fragment that does not contain negation will predicate its main temporal referent in the same way as the utterance it is connected to. E.g., a fragment "2pm" following a '*good-time* question' is a *good-time* as well.)

The next block of three modules now implements the core of the theory; here the rhetorical relations are inferred and underspecification is removed. The two different orders of modules in the alternate paths shown in Figure 9.3 reflect the two ways underspecification can be resolved in SDRT.[9] The first path—the one shown on the top in the figure—is tried first; here all available information is used to infer speech acts first, and the semantics of the relations (and its influence on the speech act related goals (SARGs) and hence on the domain and discourse plans) is then used to resolve underspecification. If this route does not lead to a result (i.e., if no relations are inferred or no underspecification is resolved), then a different route is tried, where possible resolutions of underspecification are generated first and then used to infer relations. This simulates the effect of the MDC constraint, if you recall that this constraint ranks *fully specified* discourse structures according to the quality of the rhetorical relations in them. In RUDI we use a simplified ranking, according to which only those relations for which there is evidence (i.e., that can be inferred) are allowed; this is what is achieved by this order. As discussed in the previous chapter, the MDC is problematic for our phenomenon of fragments where there are potentially uncountably infinitely many fully specified discourse structures and so a ranking of all the possibilities is undecidable (let alone efficiently computable). We will describe in the next section how we work around this problem.

Figure 9.4 shows a 'magnified' view of the two modules `speech_acts`, which is responsible for inferring speech act types, and `resolve`, where the underspecification is resolved (and for this some elements of the intended model are computed, as described above). Both modules call external reasoners, `speech_acts` a theorem prover (CETP, (Schlangen & Lascarides 2002*a*)) for the non-monotonic logic used in SDRT, and `resolve` one that encodes the domain specific knowledge about calendar terms (the *TEmpOral DOmain Reasoner* TEODOR) and another one that resolves fragments, as described in the next section.

[9] Cf. the description in Section 8.1.


Figure 9.4: Detailed View of Two Core Modules

The domain specific rules are encapsulated in these external modules (as well as in `sarg`). To give an example of how the semantics of the rhetorical relations is made specific for this domain, let us look at the rule for *Q-Elab*. We defined this relation in the previous chapter as connecting an indicative $\alpha$ and a question $\beta$, where all answers to that question elaborate a plan to reach a SARG of $\alpha$. In this domain, we have fixed the goals to be times, and the plan is to narrow down these times, and so a *Q-Elab* will be a question that asks something about a time that is included in the goal of its antecedent $\alpha$. This way we can "hard-wire" reasoning about plans into the semantics of the relations and do not have to do costly plan-based reasoning.[10] (423) shows the rules involving this relation as they are implemented in the system, first in the notation familiar from the previous chapter and then—to illustrate that RUDI fairly straightforwardly implements the theory—in the format used by the theorem prover.

(423) a. `Inferring` *Q-Elab*

$(?(\alpha,\beta,\lambda) \wedge \beta :?) > \textit{Q-Elab}(\alpha,\beta,\lambda)$

`Consequences of` *Q-Elab*

$\textit{Q-Elab}(\alpha,\beta,\lambda) \rightarrow \textit{temp\_overlap}(t_\alpha, t_\beta)$

b. Inferring *Q-Elab*

```
(at(a,b) & int(b)) > qelab(a,b)
```

Consequences of *Q-Elab*

```
qelab(a,b) -> temp_inc(t(a), t(b))
```

As these rules show, inferring *Q-Elab* (which can be done on the basis of information about sentence moods) gives us information about temporal relations between referents, and thus might resolve bridging relations that were underspecified by the grammar.

For certain other speech act types, information about the denotation of temporal referents is required to infer whether they hold (this was shown with (422) above). The rules for two relations of this kind, *Plan-Elaboration* and *Plan-Correction*, are shown below.

(424) Inferring *Plan-Correction*

[10] This move of course will in general only be possible in such simple and restricted domains; in more complicated domain proper plan-based reasoning might have to be interfaced with these rules.

```
(at(a,b) & msg(b) & good_time(t(a)) &
bad_time(t(b)) & temp_inc(t(b),t(a)))
     > pcorr(a,b)
```

Inferring *Plan-Elaboration*

```
(at(a,b) & indic(b)) & bad_time(t(b))
& temp_overlap(t(a), t(b))
    > pelab(a,b)
```

This illustrated how the systems uses information about speech acts to resolve bridging relations, or, in other cases, information about bridging relations to resolve (infer) speech acts. In the next section we will show how information about speech acts resolves fragment.

## 9.3 Resolving Fragments using RUDI

RUDI implements rules for the resolution of several kinds of fragments, both *resolution-via-identity* and (simple cases of) *resolution-via-inference*. At the moment, it can deal only with NP- and PP-fragments. Before we go into the details of this part of the system, note that at all levels RUDI works on descriptions. In the previous chapter, we have explained how SDRT constrains the sets of described discourse-representations in such a way that, if one want to 'extract' BLFs, the pragmatically preferred one is produced. In the proceduralisation of this, we never leave the level of descriptions, but rather work the consequences of the inferred rhetorical relations directly into the descriptions. The result is a more specified description, which, if the rules are correct, only describes intended solutions, and entails the less specific description.

We now illustrate with an example of a *QAP*-fragment how RUDI resolves fragments, and then show how the same basic method is used in the resolution of other types of fragments as well. In the following, we assume that we have inferred that a certain utterance, a fragment, is connected via *IQAP* to a question. The rule (425) then triggers the application of the operation *q-id-resolve* ('resolve-via-identity, question variant').[11]

(425) $IQAP(\alpha,\beta,\gamma) \wedge frag(\beta) \rightarrow q\text{-}id\text{-}resolve(\beta)$

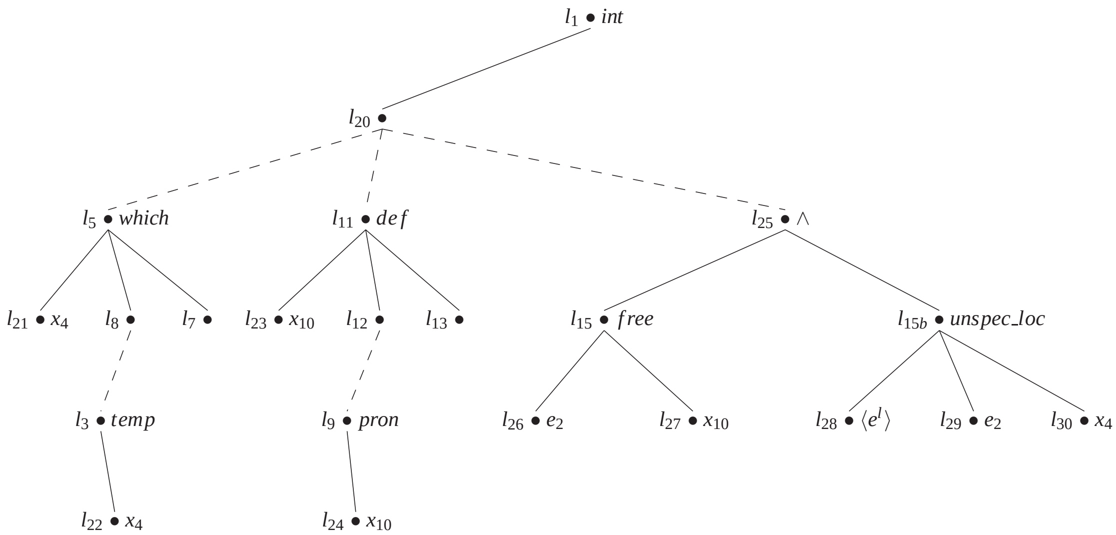
This operation consists of two steps: 'abstraction' over elements of $\alpha$ and 'application' of the resulting abstract to $\beta$ (and '$\beta$-reduction' of the result of that step).[12] We will show how this works with an

[11]Note that we do *not* model *res-via-inference* short-answers at all in RUDI; as explained above, their inference is potentially open-ended and will require information sources that we do not have available in the system. We will however model an (in this domain) restricted instance of *res-via-inference*, namely *Q-Elab*. More on this later.

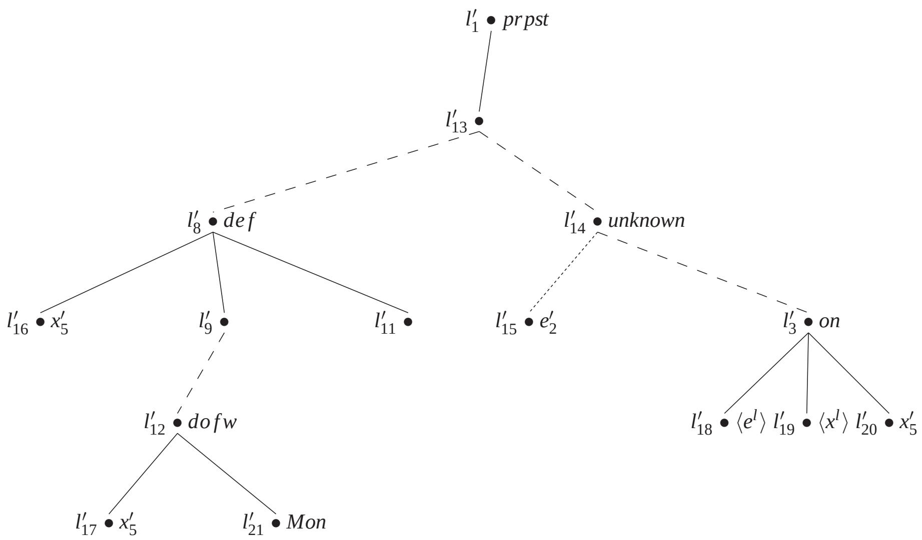
[12]We use scare-quotes on the names of the operations here, because even though the operations are similar to those of the $\lambda$-calculus, we do not want to claim that they have all the same properties. However, see (Bodirsky, Erk, Koller & Niehren 2001)

example. (426) gives the constraint-tree representation of the ULFs for a question and a reply.

(426) a. A: When are you free?



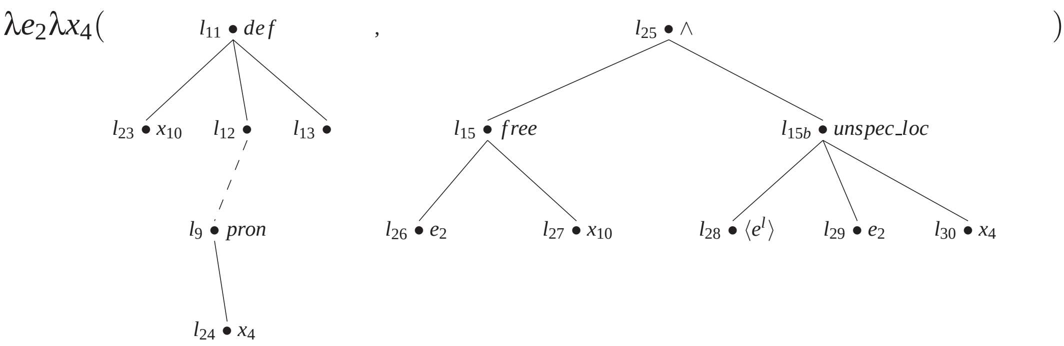

b. B: On Monday.




The step of abstraction over the *wh*-element in (426-a) leaves us with the following tree-fragments (the lambda-symbols represent that we record which variable was bound by the 'which'-quantifier, and also what the main event-variable is).

for a formalisation of β-reduction on underspecified representations, in the framework of CLLS.

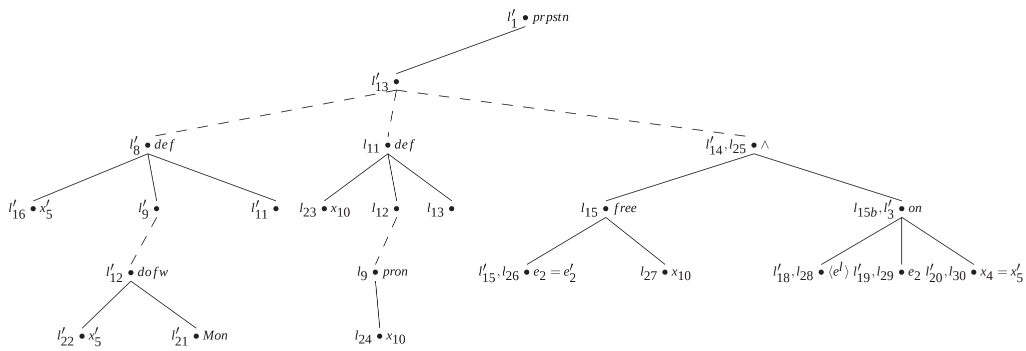
(427)




We describe what the operation of "application" does. First, the tree fragments from abstract and fragment are grouped together, then the argument(s) of *unknown* are equated with those bound by the lambda-operators (i.e. here $e_2 = e'_2$), *unknown_rel* is removed, and its handle is identified with one from the abstract (here, $l'_{14} = l_{25}$). Then the operations tests whether the constraints expressed by *unknown* are satisfied. In the example above, this is the case, since we can match *unspec_loc_rel* and *on_rel*, and so satisfy the constraints below $l'_{14}$ (that *on_rel* is subordinated by the handle of *unknown*).[13] Finally, the operation swaps any occuring first and second person pronouns, in case $\alpha$ and $\beta$ were uttered by different speakers. I.e., the representation of "you" is changed into that of "I", etc.

The result of this operation for (427) and (426-b) is shown below in (428). Note that for illustrative purposes, in this representation we have left in the original labels; the system of course assigns 'fresh' labels to the structure (otherwise, the discourse representation would not be well-formed).

(428)




Note that given the way this operation is defined, PP-fragments are blocked from answering questions where the *wh*-element is an NP, because in such cases the preposition relation in the fragment cannot be matched with one in the antecedent utterance. This alone already captures the syntactic parallelism for English, since as mentioned above, functional prepositions are represented in the ULF produced by the

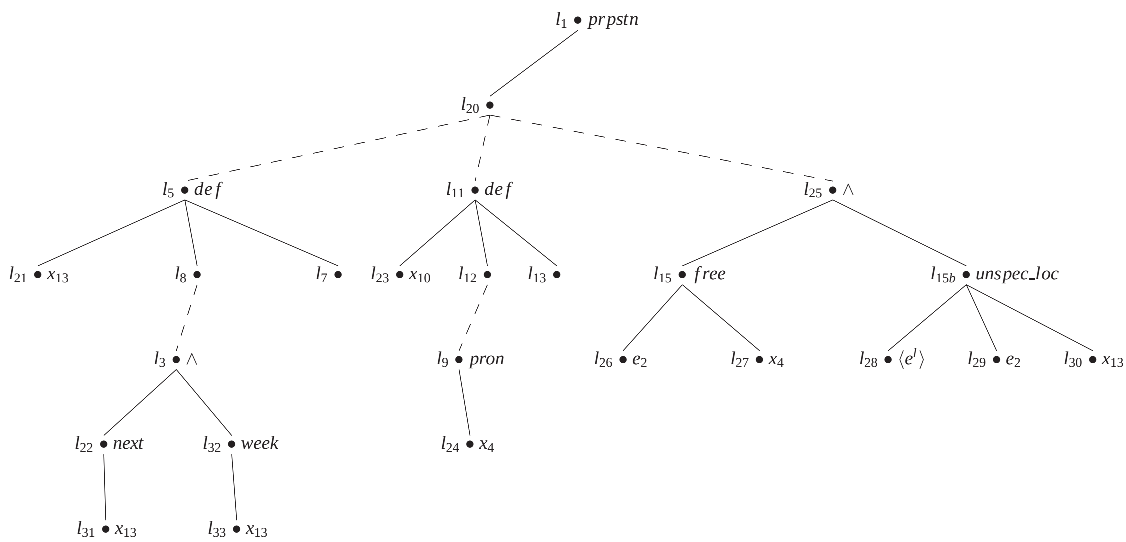
[13] I.e., the following identifications are made: $l'_3 = l_{15b}, l'_{18} = l_{28}, l'_{19} = l_{29}, l'_{20} = l_{30}$.

ERG, and for them a similar matching constraint must hold.[14]

The operation we described above is a special case of *resolution-via-identity*, because when resolving *QAP*-fragments we know which element of α to abstract over: the *wh*-element. This is generally not the case, and there can be several possibilities for abstraction. In the following we show how the generalised operation of abstraction resolves the plan-elaboration in (429-a). The representation of the fragment was already shown in (426-b), so we only show that of α here.

(429) a. (α) I am free next week.
(β) On Monday.

b.




Again, we have a rule that tells us how to resolve the fragment if we know that it is a plan-elaboration:

(430) $\textit{Plan-Elab}(\alpha,\beta,\gamma) \wedge \textit{frag}(\beta) \rightarrow \textit{id-resolve}(\beta)$

Given (429-b), there are two possible candidates for abstraction: the representation of "next week" and that of "I". So how do we stop this rule from resolving the fragment to something like "On Monday is free next week", i.e. how do we stop the fragment from 'replacing' the pronoun? Again, we realise indirectly what in the GBA is the syntactic parallelism constraint that requires categorial congruence between fragment-phrase and antecedent-element: the application of the abstract where the pronoun has been abstracted over is simply not well-formed. This can be shown as follows. Firstly, the application

[14] Note that for some other languages this will not be enough. For example, we have seen that German requires case concord between (some types of) questions and short-answers; this information would have to be represented additionally in a comparable system for German.

requires that $x_{13}$ and $x'_5$ are identified (otherwise *unspec_loc_rel* and *on_rel* cannot be made to match). However, since the quantifier that binds $x_{13}$ ("next week") is still in the abstract, this means that in the result this variable is bound twice, and this violates the well-formedness constraints on the result of the operation. Secondly, since the quantifier binding $x_4$ was removed in the abstraction step, there would be a free variable in the final representation (as argument of *free_rel*)—another violation of the well-formedness constraint.

Note that there are cases where ambiguity is wanted. For the fragment in (431) below, our system generates three possible resolutions, as indicated in brackets. In the absence of information about focus/background-partitioning (and knowledge about giving-actions), which would disambiguate the fragment, we think that this is correct.[15]

(431) A: Tom gave Jerry a cake.
B: No, Sandy.
(= *Sandy gave Jerry a cake.* or *Tom gave Sandy a cake.* or *Tom gave Sandy to Jerry.*)

Using the same operation, RUDI also resolves *Continuation*s like (432-b). (432) also gives a summary of the examples we have seen, together with the name of the operation that resolves the fragments of these types in the system.

| | | | | |
|---|---|---|---|---|
| (432) | a. | A: When are you free on Monday? — B: At 2pm. | *QAP* | *q-id-resolve* |
| | b. | A: I am free on Monday. On Tuesday as well. | *Cont* | *id-resolve* |
| | c. | A: I am free on Monday. At 2pm. | *P-Elab* | *id-resolve* |

We close this section with a look at one kind of *resolution-via-inference* fragments, namely *Q-Elab*s. We have shown in (423) above that *Q-Elab* is inferred in the system (just as in SDRT) simply on the basis of sentence modes. (Note that fragments in our approach are sentence-like, and so do have a sentence mode.) Hence, we can infer this relation for example in the mini-dialogue shown below.

(433) A: Let's meet on Monday.
B: (OK.) 2pm?

As discussed in detail in the previous chapters, this fragment cannot simply be resolved by identifying 'missing' bits with bits from the context. For example, "let's meet at 2pm?" is not a natural paraphrase of the intended content. Indeed, it is difficult to say what a natural paraphrase would be—much unlike for the fragments in (432) above. Here now RUDI takes advantage of the fact that it works on descriptions

[15] Unfortunately, our operation of abstraction and application does not allow us to keep this ambiguity in the description, i.e. to produce only one description that captures this relation between $\alpha$ and $\beta$ and constrains its result in this way, without having to produce separate descriptions. We leave this to future work, which possibly could start from the work of (Bodirsky et al. 2001).

of logical forms, and so underspecification doesn't *have* to be resolved, as long as the discourse goal can be reached: the system resolves such fragments with a generic predicate "*schedule*" that abstracts over the differences between "can we meet at 2pm?", "are you free at 2pm?", etc. The rules that triggers this resolution accordingly is as shown below, where *gen-resolve* stands for 'generic-resolution'.

(434) $\textit{Q-Elab}(\alpha,\beta,\gamma) \wedge frag(\beta) \rightarrow \textit{gen-resolve}(\beta)$

Observe that definites in such fragments will be bridged as a consequence of inferring *Q-Elab*, and so the goal of the fragment is also correctly specified; all this happens without resolving *unknown_rel*.

## 9.4 Summary

In this chapter we have presented an experimental implementation of the theory presented in the previous chapters. We have shown that for a suitably restricted domain the ideas developed in those chapters can be implemented in a relatively straightforward fashion. We have chosen this simple domain to avoid complications through orthogonal issues like reasoning about plans – issues which, as we should stress, of course are not trivial and important research topics on their own. By chosing such a simple domain (with the simple plan of 'zooming in' on a time), we were able to integrate this reasoning into the speech act semantics. In more complicated domains, the interface between reasoning about plans and reasoning about speech acts of course would have to be more complicated. The main idea realised in the system, however, would remain unchanged, namely that sometimes information flows from inferring speech acts to resolving underspecification, and sometimes from resolving underspecification to inferring speech acts.

The system is at the time of writing being extended (so far it only recognises a few speech act types) and made more robust as part of the Edinburgh-Stanford collaboration ROSIE (Robust Semantic IntErpretation), where it is to be used to prepare a corpus annotated with deep discourse semantic information. We expect that a publicly available version will be available as a result of that effort.

This concludes our brief description of the system, and indeed the presentation of our approach to fragments. In the next chapter, we will briefly summarise what we have done, and point out where future work could add.

# Chapter 10

# Conclusion and Outlook

In this chapter we briefly look back at what we have done over the last nine chapters; we also point out where further work is needed, and discuss interesting directions in which to expand this work.

## 10.1 Conclusion

*"In conclusion, it would be ridiculous to claim that this paper has more than scratched the surface of the problem of fragments." (Morgan 1973, p. 748)*

If nothing else, we hope we have at least deepened Morgan's (1973) "scratch on the surface of the problem", giving us a brief glimpse at what lies beneath the surface. We have presented a detailed theory of how a wide variety of fragments are interpreted in discourse. Past work has often concentrated only on certain kinds of fragments (mostly short-answers) and has placed the process of recovering their intended meaning in only one module of interpretation, be that syntax (Morgan 1973), semantics (e.g. (Krifka 1999)), or pragmatics (Carberry 1990, Barton 1990). Following (Ginzburg 1999*b*), we have claimed that information encompassing more than one level has to be used to model the licensing of fragments. In contrast to (Ginzburg 1999*b*), however, we have shown that it is possible, and desirable, to present a compositional account of fragments.

Here's a brief recap of the route we took. We began with discussing the information sources that are relevant when interpreting fragments, namely their syntax and the context they were uttered in, in particular the way they are **rhetorically connected** to (bits of) that context. With this as our starting hypothesis, we set out to survey what kinds of fragments one can find in a corpus of dialogue material, and from that survey we derived a comprehensive **taxonomy of fragments**. One important distinction we introduced there was that of **resolution-via-identity**-fragments versus **resolution-via-inference** ones, referring to whether all material needed for the resolution of the intended meaning of fragments has to be linguistically explicit or whether it does not. An interesting by-product of the corpus study was the finding that in typical dialogues about 10% of all utterances are fragmental; this result confirmed earlier studies.

In our attempt to devise a model of how this variety of fragments is to be accounted for, we then first turned to a phenomenon which at least on first view seems closely related, namely VP-ellipsis. Our brief review of the literature on that phenomenon revealed some differences, and gave us the structure for the review of the literature more specifically on fragments, by suggesting a classification of approaches according to the level of linguistic structure they work on: syntax, semantics or pragmatics. As a result of this review, we concluded that **there are compelling arguments both for and against purely syntactic and purely semantic approaches**. We then reviewed a more recent approach, that of Jonathan Ginzburg and colleagues, which addresses this problem by formulating a resolution-mechanism that is part of the grammar, and hence has simultaneous access to syntactic, contextual and semantic information. We argued against this approach, showing that **by allowing the grammar access to non-linguistically encoded information such as information about speech acts that were performed, one creates as many problems as one solves**.

This first part of the thesis prepared the ground for the development of our own approach, giving us clear desiderata: we wanted to explain the puzzling data, and we wanted to realize, as far as possible, a certain design feature, namely modularity. To that end, we formulated **a compositional semantics for fragments**, i.e. we formalised the meaning they have independently from the *use* that is being made with them. This we did with the help of the technical tool called **underspecified semantic representation**: we defined a formalism in which *descriptions* of logical forms can be expressed, in particular, descriptions of the information that is linguistically encoded in fragments. We then defined **a syntax for fragments**, in the framework of Head-Driven-Phrase-Structure-Grammar, and we implemented the analysis in an existing wide-coverage grammar and evaluated it on dialogue material. Finally, we showed how **our taxonomy of fragment types can be formalised in an existing theory of discourse interpretation** (SDRT), and how, **with one modification—namely giving discourse interpretation limited access to syntactic information—this theory can explain how our underspecified representations for fragments are resolved in context**.

In the following section we briefly discuss a possible objection to our approach, and in Section 10.3 we discuss further lines of research suggested by the work presented here.

## 10.2 A Possible Objection

• *"How does the need for access to syntactic information in the constraints on* G-Parallel *fit with the purported goal of modularity, and is this syntactic constraint not a bit* ad hoc *anyway?"*

In a paper from the 1970s (Morgan 1975), Jerry Morgan argued that one of the two following claims must be false (he didn't argue which one):[1] (a) pragmatics "is 'grammar-free' in that none of the principles involved depends on (or is sensitive to) the operation of syntactic rules or matters of superficial form", and (b) "the syntactico-semantic component is 'pragmatically-transparent' ", which we can understand for our present purposes to mean that the grammar does not have access to contextual information.

The data we have discussed here regarding fragments seem to support this observation: some information beyond logical form permeate the boundaries between these modules. But, as we have argued at length, if that is so, it seems advantageous to allow the pragmatic module access to syntactic information rather than *vice versa*. We will not repeat the argument here, and just remind the reader of the basic point that it seems more principled to just add one further information source to a module that already has access to several different sources (recall that the glue-logic in SDRT has limited access to the content of the discourse, cognitive states, lexical semantics etc.) rather than opening up a wide field of contextual information, which in turn needs to be supported by highly complex non-monotonic reasoning, to a module (the grammar) which otherwise could work very well with very constrained information.

[1] Both quotes are from (Morgan 1975, p.7).

To address the question of whether this is *ad hoc*, we would like to point out that there is some independent motivation for this strategy. As briefly discussed in Chapter 3, Kehler's (2002) approach to ellipsis similarly postulates that the process of establishing coherence can require accessing syntactic material—in his case, elliptical utterances that are connected with certain coherence relations are resolved via syntactic reconstruction, whereas others are resolved by semantic reconstruction.

## 10.3 Further Work

As is probably unavoidable even (or especially) in such a long work as a PhD thesis, there are a number of places in this thesis where details need to be filled in, as well as there being places which seem promising as starting points for further research. In the following we list some of both, in order of increasing complexity.

- In the syntactic and semantic analysis, we have concentrated on some paradigmatic types of fragments, mostly NP-, PP- and VP-fragments. We have not given an analysis for ADJ-fragments, free relatives ("I took a course there. — Taught by Montague himself?"), or fragments beginning with a conjunction ("and then Sandy."), to give just a few examples. However, we have no reason to believe that those types would pose fundamental problems to our approach.
- In the taxonomy of fragment types we have concentrated on relations involving propositions or questions, and have mostly ignored those involving imperatives. Again, this was done because of time constraints (and partially also because such relations did not occur in our corpus)—we don't think that there are any principled reasons why our model could not be extended to deal with these in a straightforward way.
- Our syntactic rule that only phrases can be fragments undergenerates. There seems to be a certain context, involving alternatives, in which non-phrasal projections are allowed, as shown in the following:

  (435) A: Do you have a cat or a dog?
  B: Cat.

- Even though we have occasionally used German examples, we have mostly concentrated on English. It would be interesting to conduct a proper cross-linguistic study of non-sentential utterances, comparing frequencies and behaviour. For example, (Morgan 1989) mentions Korean data that suggest that in this language one can choose whether to obey syntactic-parallelism or not, where this choice has pragmatic effect—an interesting challenge for our approach.
- At the more speculative end of the spectrum of possible topics for further work, we find the phe-

nomenon of ‘situationally controlled fragments’, i.e. fragments of the kind “one coffee, please”. We have explicitly excluded these from the domain of our study, but of course it would be a good validation for our theory if it could be extended to have something to say about these as well. Indeed, one might speculate that a theory of coherent *behaviour* in general (not just linguistic behaviour), could provide the right antecedents for the resolution of fragments like these.

However, we will not further speculate here, and so we close with the hope that what we do offer here is at least a solid basis for pursuing these and other related questions.

# Appendix A

# The Implemented Grammar of Fragments

## A.1 The Grammar Rules

```
;;
;; This files lists the modifications to the ERG that are needed
;; to get analyse fragment as described in this thesis.
;;
;; The last version of the ERG the modifications were tested on is
;; that from November 11th 2002.
;;


;; append to file: syntax.tdl
;; das, 14/02/02 -- 05/02/03, Rules for Fragments
;;

;; the mother of all fragment construction types: frg.
;; Its constraints give the specifications common to all fragment-signs:
;; -- the resulting sign is syntactically a sentence (s_cat_fin_unspec),
;;    which is a verb with its valency requirements fulfilled.
;; -- MC + forbidds selection by other verbs, eg rules out ‘Peter said S’.
;;    this has to be removed to get the ‘embedded-fragments’ version
```

```
;;    of the grammar.
;; -- there are no gaps and other non-local features.
;; -- the semantic contribution of all types of fragment construction
;;    is a list containing of a message relation (either prpstn_rel or
;;    int_rel, to be specified by the frg_mood-hierarchy) and the unknown_rel;
;;    also we already know that there will be one qeq relation, namely the
;;    one between the message-rel and the unknown-rel.
;; -- finally, two new top-level features are declared,
;;    FRAGH-DTR & FRAGNH-DTR. They provide uniform access to the sign of the
;;    fragment argument (eg the NP in an NP-fragment), regardless of whether
;;    the fragment contains a modification or not (‘‘probably Sandy’’).

frg := non_headed_phrase & rule &
  [
    SYNSEM synsem_min &
           [ LOCAL [ CAT s_cat_fin_unspec &
                         [ HEAD verb & [ AUX -,
                                         INV -,
                                         MOD < > ],
                           VAL [ SUBJ <>,
                                 SPR  <>,
                                 COMPS <>,
                                 SPEC <>]
                         ],
                     CONJ cnil,
                     KEYS.MESSAGE <! #msg !> ],
             NON-LOCAL non-local_none
            ],
    C-CONT [ TOP #ctop,
             INDEX #event,
             E-INDEX #event,
             LISZT [ LIST < message & #msg & [ HANDEL #ctop ],
                            unknown_rel & [ EVENT #event ]
                          > ],
             H-CONS [ LIST < qeq, ... > ]
           ],
    FRAGH-DTR sign & [SYNSEM [ LOCAL.CONJ cnil,
                               NON-LOCAL.SLASH 0-dlist]],
```

```
    FRAGNH-DTR sign
  ].



;;
;; Dimension ‘type of fragment construction’ (VP-ADV yes/no)
;;

frg_md :< frg.


;; Non-modified fragments.
;; -- they are unary-phrases (only one argument), and
;; -- this argument is the fragment-head.
;; -- since there is no relation intervening, the messag-rel can be
;;    connected to the unknown-rel.

fragment_nm := frg_md & unary_phrase &
  [
    FRAGH-DTR #fragh,
    ARGS < #fragh >,
    C-CONT [ LISZT [ LIST < message & [ SOA #soa ],
                           unknown_rel & [HANDEL #ukhandle] > ],
             H-CONS [ LIST < qeq & [ SC-ARG #soa,
                                     OUTSCPD #ukhandle ], ... > ]
           ]
  ].


;; Modified Fragments:
;; -- They are binary phrases, where
;; -- the first argument (the adverb) is the fragment-non-head daughter,
;;    and the second is the (fragment-)head.
;; -- The relation(s) introduced by the adverb are positioned ‘between’
;;    the message-rel and the unknown-rel in the scope-relation-graph.

fragment_m := frg_md & basic_binary_phrase &
```

```
  [
    FRAGH-DTR #fragh,
    FRAGNH-DTR #fragnh &
         [ SYNSEM scopal_vp_aux_adverb_synsem &
              [ LOCAL [
                        CAT.HEAD.MOD.FIRST.LOCAL.CONT.TOP #ukhandle,
                        CONT.TOP #advtop ] ]
         ],
    ARGS < #fragnh, #fragh>,
    C-CONT [
             LISZT [ LIST < message & [ SOA #soa ],
                            unknown_rel & [ HANDEL #ukhandle ] > ],
             H-CONS [ LIST < qeq & [ SC-ARG #soa,
                                     OUTSCPD #advtop ],
                             ... > ]
           ]
  ].


fragment_mscop := fragment_m &
  [ FRAGNH-DTR.SYNSEM scopal_vp_aux_adverb_synsem ].

fragment_mneg := fragment_m &
  [ FRAGNH-DTR.SYNSEM.LOCAL [ CAT.HEAD negadv,
                              CONJ cnil ]
                            ].



;;
;; Dimension ‘sentence mood’
;;

frg_mood :< frg.

;; declarative fragments are those with an empty QUE-list, whereas
;; interrogative fragments have a non-empty QUE-list. This restriction
;; means that only wh-phrases (which have such an non-empty QUE-list)
;; are parsed as declaratives. See remarks in thesis.
```

```
frag_decl := frg_mood &
  [ C-CONT.LISZT.LIST < prpstn_rel, ...>,
   FRAGH-DTR.SYNSEM.NON-LOCAL  [QUE 0-dlist,
                                REL 0-dlist]].
frag_int := frg_mood &
  [ C-CONT.LISZT.LIST < int_rel, ...>,
   FRAGH-DTR.SYNSEM.NON-LOCAL  [QUE 1-dlist,
                                REL 0-dlist]].




;;
;; Dimension ‘type of fragment phrase’
;;
;; This is the type with the most sub-types, one for each (syntactic) type
;; of fragment (eg NP- or VP-fragment).

frg_arg :< frg.



;; NP-fragments have as argument the semantic index of the fragment-phrase
;; (the fragment-head).

np_fragment := frg_arg &
  [
   FRAGH-DTR.SYNSEM [LOCAL [ CONT.INDEX #daughter_ind,
                              CAT.HEAD noun,
                              CAT.VAL [COMPS <>,
                                       SPR *olist*]]],
   C-CONT [ LISZT [ LIST < message,
                            unknown_rel & [
                                           ARG #daughter_ind] > ]
                 ]
   ].
```

```
;; Supertype for PPs (both lexical and functional):
;; they have in common that the argument phrase (the PP) is outscoped by
;; the unknown-rel.

pp_fragment := frg_arg &
  [
   FRAGH-DTR.SYNSEM [LOCAL [  CONT.TOP #ppltop,
                              CAT.HEAD prep,
                              CAT.VAL [COMPS <>]]],
   C-CONT [
             LISZT [ LIST < message,
                            unknown_rel & [HANDEL #ukhandle ] > ],
             H-CONS [ LIST < qeq,
                             geq & [ SC-ARG #ukhandle,
                                     OUTSCPD #ppltop ] > ]
                 ]
   ].


;; lexical PP-fragments
;; --- selected via KEYS.KEY, which makes the relation of the preposition
;;     accessible. independent_rel is the type of prepositions that are not
''     verb-particles.

pp_l_fragment := pp_fragment &
  [
   FRAGH-DTR.SYNSEM.LOCAL [
                              KEYS.KEY independent_rel
                          ]
  ].

;; functional PP-fragments
;; --- also selected via KEYS.KEY.
;;     the type unsurprisingly is called 'selected_rel'.

pp_f_fragment := pp_fragment &
  [
   FRAGH-DTR.SYNSEM.LOCAL [
```

```
                              KEYS.KEY selected_rel &
                                       [ ARG3 #daughter_ind ]
                          ],
   C-CONT [ LISZT [ LIST <  message,
                            unknown_rel & [ ARG #daughter_ind ] >
                  ]
          ]
   ].




;; The rules for VP-fragments.


;; VP[inf], (eg. in  ‘‘What did he force you to do? --- To kick Sandy’’)
;; -- the top-handle of the fragment phrase becomes the argument of the
;;    unknown-rel.

vp_inf_fragment := frg_arg &
  [
   FRAGH-DTR.SYNSEM [ LOCAL [ CONT.TOP #daughter_top,
                              CAT [ HEAD comp & [VFORM inf],
                                    VAL.COMPS *olist*]],
                      NON-LOCAL non-local_none],
   C-CONT [ LISZT [ LIST <  message,
                            unknown_rel & [
                                           ARG #daughter_top] > ]
                 ]
   ].


;; VP[bse]-fragments (eg. in ‘‘What did he make you do? --- Kick Sandy.)
;;
;; -- Similarly, the top-handle of the fragment phrase becomes the
;;    argument of the unknown-rel.
;; Only difference to previous rule is specification of that argument-phrase.
```

```
vp_bse_fragment := frg_arg &
  [
   FRAGH-DTR.SYNSEM [LOCAL bse_verb &
                         [ CONT.TOP #daughter_top,
                           CAT vp_cat & [ HEAD.VFORM bse_only,
                                          VAL [COMPS *olist*] ] ] ],
   C-CONT [ LISZT [ LIST < prpstn_rel,
                         unknown_rel & [
                                         ARG #arghandle] > ],
            H-CONS [ LIST < qeq,
                          qeq & [ SC-ARG #arghandle,
                                  OUTSCPD #daughter_top ] > ]
          ]
   ].



;; S[comp]-rule, for things like ‘what did Peter say? -- that Sandy sleeps.’’
;;
;; -- Again daugther-top is identified with the argument to unknown-rel.

s_comp_fragment := frg_arg &
  [
   FRAGH-DTR.SYNSEM [LOCAL [ CONT.TOP #daughter_top,
                           CAT.HEAD comp,
                           CAT.HEAD.VFORM fin]],
   C-CONT [ LISZT [ LIST < prpstn_rel,
                         unknown_rel & [
                                        ARG #daughter_top] > ]
                 ]
   ].


;;
;; The rules inhert from one instance of the three dimensions, respectively.
;; (but note that there are different versions of modification)

np_nm_d_fragment := np_fragment & frag_decl & fragment_nm.
```

```
np_mneg_d_fragment := np_fragment & frag_decl & fragment_mneg.
np_mscop_d_fragment := np_fragment & frag_decl & fragment_mscop.


pp_l_nm_d_fragment := pp_l_fragment & frag_decl & fragment_nm.
pp_l_mneg_d_fragment := pp_l_fragment & frag_decl & fragment_mneg.
pp_l_mscop_d_fragment := pp_l_fragment & frag_decl & fragment_mscop.
pp_f_nm_d_fragment := pp_f_fragment & frag_decl & fragment_nm.
pp_f_mneg_d_fragment := pp_f_fragment & frag_decl & fragment_mneg.
pp_f_mscop_d_fragment := pp_f_fragment & frag_decl & fragment_mscop.


;; note: we do not allow modified interrogatives.
;; something like ‘possibly who?’ only has an echo-question reading.
;; recall that we only allow wh-phrases as interrogatives, and
;; so do not handle the (legit) fragment ‘‘not Peter?’’ for
;; other reasons.
np_nm_i_fragment := np_fragment & frag_int & fragment_nm.
pp_l_nm_i_fragment := pp_l_fragment & frag_int & fragment_nm.
pp_f_nm_i_fragment := pp_f_fragment & frag_int & fragment_nm.


vp_inf_nm_d_fragment := vp_inf_fragment & frag_decl & fragment_nm.
vp_inf_mneg_d_fragment := vp_inf_fragment & frag_decl & fragment_mneg.
vp_inf_mscop_d_fragment := vp_inf_fragment & frag_decl & fragment_mscop.

vp_bse_nm_d_fragment := vp_bse_fragment & frag_decl & fragment_nm.
vp_bse_mneg_d_fragment := vp_bse_fragment & frag_decl & fragment_mneg.
vp_bse_mscop_d_fragment := vp_bse_fragment & frag_decl & fragment_mscop.

s_comp_nm_d_fragment := s_comp_fragment & frag_decl & fragment_nm.
s_comp_mneg_d_fragment := s_comp_fragment & frag_decl & fragment_mneg.
s_comp_mscop_d_fragment := s_comp_fragment & frag_decl & fragment_mscop.



;; append to file: constructions.tdl
;; das, 14/02/02 -- 05/02/03, Rules for Fragments
```

```
;;


;; the actual grammar rules:

np_nm_d_fragment_r := np_nm_d_fragment &
  [ RULE-NAME 'np_nm_d_fragment_r ].
np_mneg_d_fragment_r := np_mneg_d_fragment &
  [ RULE-NAME 'np_mneg_d_fragment_r ].
np_mscop_d_fragment_r := np_mscop_d_fragment &
  [ RULE-NAME 'np_mscop_d_fragment_r ].

pp_l_nm_d_fragment_r := pp_l_nm_d_fragment &
  [ RULE-NAME 'pp_l_nm_d_fragment_r ].
pp_l_mneg_d_fragment_r := pp_l_mneg_d_fragment &
  [ RULE-NAME 'pp_l_mneg_d_fragment_r ].
pp_l_mscop_d_fragment_r := pp_l_mscop_d_fragment &
  [ RULE-NAME 'pp_l_mscop_d_fragment_r ].
pp_f_nm_d_fragment_r := pp_f_nm_d_fragment &
  [ RULE-NAME 'pp_f_nm_d_fragment_r ].
pp_f_mneg_d_fragment_r := pp_f_mneg_d_fragment &
  [ RULE-NAME 'pp_f_mneg_d_fragment_r ].
pp_f_mscop_d_fragment_r := pp_f_mscop_d_fragment &
  [ RULE-NAME 'pp_f_mscop_d_fragment_r ].

np_nm_i_fragment_r := np_nm_i_fragment &
  [ RULE-NAME 'np_nm_i_fragment_r ].
pp_l_nm_i_fragment_r := pp_l_nm_i_fragment &
  [ RULE-NAME 'pp_l_nm_i_fragment_r ].
pp_f_nm_i_fragment_r := pp_f_nm_i_fragment &
  [ RULE-NAME 'pp_f_nm_i_fragment_r ].

vp_inf_nm_d_fragment_r := vp_inf_nm_d_fragment &
  [ RULE-NAME 'vp_inf_nm_d_fragment_r ].
vp_inf_mneg_d_fragment_r := vp_inf_mneg_d_fragment &
  [ RULE-NAME 'vp_inf_mneg_d_fragment_r ].
vp_inf_mscop_d_fragment_r := vp_inf_mscop_d_fragment &
  [ RULE-NAME 'vp_inf_mscop_d_fragment_r ].
```

```
vp_bse_nm_d_fragment_r := vp_bse_nm_d_fragment &
  [ RULE-NAME 'vp_bse_nm_d_fragment_r ].
vp_bse_mneg_d_fragment_r := vp_bse_mneg_d_fragment &
  [ RULE-NAME 'vp_bse_mneg_d_fragment_r ].
vp_bse_mscop_d_fragment_r := vp_bse_mscop_d_fragment &
  [ RULE-NAME 'vp_bse_mscop_d_fragment_r ].


s_comp_nm_d_fragment_r := s_comp_nm_d_fragment &
  [ RULE-NAME 's_comp_nm_d_fragment_r ].
s_comp_mneg_d_fragment_r := s_comp_mneg_d_fragment &
  [ RULE-NAME 's_comp_mneg_d_fragment_r ].
s_comp_mscop_d_fragment_r := s_comp_mscop_d_fragment &
  [ RULE-NAME 's_comp_mscop_d_fragment_r ].




;; append to file: fundamentals.tdl
;; das, 14/02/02 -- 05/02/03, Rules for Fragments
;;

;; added type `hcons_pr', which is a new supertype of qeq
;; (the ERG-native outscopes-variant) and geq, which we introduce
;; and which stands for `greater or equal', i.e. is `normal'
;; outscopes.

hcons_pr := scp_pr &
  [ OUTSCPD handle ].

qeq := hcons_pr.

geq := hcons_pr.



;; we also have to declare unknown_rel:
```

```
unknown_rel :< event_arg_rel.
```

## A.2 The Test-Suite for Fragments

```
;;;
;;; NP_frag
;;;
Peter.
The man in the garden.
The dog.
Dogs.
one to three.
;;;
;;; PP_frag
;;;
On Sandy.
With Sandy.
Of Sandy.
;;;
;;; VP_frag
;;;
kick Sandy.
to kick Sandy.
that Peter kicks Sandy.
*kicks Sandy.
*sounds good.
*to kicks Sandy.
*that Peter kick Sandy.
;;;
;;; ADV_frag
;;;
slowly.
possibly.
maybe.
```

```
;;;
;;; ADJ_frag
;;;
cool.
good.
;;;
;;; Int_frag
;;;
whose father?
who?
which party?
when?
where?
why?
how?
how many?
;;;
;;; Slashed_frag
;;;
*of
*the father of
*kick
*to give
*that Peter gives
*a picture of.
;;;
;;; Modified_frag
;;;
probably kick sandy.
probably to kick sandy.
probably probably to kick Sandy.
probably that Peter kicked Sandy.
;;;
;;; Conj_frag
;;;
and Sandy.
and to kick Sandy.
or maybe Peter.
```

```
and then to Paris.
and died.
```

# Appendix B

# Publications

Some of the results reported here have been published before. We give a list of the relevant publications here, and refer the interested reader to our web-page,[1] from where electronic copies of these papers can be obtained.


- Schlangen, D. and A. Lascarides, [2003] The Interpretation of Non-Sentential Utterances in Dialogue, to appear in *Proceedings of the 4th SIGdial Workshop on Discourse and Dialogue*, Sapporo, Japan, July 2003
- Schlangen, D. and A. Lascarides, [2003] A Compositional and Constraint-Based Approach to Non-Sentential Utterances, to appear in *Proceedings of the 10th International Conference on Head-Driven Phrase Structure Grammar*, Michigan, USA, July 2003
- Schlangen, D., A. Lascarides and A. Copestake, [in press] Resolving Underspecification using Discourse Information, in Hannes Rieser and Peter Kühnlein (eds.), *Formal Semantics and Pragmatics of Dialogue*, John Benjamins
- Schlangen, D. and A. Lascarides, [2002] Resolving Fragments using Discourse Information, *Proceedings of the 6th International Workshop on Formal Semantics and Pragmatics of Dialogue (EDILOG 2002)*, pp.161–168, Edinburgh, UK, September 2002.
- Schlangen, D., [2002] A Compositional Approach to Short-Answers, in *Proceedings of the Student Research Workshop at the 40th ACL*, pp.54–59, Philadelphia USA, July 2002.
- Schlangen, D. and A. Lascarides, [2002] CETP: An automated theorem prover for a fragment of common sense entailment, May 2002, Edinburgh Informatics Report Series, EDI-INF-RR-0119
- Schlangen, D., A. Lascarides and A. Copestake, [2001] Resolving Underspecification using Discourse Information, *Proceedings of the 5th International Workshop on Formal Semantics and


[1] `http://www.cogsci.ed.ac.uk/~das`

*Pragmatics of Dialogue (BI-DIALOG 2001)*, Bielefeld Germany, June 2001.

# Bibliography

Allen, J. F. & Litman, D. J. (1990), Discourse processing and common sense plans, *in* Cohen, Morgan & Pollack (1990).

Alshawi, H., ed. (1992), *The CORE Language Engine*, MIT Press, Cambridge, Massachusetts.

Anderson, A., Bader, M., Bard, E., Boyle, E., Doherty, G. M., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., Thompson, H. S. & Weinert, R. (1991), 'The HCRC maptask corpus', *Language and Speech* **34**, 351–366.

Asher, N. (1993), *Reference to Abstract Objects in Discourse*, Studies in Linguistics and Philosophy, Kluwer Academic Publisher, Dordrecht.

Asher, N. (1995), Commonsense entailment: A conditional logic for some generics, *in* G. Crocco, L. F. nas del Cerro & A. Herzig, eds, 'Conditionals: From Philosophy to Computer Science', Oxford University Press, Oxford, pp. 103–145.

Asher, N. (1998), Varieties of discourse structure in dialogue, *in* 'Twendial (Proceedings of the Second International Workshop on Dialogue)', Twente, the Netherlands.

Asher, N. & Fernando, T. (1999), Labeled representations, underspecification and disambiguation, *in* H. Bunt & R. Muskens, eds, 'Computing Meaning', Kluwer Academic Press, Dordrecht, pp. 33–55.

Asher, N., Hardt, D. & Busquets, J. (2001), 'Discourse parallelism, ellipsis and ambiguity', *Journal of Semantics* **18**(1).

Asher, N. & Lascarides, A. (1995), 'Lexical disambiguation in a discourse context', *Journal of Semantics* **12**(1), 69–108.

Asher, N. & Lascarides, A. (1998*a*), 'Bridging', *Journal of Semantics* **15**(1), 83–113.

Asher, N. & Lascarides, A. (1998*b*), 'The semantics and pragmatics of presupposition', *Journal of Semantics* **15**, 239–299.

Asher, N. & Lascarides, A. (1999), Cognitive states, discourse structure and the content of dialogue, *in* 'Proceedings to Amstelogue 1999', Amsterdam.

Asher, N. & Lascarides, A. (2001), 'Indirect speech acts', *Synthese* **128**(1), 183–228.

Asher, N. & Lascarides, A. (2003), *Logics of Conversation*, Cambridge University Press.

Asher, N. & Morreau, M. (1991), Commonsense entailment, *in* J. Mylopoulos & R. Reiter, eds, 'Proceedings of the Twelfth International Joint Conference on Artificial Intelligence', Morgan Kaufmann, Los Altos, California.

Aston, G. & Burnard, L. (1998), *The BNC Handbook*, Edinburgh University Press, Edinburgh, UK.

Austin, J. (1962), *How To Do Things With Words*, Clarendon Press, Oxford.

Barton, E. L. (1990), *Nonsentential Constituents*, John Benjamins, Amsterdam / Philadelphia.

Blackburn, P. & Bos, J. (2000), A course in computational semantics. To be published by CSLI publications.
*http://www.comsem.org

Bodirsky, M., Erk, K., Koller, A. & Niehren, J. (2001), Underspecified beta reduction, *in* 'Proceedings of the 39th ACL', Toulouse, pp. 74–81.

Bolinger, D. (1972), *That's That*, Mouton, The Hague, Paris.

Bos, J. (1994), Presupposition and vp-ellipsis, *in* 'Proceedings of COLING-1994', Kyoto, Japan.

Bos, J. (1996), Predicate logic unplugged, Verbmobil Report 103, Universität des Saarlandes.

Bouma, G., Malouf, R. & Sag, I. A. (2001), 'Satisfying constraints on extraction and adjunction', *Natural Language and Linguistic Theory* **19**(1), 1–65.

Branigan, H., Pickering, M. & Cleland, A. (2000), 'Syntactic coordination in dialogue', *Cognition* (75), B13–B25.

Bresnan, J. (1994), 'Locative inversion and the architecture of universal grammar', *Language* **70**(1), 72–131.

Bresnan, J. (2001), *Lexical-functional Syntax*, Blackwell.

Brewka, G., Dix, J. & Konolidge, K. (1997), *Nonmonotonic Reasoning: An Overview*, CSLI Lecture Notes Nr. 73, CSLI Publications, Stanford.

Bunt, H. (1994), 'Context and dialogue control', *Think Quarterly* (3), 19–31.

Büring, D. (1999), Topic, *in* P. Bosch & R. van der Sandt, eds, 'Focus: Linguistic, Cognitive and Computational Perspectives', Studies in Natural Language Processing, Cambridge University Press, Cambridge, UK, pp. 142–165.

Carberry, S. (1989), 'A pragmatics-based approach to ellipsis resolution', *Computational Linguistics* **15**(2), 75–96.

Carberry, S. (1990), *Plan Recognition in Natural Language Dialogue*, MIT Press, Cambridge, Massachusetts.

Carbonell, J. G. (1983), Discourse pragmatics and ellipsis resolution in task-oriented natural language interfaces, *in* 'Proceedings of the 21st annual meeting for the Association for Computational Linguistics', Boston, USA, pp. 164–168.

Carlson, L. (1983), *Dialogue Games: An Approach to Discourse Analysis*, Kluwer Academic Publishers.

Clark, H. (1977), Bridging, *in* P. N. Johnson-Laird & P. C. Wason, eds, 'Thinking: Readings in Computer Science', Cambridge University Press, Cambridge, UK, pp. 411–420.

Clark, H. H. (1996), *Using Language*, Cambridge University Press, Cambridge.

Cohen, P. R., Morgan, J. & Pollack, M. E., eds (1990), *Intentions in Communication*, MIT Press, Cambridge, Mass.

Cole, P. & Morgan, J. L., eds (1975), Academic Press, New York.

Cooper, R. (1983), *Quantification and Syntactic Theory*, Reidel, Dordrecht.

Cooper, R., Engdahl, E., Larsson, S. & Ericsson, S. (2000), Accommodating questions and the nature of QUD, *in* 'Proceedings of Götalog 2000', Gotheburg, pp. 57–62.

Copestake, A. (2002), *Implementing Typed Feature Structure Grammars*, CSLI Publications, Stanford, USA.

Copestake, A. & Flickinger, D. (2000), An open source grammar development environment and broad-coverage english grammar using HPSG, *in* 'Proceedings of the 2nd Linguistic Resources and Evaluation Conference', Athens, Greece, pp. 591–600.

Copestake, A., Flickinger, D., Sag, I. & Pollard, C. (1999), Minimal recursion semantics: An introduction. Stanford University, Stanford, CA.
*http://www-csli.stanford.edu/∼aac/papers/newmrs.ps

Copestake, A., Lascarides, A. & Flickinger, D. (2001), An algebra for semantic construction in constraint-based grammars, *in* 'Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics (ACL/EACL 2001)', Tolouse, France, pp. 132–139.

Core, M. G. & Allen, J. F. (1997), Coding dialogs with the damsl annotation scheme, *in* 'Working Notes of AAAI Fall Symposium on Communicative Action in Humans and Machines', Boston, MA, pp. 28–35.

Covington, M. A. (1994), GULP 3.1: An extension of prolog for unification-based grammar, Technical Report AI-1994-06, University of Georgia.

Crouch, R. (1995), Ellipsis and quantification: A substitutional approach, *in* 'Proceedings of the Seventh Conference of the European Chapter of the Association for Computational Linguistics (EACL-95)', EACL, Dublin, Ireland, pp. 229–236.

Dalrymple, M., ed. (1999), *Semantics and Syntax in Lexical Functional Grammar: The Resource Logic Approach*, The MIT Press, Cambridge, Massachusetts.

Dalrymple, M., Shieber, S. & Pereira, F. (1991), 'Ellipsis and higher order unification', *Linguistics and Philosophy* **14**, 399–452.

Danlos, L. (1999), Event coreference between two sentences, *in* 'Proceedings of the Third International Workshop on Computational Semantics', Tilburg, The Netherlands.

de Saussure, F. (1916), *Cours de linguistique gnrale*, Payot, Paris.

Deemter, K. v. & Peters, S., eds (1996), *Semantic Ambiguity and Underspecification*, CSLI, Stanford.

Egg, M., Koller, A. & Niehren, J. (2001), 'The constraint language for lambda structures', *Journal of Logic, Language, and Information* (10), 457–485.

Egg, M., Niehren, J., Ruhrberg, P. & Xu, F. (1998), Constraints over lambda-structures in semantic underspecification, *in* 'Proceedings of the 17th International Conference on Computational Linguistics and 36th Annual Meeting of the Association for Computational Linguistics (COLING/ACL'98)', Montreal, Canada, pp. 353–359.

Engdahl, E., Larsson, S. & Ericsson, S. (2000), Focus-ground articulation and parallelism in a dynamic model of dialogue, Trindi Deliverable D4.1, University of Gothenburg.

Engdahl, E. & Vallduví, E. (1996), Information packaging in HPSG, *in* C. Grover & E. Vallduví, eds, 'Studies in HPSG', Vol. 12 of *Edinburgh Working Papers in Cognitive Science*, University of Edinburgh, pp. 1–31.

Erk, K., Koller, A. & Niehren, J. (2002), 'Processing underspecified semantic descriptions in the constraint language for lambda structures', *Journal of Language and Computation* . to appear.

Fernández, R. & Ginzburg, J. (2002), Non-sentential utterances in dialogue: A corpus-based study, *in* K. Jokinen & S. McRoy, eds, 'Proceedings of the Third SIGdial Workshop on Discourse and Dialogue', ACL Special Interest Group on Dialog, Philadelphia, USA, pp. 15–26.

Fiengo, R. & May, R. (1994), *Indices and Identity*, MIT Press, Cambridge, Mass.

Gardent, C. (1999), Unifying parallels, *in* 'Proceedings of the 37th Meeting of the ACL'.

Gardent, C. & Kohlhase, M. (1997), Computing parallelism in discourse, *in* 'Proceedings of IJCAI-1997', Japan.

Gardent, C., Kohlhase, M. & van Leusen, N. (1996), Corrections and higher-order unification, *in* 'Proceedings of KONVENS', Bielefeld, Germany.

Ginzburg, J. (1995*a*), 'Resolving questions I', *Linguistics and Philosophy* **18**, 459–527.

Ginzburg, J. (1995*b*), 'Resolving questions II', *Linguistics and Philosophy* **19**, 459–527.

Ginzburg, J. (1996*a*), Dynamics and the semantics of dialogue, *in* J. Seligman, ed., 'Language, Logic and Computation: The 1994 Moraga Proceedings', CSLI Press.

Ginzburg, J. (1996*b*), Interrogatives: Questions, facts and dialogue, *in* Lappin (1996).

Ginzburg, J. (1997), Structural mismatch in dialogue, *in* A. Benz & G. Jäger, eds, 'Proceedings MunDial 97', pp. 59–80.

Ginzburg, J. (1999*a*), Fragmenting meaning: Clarification ellipsis and nominal anaphora, *in* H. Bunt, ed., 'Proceedings of the 3rd Workshop on computational semantics', Tilburg, The Netherlands.

Ginzburg, J. (1999*b*), Semantically-based ellipsis resolution with syntactic presuppositions, *in* H. Bunt & R. Muskens, eds, 'Computing Meaning: Current issues in computational semantics', Kluwer, Dordrecht.

Ginzburg, J. & Cooper, R. (2001), Resolving ellipsis in clarification, *in* 'Proceedings of the 39th Meeting of the ACL', Tolouse, France.

Ginzburg, J. & Gregory, H. (2001), SHARDS: Fragment Resolution in Dialogue, *in* H. Bunt, I. van der Silius & E. Thijsse, eds, 'Proceedings of the 4th International Conference of Computational Semantics', Tilburg, The Netherlands, pp. 156–172.

Ginzburg, J. & Sag, I. A. (2001), *Interrogative Investigations: The Form, Meaning, and Use of English Interrogatives*, number 123 *in* 'CSLI Lecture Notes', CSLI Publications, Stanford.

Green, G. M. (1996), The structure of context: The representation of pragmatic restrictions in hpsg, *in* J. Yoon, ed., 'Proceedings of the 5th annual meeting of the Formal Linguistics Society of the Midwest', Studies in the Linguistic Sciences.

Grice, H. (1957), 'Meaning', *Philosophical Review* (67), 377–388.

Grice, H. (1969), 'Utterer's meaning and intentions', *Philosophical Review* **68**(2), 147–177.

Grimshaw, J. (1982), Grammatical relations and subcategorization, *in* A. Zaenen, ed., 'Subject and Other Subjects: Proceedings of the Harvard Conference on Grammatical Relations', IULC, Bloomington, pp. 35–55.

Groenendijk, J. & Stokhof, M. (1984), Studies on the Semantics of Questions and the Pragmatics of Answers, PhD thesis, University of Amsterdam, Amsterdam.

Groenendijk, J. & Stokhof, M. (1991), ‘Dynamic predicate logic’, *Linguistics and Philosophy* **14**(1), 39–100.

Groenendijk, J. & Stokhof, M. (1997), Questions, *in* van Benthem & ter Meulen (1997), chapter 19, pp. 1055–1124.

Grover, C., Brew, C., Manandhar, S., Moens, M. & Schöter, A. (1995), Priority union and generalization in discourse grammar, *in* A. Schöter & C. Vogel, eds, ‘Nonclassical Feature Systems’, Vol. 10 of *Edinburgh Working Papers in Cognitive Science*, Dep. of Cognitive Science, Edinburgh.

Haegeman, L. (1994), *Introduction to Government and Binding Theory*, 2 edn, Blackwell, Oxford.

Haïk, I. (1987), ‘Bound variables that need to be’, *Linguistics and Philosophy* **11**, 503–530.

Halliday, M. & Hasan, R. (1976), *Cohesion in English*, number 9 *in* ‘English Language Series’, Longman’s, London.

Hardt, D. (1997*a*), ‘Dynamic interpretation of verb phrase ellipsis’, *Linguistics and Philosophy* **22**(2), 187–224.

Hardt, D. (1997*b*), ‘An empirical approach to vp ellipsis’, *Computational Linguistics* **23**(4), 525–541.

Hardt, D. (1997*c*), VPE as proform: Some consequences for binding, *in* ‘Proceedings of CSSP 97’.

Hausser, R. & Zaefferer, D. (1978), Questions and answers in a context-dependent montague grammar, *in* F. Guenthner & S. Schmidt, eds, ‘Formal Semantics and Pragmatics for Natural Languages’, Reidel, Dordrecht, pp. 339–358.

Heeman, P. A. & Allen, J. F. (1999), ‘Speech repairs, intonational phrases and discourse markers: Modeling speakers’ utternaces in spoken dialogue’, *Computational Linguistics* **25**(4), 527–571.

Heim, I. (1982), The Semantics of Definite and Indefinite Noun Phrases, PhD thesis, University of Massachusets at Amherst.

Higgins, R. (1973), J.Edmond’s analysis of extraposition, *in* J. Kimball, ed., ‘Syntax and Semantics II’, Seminar Press, New York.

Hobbs, J. (1990), *Literature and Cognition*, CSLI lecture notes, CSLI publications.

Hobbs, J. & Kehler, A. (1997), A theory of parallelism and the case of vp ellipsis, *in* ‘Proceedings of ACL-97’, pp. 394–401.

Hobbs, J. R., Stickel, M., Appelt, D. & Martin, P. (1993), ‘Interpretation as abduction’, *Artificial Intelligence* **63**, 69–142.

Huet, G. P. (1975), 'A unification algorithm for typed λ-calculus', *Theoretical Computer Science* **1**, 27–57.

Janssen, T. M. (1997), Compositionality, *in* van Benthem & ter Meulen (1997), chapter 7, pp. 417–473.

Kamp, H. (1981), A theorie of truth and representation, *in* J. Groenendijk, T. Janssen & M. Stokhof, eds, 'Formal Methods in the Study of Language', Vol. 1 of *Mathematical Centre Tracts*, Amsterdam, pp. 277–322.

Kamp, H. (1991), Prolegomena to a structural account of belief and other attitudes, *in* 'Proceedings of the 2nd Summerschool in Logic, Language and Information', EALLI, Suttgart.

Kamp, H. & Reyle, U. (1991), A calculus for first order discourse representation structures, Arbeitspapiere des Sonderforschungsbereichs 340 16-1991, Universität Stuttgart, Stuttgart.

Kamp, H. & Reyle, U. (1993), *From Discourse to Logic*, Kluwer, Dordrecht.

Kehler, A. (1993), The effect of establishing coherence in ellipsis and anaphora resolution, *in* 'Proceedings of ACL-93', pp. 62–69.

Kehler, A. (1994), Common topics and coherent situations: Interpreting ellipsis in the context of discourse inference, *in* 'Proceedings of ACL-94', pp. 50–57.

Kehler, A. (2002), *Coherence, Reference, and the Theory of Grammar*, CSLI Publications.

Klein, E. (1987), Vp-ellipsis in Dr-theory, *in* J. Groenendijk, D. de Jongh & M. Stokhof, eds, 'Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers', Dordrecht.

Klein, E. & Stainton-Ellis, K. (1989), 'A note on multiple vp ellipsis', *Language* pp. 1121–1126.

Klein, W. (1993), Ellipse, *in* J. Jacobs, A. von Stechow, W. Sternefeld & T. Vennemann, eds, 'Syntax: Ein internationales Handbuch zeitgenössischer Forschung', de Gruyter, Berlin, chapter 36, pp. 763–798.

Klein, W. & von Stutterheim, C. (1987), 'Quaestio und referentielle Bewegung in Erzählungen', *Linguistische Berichte* .

Krifka, M. (1999), For a Structured Account of Questions and Answers, *in* C. Smith, ed., 'Proceedings of the Workshop on Spoken and Written Text', University of Texas at Austin.

Kuno, S. (1987), *Functional Syntax*, University of Chicago Press, Chicago and London.

Lappin, S., ed. (1996), *The Handbook of Contemporary Semantic Theory*, Blackwell, Oxford.

Lappin, S. & Shih, H.-H. (1996), A generalized reconstruction algorithm for ellipsis resolution, *in* 'Proceedings of COLING-96', University of Copenhagen, pp. 687–692.

Lascarides, A. (2001), Imperatives in dialogue, *in* 'Proceedings of the 5th International Workshop on Formal Semantics and Pragmatics of Dialogue (BI-DIALOG 2001)', Bielefeld, Germany, pp. 1–16.

Lascarides, A. & Asher, N. (1993), 'Temporal interpretation, discourse relations and commonsense entailment', *Linguistics and Philosophy* **16**(5), 437–493.

Lascarides, A. & Asher, N. (in press), Segmented discourse representation theory: Dynamic semantics with discourse structure, *in* H. Bunt & R. Muskens, eds, 'Computing Meaning: Volume 3', Kluwer Academic Publishers, Dordrecht.

Levelt, W. & Kelter, S. (1982), 'Surface form and memory in question answering', *Cognitive Psychology* **14**, 78–106.

Levin, B. (1993), *English Verb Classes and Alternations: A Preliminary Investigation*, University of Chicago Press, Chicagoe.

Levinson, S. C. (1983), *Pragmatics*, Cambridge Textbooks in Linguistics, Cambridge University Press.

Lewis, D. (1969), *Convention*, Harvard University Press.

Marcus, M. P., Santorini, B. & Marcinkiewicz, M. A. (1993), 'Building a large annotated corpus of english: the penn treebank', *Computational Linguistics* **19**, 313–330.

Matthews, P. H. (1997), *Oxford Concise Dictionary of Linguistics*, Oxford University Press, Oxford, UK.

McKelvie, D. (1998), The syntax of disfluency in spontaneous spoken language, HCRC Research Paper HCRC/RP-95, Human Communication Research Centre, Edinburgh.

McRoy, S. & Hirst, G. (1995), 'The repair of speech act misunderstandings by abductive inference', *Computational linguistics* **21**(4), 435–478.

Milward, D. (2000), Distributing representation for robust interpretation of dialogue utterances, *in* 'Proceedings of the Meeting of the Association for Computational Linguistics / ACL 2000', pp. 133–141.

Moore, J. & Pollack, M. E. (1992), 'A problem for rst: The need for multi-level discourse analysis', *Computational Linguistics* pp. 537–544.

Morgan, J. (1978), Two types of convention in indirect speech acts, *in* P. Cole, ed., 'Syntax and Semantics 9: Pragmatics', Academic Press, New York, pp. 261–280.

Morgan, J. L. (1973), Sentence fragments and the notion 'sentence', *in* 'Issues in Linguistics: Essays in honour of Henry and Rene Kahane', UIP, Urbana.

Morgan, J. L. (1975), Some interactions of syntax and pragmatics, *in* Cole & Morgan (1975), pp. 5–18.

Morgan, J. L. (1989), Sentence fragments revisited, *in* B. Music, R. Graczyk & C. Wiltshire, eds, 'Papers from the 25th regional meeting of the Chicago Linguistics Society Part Two: Parasession on Language in Context', Chicago, USA, pp. 228–241.

Nerbonne, J. (1996), Computational semantics: Linguistics and processing, *in* Lappin (1996).

Niehren, J. & Koller, A. (1998), Dominance constraints in context unification, *in* 'Third International Conference on Logical Aspects of Computational Linguistics', Grenoble.

Oepen, S. & Flickinger, D. P. (1998), 'Towards systematic grammar profiling: Test suite technology ten years after', *Journal of Computer Speech and Language: Special Issue on Evaluation* **12**(4), 411–437.

Oepen, S., Flickinger, D., Toutanova, K. & Manning, C. D. (2002), LinGO redwoods. a rich and dynamic treebank for HPSG, *in* 'Proceedings of the First Workshop on Treebanks and Linguistic Theories (TLT2002)', Sozopol, Bulgaria, pp. 139–149.

Paul, H. (1880), *Prinzipien der Sprachgeschichte*, Niemeyer, Tübingen.

Perrault, R. (1990), An application of default logic to speech act theory, *in* Cohen et al. (1990).

Pickering, M. & Garrod, S. (in press), 'Toward a mechanistic psychology of dialogue', *Behavioral and Brain Sciences* .

Poesio, M. (1994), Discourse Interpretation and the Scope of Operators, Ph.d. thesis, University of Rochester, Rochester, New York.

Pollard, C. & Sag, I. (1994), *Head-Driven Phrase Structure Grammar*, CSLI / The University of Chicago Press, Chicago and London.

Pollard, C. & Sag, I. A. (1987), *Information-Based Syntax and Semantics, Volume 1: Fundamentals*, number 13 *in* 'CSLI Lecture Notes', CSLI Publications, Stanford, USA.

Prüst, H., Scha, R. & van den Berg, M. (1994), 'Discourse grammar and verb phrase anaphora', *Linguistics and Philosophy* **17**, 261–327.

Pustejovsky, J. (1991), 'The generative lexicon', *Computational Linguistics* **17**(4), 409–441.

Rath, R. (1979), *Kommunikationspraxis*, Göttingen.

Reyle, U. (1993), 'Dealing with ambiguities by underspecification: Construction, representation and deduction', *Journal of Semantics* **10**, 123–179.

Roberts, C. (1996), Information structure in discourse: Towards an integrated formal theory of pragmatics, *in* J. H. Yoon & A. Kathol, eds, 'Papers in Semantics', Vol. 49 of *OSUWPL*, The Ohio State University Department of Linguistics.

Rooth, M. (1996), Focus, *in* Lappin (1996), pp. 271–298.

Ross, J. (1967), Constraints on Variables in Syntax, PhD thesis, MIT.

Ross, J. R. (1969), Guess who?, *in* R. Binnick, D. A., G. Green & J. Morgan, eds, 'Proceedings of CLS 5', pp. 252–286.

Ross, J. R. (1970), Gapping and the order of constituents, *in* M. Bierwisch & K. Heidolph, eds, 'Progress in Linguistics', Mouton, The Hague.

Sag, I. (1976), Deletion and Logical Form, PhD thesis, Massachusets Institute of Technology.

Sag, I. A. (1997), 'English relative clause constructions', *Journal of Linguistics* **33**(2), 431–484.

Sag, I. & Wasow, T. (1999), *Syntactic Theory: A Formal Introduction*, CSLI Publications, Stanford.

Schlangen, D. (2002), A compositional approach to short answers in dialogue, *in* G. Mann & A. Koller, eds, 'Proceedings of the Student Research Workshop at the 40th ACL', Philadelphia, USA.

Schlangen, D. & Lascarides, A. (2002*a*), CETP: An automated theorem prover for a fragment of common sense entailment, Informatics Research Report EDI-INF-RR-0119, Edinburgh University.

Schlangen, D. & Lascarides, A. (2002*b*), Resolving fragments using discourse information, *in* J. Bos, M. E. Foster & C. Matheson, eds, 'Proceedings of the 6th International Workshop on Formal Semantics and Pragmatics of Dialogue (EDILOG 2002)', Edinburgh, pp. 161–168.

Schlangen, D., Lascarides, A. & Copestake, A. (2001), Resolving underspecification using discourse information, *in* P. Kühnlein, H. Rieser & H. Zeevat, eds, 'Proceedings of the 5th International Workshop on Formal Semantics and Pragmatics of Dialogue (BI-DIALOG 2001)', Bielefeld, pp. 79–93.

Schwabe, K. (1994), *Syntax und Semantik situativer Ellipsen*, Studien zur deutschen Grammatik, Gunter Narr Verlag.

Searle, J. (1975), Indirect speech acts, *in* Cole & Morgan (1975), pp. 59–82.

Shieber, S. (1986), *An Introduction to Unification-based Approaches to Grammar*, CSLI publications.

Sperber, D. & Wilson, D. (1986), *Relevance: Communication and Cognition*, Blackwell, Oxford.

Steedman, M. (1990), 'Gapping as constituent coordination', *Linguistics and Philosophy* **13**, 207–263.

Svartvik, J., ed. (1990), *The London Corpus of Spoken English: Description and Research*, number 82 *in* 'Lund Studies in English', Lund University Press.

Thompson, B. H. (1980), Linguistic analysis of natural language communication with computers, *in* 'Proceedings of the Eighth International Conference on Computational Linguistics', Tokyo, pp. 190–201.

Tichy, P. (1978), 'Questions, answers, and logic', *American Philosophical Quarterly* **15**, 275–284.

Traum, D., Bos, J., Cooper, R., Larsson, S., Lewin, I., Matheson, C. & Poesio, M. (1999), A model of dialogue moves and information state revision, Trindi Deliverable D2.1, University of Gothenburg.

Traum, D. & Hinkelman, E. (1992), 'Conversation acts in task-oriented spoken dialogue', *Computational Intelligence* **8**(3), 575–599.

Traum, D. R. (1999*a*), Computational models of grounding in collaborative systems, *in* 'Proceedings of the AAAI Fall Symposium on psychological models of communication', pp. 124–131.

Traum, D. R. (1999*b*), Speech acts for dialogue agents, *in* M. Wooldridge & A. Rao, eds, 'Foundations and Theories of Rational Agents', Kluwer, Dordrecht, pp. 169–201.

Tseng, J. L. (2000), The Representation and Selection of Prepositions, PhD thesis, Institute for Communicating and Collaborative Systems / Division of Informatics, Edinburgh, Scotland.

van Benthem, J. F. (1996), *Exploring Logical Dynamics*, Studies in Logic, Language, and Information, CSLI Publications / University of Chicago Press, Stanford / Chicago.

van Benthem, J. & ter Meulen, A., eds (1997), *Handbook of Logic and Language*, Elsevier, Dordrecht.

van Leusen, P. N. (1994), The interpretation of corrections, Technical Report LP-1994-21, Universitat van Amsterdam, Amsterdam, The Netherlands.

von Stechow, A. & Zimmermann, T. E. (1984), 'Term answers and contextual change', *Linguistics* **22**, 3–40.

Walker, M. A. & Moore, J. D. (1997), 'Empirical studies in discourse', *Computational Linguistics* **20**(2), 1–12.

Webber, B. (1978), A Formal Approach to Discourse Anaphora, PhD thesis, Harvard University.

Wilcock, G. (1997), Lexicalization of context, *in* 'Proceedings of the 4th international conference on HPSG', Ithaca, NY.

Williams, E. S. (1977), 'Discourse and logical form', *Linguistic Inquiry* **8**(1), 101–139.

Wittgenstein, L. (1953/84), *Tractatus Logicus Philosophicus und Philosophische Untersuchungen*, Vol. 1 of *Werkausgabe*, Suhrkamp, Frankfurt am Main.

Yanofsky, N. (1978), NP utterances, *in* D. F. Farkas, W. M. Jacobsen & K. W. Todrys, eds, 'Papers from the fourteenth regional meeting of the Chicago Linguistics Society', Chicago Linguistics Society, Chicago, pp. 491–502.