The interpretation of non-sentential utterances in dialogue

David Schlangen   Alex Lascarides
School of Informatics, University of Edinburgh
{das|alex}@cogsci.ed.ac.uk

Abstract

We present an overview of a comprehensive formal theory of the interpretation of sentential fragments, which has as components an empirically validated taxonomy, an analysis of the syntax and compositional semantics of fragments, and a formalisation of their contextual interpretation. We also briefly describe an implementation of this theory, and quantify the potential practical use of handling fragments in dialogue systems.

1 Introduction

In dialogue settings, people frequently produce utterances which, despite being non-sentential, convey propositions, questions or requests. For instance, B's utterance in (1) of the NP "John" conveys in this context the proposition "John came to the party."

(1)   A: Who came to the party?
       B: John.

Clearly, the interpretation of such non-sentential utterances or *fragments*, as they are traditionally called (eg. (Morgan, 1973)) is highly context dependent. In this paper we present an overview of a comprehensive formal theory of the interpretation of fragments.[1]

The theory has as components an empirically validated taxonomy, an analysis of the syntax and compositional semantics of fragments, and a formalisation of their contextual interpretation. We also briefly describe an implementation of this theory,

---

[1] More details can be found in (Schlangen, 2003).

and quantify the potential practical use of handling utterances of this kind in dialogue systems.

The main thesis of our approach 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 a meaningful connection of the content of the current utterance to its discourse context. We will show that the constraints on the form and content of fragments follow from such connections.

There has recently been some renewed interest in fragments. For example, (Carberry, 1990) offers an approach where computationally expensive plan-recognition techniques are employed for their interpretation. As we will show, this fails to predict certain empirical facts and further, we will show that the complex reasoning with cognitive states that she employs can often be replaced with much simpler inferences based on linguistic information. (Ginzburg and Sag, 2001) on the other hand uses grammar-based methods which, as we will show, are too weak to explain the interpretation of certain kinds of fragments where their missing content is linguistically implicit and has to be inferred. Moreover, we will show that the non-compositionality of that approach has certain disadvantages.

The remainder of the paper is organised as follows. We first describe our empirically validated, two-dimensional taxonomy of fragment-types and say something about the distribution of these types in corpora. In our theory, the type instantiated by a given fragment determines how it is resolved, but before we formalise this in Section 5, we present in Sections 3 and 4 respectively an analysis of the semantics and syntax of such utterances in isolation. This separation of grammar and resolution has cer-

tain advantages, as we discuss in comparison to previous work.

2 A Taxonomy of Fragments

2.1 The Taxonomy

We classify fragments according to two dimensions. In the first dimension, the criterion for classification is the source of the material that is needed for resolution. In examples like (1), all information that is required to resolve the meaning of the fragment (we will soon say something about what kind of information we assume this to be, syntactic or semantic) is present in the context: the resolution "John came to the party" can be computed by identifying the 'missing' content in "John" with elements from the utterance it is related to, ie. the question. We call this type of resolution *resolution-via-identity*. However, there are also examples where required information is linguistically implicit, as in (2).

(2)   A: Why did John leave?
        B: Exams.

If John is a student and this fact is mutually known by the dialogue participants, then B's utterance is presumably resolved to something like "John left because he has to take an exam soon". This resolution contains elements that are not linguistically explicit in the utterance to which the fragment is a reply; the resolution has to be *inferred*, from both linguistic sources (the content of the fragment and the content of its discourse context) and also extralinguistic sources (knowledge about preparing for exams, for example). Hence we call this type of resolution *resolution-via-inference*. We will show below that there are different constraints on fragments of these types.

In the second dimension fragments are categorised by their discourse-function or rhetorical role with respect to elements of the context. For example, in both (1) and (2), the fragment provides an answer to a question, and so we call the type these fragments instantiate *Question-Answer Pair* (*QAP*).[2] As this name suggests, strictly speaking this dimension does not classify fragments as such, but rather the relation a given fragment has to elements of its context. To make this explicit, our types are defined as relations, that is, the element of the

context the fragment connects to is explicitly part of what individuates the type.[3]

As a theoretical backbone for the taxonomy, we used the rhetorical relations defined by a theory of discourse structure called SDRT (Asher, 1993; Asher and Lascarides, 2003), which is also the framework in which we formalised the resolution of fragments. However, we also used corpora to motivate the set of fragment-types in our taxonomy. We manually identified all instances of non-sentential utterances in a number of transcripts of dialogues and classified their relation to the context.[4] The resulting list of relations is given in Tables 1 and 2, together with informal definitions of the semantics of each relation (where β is the fragment and α the utterance it is related to) and an example instance. We will not go into more details of the types here; in Section 5 we will return to a select few and give a formal definition of their semantics. Note that we do not claim that this set of fragment-types is exhaustive; we discuss in the next section the coverage that can be achieve with it on test data.[5] As a final point, note that we subsume what is often called "clarification question" (eg. (Ginzburg and Cooper, 2001)) under *Elaboration* to stress the similarity with 'normal' elaborations. The subscripts $_p$ and $_q$ indicate the message type of α and β; e.g. *Elab$_{pq}$* is an elaboration of a proposition, performed with a question.

2.2 Corpus Study

To test the coverage that can be achieved with our taxonomy, we analysed 5087 items of general free conversation from 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).[6] We proceeded

---

[2]That both (1) and (2) instantiate this type shows that the dimensions are indeed independent, since as discussed above in the first dimension they belong to different classes.

[3]All taxonomies of fragment-types that are known to us use classes that are at least partially determined by the rhetorical function of the fragment (eg., that of (Carberry, 1990) as well as that of (Fernández and Ginzburg, 2002)); nevertheless, to our knowledge ours is the only one to make this inherent relationality explicit in the formal definition of the classes.

[4]For this we used two dialogue transcripts from the BNC (Aston and Burnard, 1998) and five dialogues from the Verbmobil corpus (Wahlster, 2000).

[5]A systematic omission are relations that connect requests. In the dialogues we looked at (which were negotiations and free conversation) these did not occur; however, in more task-oriented dialogues they will presumably be quite frequent. Note that there is no principled reason for not including them, and our approach could easily be extended accordingly.

[6]We held out this data in the compilation stage, and so 'training' and 'test-data' are disjoint.

Relation	Definition, Example
QAP	β provides a direct answer to α. "*A: Who came to the party? — B: Peter.*"
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?*"
$Elab_{pp}$	β elaborates on some aspect of the indicative α. "*A: I talked to Peter. Peter Miller.*" "*A: I talked to Peter. Yesterday.*"
$Elab_{qp}$	β elaborates on the content of α. "*A: Who did you talk to? Yesterday.*" (= "*Who did you talk to yesterday?*")
$Elab_{pq}$	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?*"
$Elab_{qq}$	Any answer to β elaborates on the content of α. "*A: Did you talk to Peter? — B: Peter Miller?*"
Contr	α and β have a *contrasting theme*. "*(A: Are they in the cupboard?) — B: (α :) No, (β :) in the fridge.*"
Cont	β continues a topic of α. "*A: I am free on Monday. And on Wednesday.*"
Q-Cont	The question β continues a topic of the question α. "*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. "*A: Can you come on Tuesday? Or Wednesday?*" (= "*When can you come, Tuesday or Wednesday?*")
Expl	β explains e_α. "*A: Peter left early. Exams.*"
$Expl_q$	All answers to β explain e_α. "*A: Peter left early. — B: Exams?*"
$Expl_q^*$	All answers to β explain why α has been uttered. "*A: Are you married? — B: Why?*"
Res	α explains e_β. "*A: He had a stroke. And died.*"
Res_q	Answers to β are explained by α. "*A: He had a stroke. — B: And died?*"

Table 1: Speech act types that can be realized with NSUs

Relation	Definition, Example
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.*"
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?*"
Ack	β entails that $Agent(\beta)$ has accepted or achieved $Agent(\alpha)$'s goal behind uttering α. "*A: Let's meet on Monday. — B: OK.*"
Plan-Corr	β indicates that $Agent(\beta)$ doesn't accept or is unable to help achieve $Agent(\alpha)$'s goal behind α. "*A: Let's meet on Monday. — B: No.*"
Ack_q	positive answers γ to β entails $Ack(\alpha, \gamma)$, negative $Plan\text{-}Corr(\alpha, \gamma)$. "*A: Let's meet on Monday. OK?*"
Comnt	β indicates a propositional attitude of $Agent(\beta)$ towards the content of α. "*A: I talked to Peter. — B: Awesome!*"
$Comnt_q$	Answers to β indicate a propositional attitude of $Agent(\beta)$ towards the content of α. "*A: I talked to Peter. — B: Really?*"
Narr	e_β occurs after e_α, … "*A: He went to Italy. And (then) to Spain.*"
$Narr_q$	Answers γ to β entail $Narr(\alpha, \gamma)$. "*A: He went to Italy. — B: And then?*"

Table 2: Speech act types that can be realized with NSUs (contd)

in two stages, first semi-automatically identifying non-sentential utterances (using the wide-coverage grammar described below as a filter for sententiality) which we then classified according to the taxonomy, or with `other` if no decision could be made.[7]

Results Numbers about the frequency of fragments in our corpus and about the achieved coverage overall are presented in Table 3. For reasons of space, we cannot show the detailed distribution of the classes here; what is noteworthy about it is that the majority of fragment instances is concentrated in a few classes, with the rest of the classes only represented by a few examples each. The most frequent type is QAP, followed by $Elab_{pq}$ and $Elab_{pp}$,

[7]For this study we only used one annotator, so we cannot present inter-annotator agreement measures. Further studies are underway at the moment that will give us such data.

whereas types like *Expl*$_q^*$, *Q-Alt* and *Contr* were found only a few times.

SMALLCAPS SUMMARY

items analysed	9142	
Fragments	931	(= 10.2%)
classfd.	865	(= 93% of fragments)
other	66	(= 7% of fragments)

Table 3: Results of Annotation

Discussion The overall percentage of fragments we have found seems to confirm the results of earlier studies (Thompson, 1980; Fernández and Ginzburg, 2002), which also classified as fragments around 10% of the utterances in the dialogues they looked at. (Fernández and Ginzburg, 2002) also offers a taxonomy of fragment types; the authors claim to have reached a coverage which is much higher than what we achieved (99% compared to our 93%). We think this can partially be explained by the fact that the classes they use are more surface-oriented. For example, they have a class 'sluice', which is defined as "bare question-denoting *wh*-phrases" (Fernández and Ginzburg, 2002, p.16). We make a finer distinction, splitting this class further according to the rhetorical function played by the fragmental question (eg., *Expl*$_q$, *Expl*$_q^*$, *Narr*$_q$, *Elab*$_q$). Hence, to classify a given fragment we need more information about its rhetorical function; information which sometimes is difficult to recover from a transcript. While a surface-oriented approach to defining classes is an advantage for annotation, it is a hindrance for formally defining their semantics, as we will argue below.

In conclusion, we think that this corpus study has shown that our taxonomy has satisfactory empirical motivation. In the next two sections we look at fragments in isolation. We return to the taxonomy when we give an example of the formal semantics of relations and show how they determine the resolution of the fragments.

3 A Compositional Semantics for Fragments

For compositional semantic analysis we use Minimal Recursion Semantics (MRS, (Copestake et al., 1999)), a language in which partial descriptions of formulae of a logical language (the *base language*) can be expressed. This allows one to leave certain semantic distinctions unresolved, reflecting the idea that syntax supplies only partial information

about meaning. Technically this is achieved via a strategy that has become standard in computational semantics (e.g., (Reyle, 1993)): one assigns labels to bits of base language formulae so that statements about their combination can remain 'underspecified'. The (first-order) models of formulae of this latter language then can be seen as standing in a direct relation to formulae of the base language; $M \models \phi$ then means that the (unique) base-language formula corresponding to M is described by the MRS ϕ.[8] By way of example, (3) shows an MRS-representation of "Everyone loves someone", where so-called elementary predications (EPs) are labelled with *handles* (h_n), with h being the top handle that outscopes all others; '$h_1 =_q h_2$' stands for an 'outscopes' relation between EPs where only quantifiers can be scoped in between h_1 and h_2; *prpstn_rel* signals that the MRS describes a proposition.

(3) $\langle h, e, \{h\colon prpstn\_rel(h_1), h_2\colon love\_v\_rel(e, x_1, x_2),$
       $h_6\colon \_every\_rel(x_1, h_8, h_9),$
       $h_{10}\colon \_person\_rel(x_1),$
       $h_{11}\colon \_some\_rel(x_2, h_{12}, h_{13}),$
       $h_{14}\colon \_person\_rel(x_2)\},$
       $\{ \ h_1 =_q h_2, h_8 =_q h_{10}, h_{12} =_q h_{14} \}\rangle$

The compositional semantics of fragments leaves more information unresolved than just semantic scope, however. All we know about the meaning of fragments like those in (1) and (2) independent from their context is: (a) they will resolve to a proposition,[9] of which (b) the main predicate is unknown, but (c) one participant in the main event is specified although its exact role isn't. We represent this with an anaphoric relation *unknown_rel*, and so the NP-fragment "John." (regardless of the context it stands in) is represented as:

(4) $\langle h, e, \{h\colon prpstn\_rel(h_1),$
       $h_2\colon unknown\_rel(e, x),$
       $h_6\colon def\_np\_rel(x, h_8, h_9),$
       $h_{10}\colon named(x, \text{John})\},$
       $\{ \ h_1 =_q h_2, \ h_8 =_q h_{10} \ \} \rangle$

The *unknown_rel* acts as a 'place-holder' for a potentially complex sub-formula; more precisely it is a constraint on the form of the described (base-

---

[8]Such a semantics if given to MRS in (Copestake et al., 2001); we follow the similar formalisation in (Asher and Lascarides, 2003). Note that we do not make any assumptions about the base language and its logic here; the descriptions are compatible with it being static first order predicate logic, or a dynamic logic like DRT (Kamp and Reyle, 1993).

[9]We make the simplifying assumption that there is an unambiguous intonation pattern indicating whether a fragment is intended to convey a proposition, a question or a request.

language) formulae, that they contain at this place a subformula, which in the case of (4) must have e and x amongst its variables. Clearly, such a description then describes an infinite number of formulae; however, all of these are potential resolutions of the fragment. For instance, (4) (partially) describes the intended resolution of the fragment in context (1)— "John came to the party."—, but it also describes for example "Carl loves John." or "Sandy thinks that Kim relies on John.", which can be resolutions in other contexts.

It is important to note that *unknown_rel* is *not* a second order variable (as it would be in an approach in the vein of (Dalrymple et al., 1991)), and it is *not* something that simply gets replaced by a predicate symbol of the same arity. Rather, *unknown_rel* is a constraint more like the $=_q$-constraints, constraining the 'shape' of the described formulae. It is anaphoric, because the sub-formula that is to be inserted at this point in the described formula is not determined by the grammar, but must be provided by the context.

4 A Grammar of Fragments

4.1 The Analysis

Our grammatical analysis of fragments is relatively straightforward: we make the assumption that fragments are phrases,[10] possibly modified by adverbs. As (5) shows, only scopally modifying adverbs are allowed.[11]

(5)   A: Who sang this song?
       B: Maybe Sandy. / *Badly Sandy.

In a pseudo phrase-structure notation, the rules simply are of the form 'S-frag → (ADV) XP'. We formalise this in a version of HPSG that allows *constructions* (Sag, 1997) (ie. phrase-types that make a semantic contribution) and that uses MRSs as semantic representations. HPSG-representations of these semantic structures consist of a feature IN-DEX whose value represents the semantic index of the sign; a feature LTOP that holds the *handle* of the sign, ie. a label for the bits of logical form introduced by it; LZT, which is a bag of labelled EPs; and H-CONS, which collects the constraints on the

order of sub-formulae.

Figure 1: "John" as a declarative fragment.

The formalisation is best explained with an example. Figure 1 shows, in a tree representation, the sign for the NP-fragment "John." It demonstrates how the NP is lifted to the level of sentences, and how the semantics of that sentence is composed.

Let's work 'top-down' to describe this Figure in detail. The root-sign in this tree has all the syntactic features of a sentence: the value of its SYN-SEM.LOCAL.CAT is of type *verb*, and all valence requirements are satisfied. It is also semantically like a sentence, in that its top-EP (with the handle 2) is of type *message* (more precisely, a *prpstn*). This

---

[10]This goes back to (Morgan, 1973); explicit rules can be found in (Barton, 1990). We ignore for now more complicated examples like 'A: Does John devour or nibble at his food? — B: Oh, John devours.'

[11]Note that the latter *is* licensed as an answer to a multiple *wh*-question like "Who sang, and how?", in which case we analyse it as a sequence of two fragments.

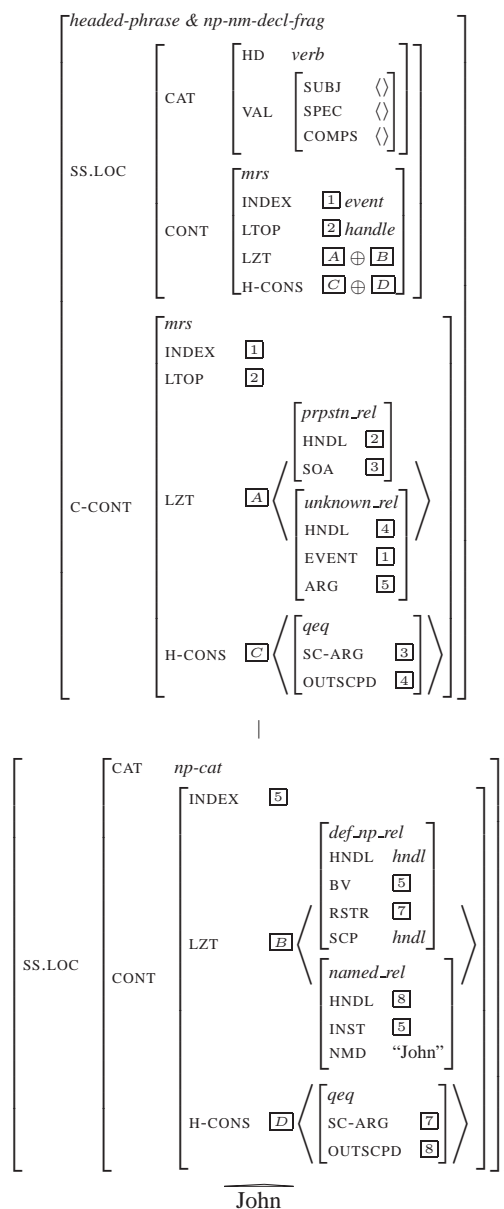EP is contributed by the fragment-rule, via the feature C-CONT (construction content). In the same way the *unknown_rel*-constraint that was introduced in the previous section is added. The connection of this constraint to the semantics of the phrase is made via co-indexation of the argument-slot of *unknown_rel* with the INDEX of the argument phrase (in Figure 1 this is ⑤).
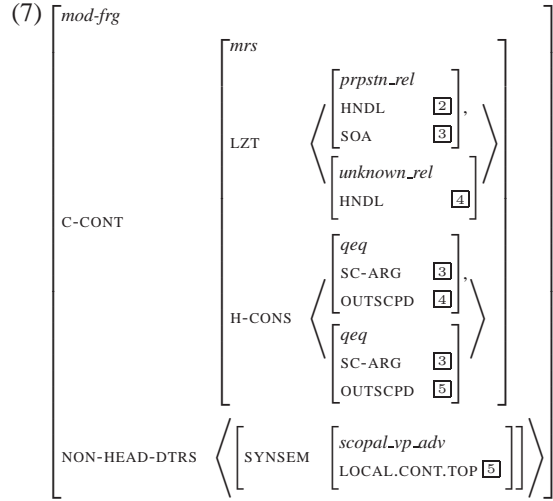
As the type-declaration in Figure 1 shows, this sign is the combination of two types, namely *headed-phrase*, which is a general type that defines the features and co-indexations in headed phrases; and *np-nm-decl-frag*, which collects the specifications particular to fragments. This type in turn inherits from three further types: *np-frag*, which specifies the particularities of fragments consisting of NPs; *nm-frag*, which specifies non-modified fragments (ie., a phrase that is not modified by an adverb); and *decl-frag*, which indicates that the fragments resolves to a proposition. These three types encapsulate properties of fragments that can vary independently; see the hierarchy in Figure 2.

We assume a generalised head-feature principle (*ghfp*) as in (Ginzburg and Sag, 2001) according to which all values for SYNSEM-features on the mother are by default token-identical to those of the daughter, and hence we have to make sure that the fragment-types override this default where appropriate. For example, the value for SYNSEM.LOCAL of fragments 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 is the whole point of the rule, and so the default of the *ghfp* to copy these specifications must be overridden. The value of SYNSEM.LOCAL.CAT will be the same for all types of fragments, namely that of a sentence. In fact, the only elements of the type instantiated in Figure 1 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 the restriction that the phrase be an NP. So the constraint unique to NP-fragments (ie., the specification of the type *np-frag*) is that shown in (6).

(6)   *np-frg:*

$$\begin{bmatrix} \text{C-CONT.LZT} & \left\langle \begin{bmatrix} \ \end{bmatrix}, \begin{bmatrix} \text{ARG} & \boxed{1} \end{bmatrix} \right\rangle \end{bmatrix} \rightarrow$$

$$\mathbf{H} \begin{bmatrix} \text{SYNSEM.LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & nominal \\ \text{VAL} & \begin{bmatrix} \text{COMPS} & \langle \rangle \\ \text{SPR} & \langle \rangle \end{bmatrix} \end{bmatrix} \\ \text{CONT.INDEX} & \boxed{1} \end{bmatrix} \end{bmatrix}$$

Figure 1 represents a non-modified fragment. In fragments that are modified by an adverb, we find an additional non-head-daughter, whose EP is scoped in as sister to the *unknown_rel*, as shown in (7).

(7)

$$\begin{bmatrix} mod\text{-}frg \\ \text{C-CONT} & \begin{bmatrix} mrs \\ \text{LZT} & \left\langle \begin{bmatrix} prpstn\_rel \\ \text{HNDL} & \boxed{2} \\ \text{SOA} & \boxed{3} \end{bmatrix}, \begin{bmatrix} unknown\_rel \\ \text{HNDL} & \boxed{4} \end{bmatrix} \right\rangle \\ \text{H-CONS} & \left\langle \begin{bmatrix} qeq \\ \text{SC-ARG} & \boxed{3} \\ \text{OUTSCPD} & \boxed{4} \end{bmatrix}, \begin{bmatrix} qeq \\ \text{SC-ARG} & \boxed{3} \\ \text{OUTSCPD} & \boxed{5} \end{bmatrix} \right\rangle \end{bmatrix} \\ \text{NON-HEAD-DTRS} & \left\langle \begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} scopal\_vp\_adv \\ \text{LOCAL.CONT.TOP} & \boxed{5} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

Finally, the last dimension organises the differences in the type of message to which the fragment will resolve. The example we have seen in Figure 1 was one of a propositional-fragment; fragmental questions or requests only differ in the type of this message-relation. To give an example, (8) shows the type *int(errogative)-frag(ment)*.

(8)

$$\begin{bmatrix} int\text{-}frag \\ \text{C-CONT.LZT} & \left\langle \begin{bmatrix} int \end{bmatrix}, \dots \right\rangle \end{bmatrix}$$

The rules in this dimension also make sure that *wh*-phrases must be *int-frag*s.

4.2   Implementation

We have implemented our analysis in a wide-coverage HPSG, the *English Resource Grammar* (ERG, see for example (Copestake and Flickinger, 2000));[12] the implementation was evaluated using the grammar-profiling tool [incr tsdb()] (Oepen and Flickinger, 1998). First, to test for possible adverse effects on the analyses of full-sentences, we ran a batch-parse of a test-suite of full sentences, the CSLI-test-suite which is distributed with [incr tsdb()]. It consists of 1348 sentences, of which 961 are marked as syntactically well-formed and 387 as ill-formed. Table 4 shows a comparison of the original ERG with our extended version containing the fragment rules, with respect to the average number of parses per sentence.

---

[12]The implementation differs slightly from the analysis described in the previous section: the ERG doesn't make use of defaults, and so we had to explicitly state what is identical between mother and daughter and what isn't.
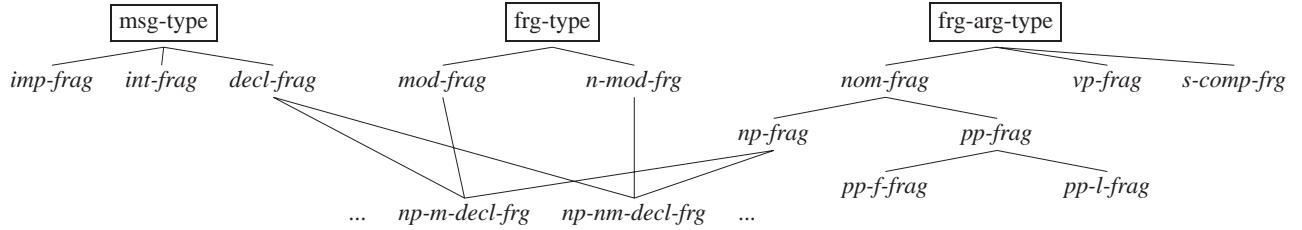
Figure 2: An extract of the construction hierarchy for fragments

Version of Grammar	Average # parses
LinGO ERG, 20/11/02	2.86
ERG+frag	3.69

Table 4: Competence comparison of the original ERG with the fragment-ERG

As these data show, the fragment rules introduce some new ambiguity, but on average less than one more parse per item. We conclude from this that adding the fragment-rules doesn't lead to an explosion of readings that would render the grammar practically unusable. What this evaluation doesn't tell us, however, is whether the additional readings (of what is meant to be full sentences) are erroneous or not. The problem is that 'fragmenthood' is not a syntactic criterion, and so some strings that can be analysed as sentences can also be analysed as fragments. E.g., "leave" can be both an imperative sentence and a VP-fragment (e.g. in the context of the question "What did John make Sandy do?").

To test the coverage of our extended grammar, we used parts of the annotated corpus described in Section 2.2. In 4037 items we identified 369 fragments, of which our grammar correctly parsed 242 (= 65.5%). A detailed study of the fragments that were not recognised showed that a useful extension would be rules for handling fragments of the form "CONJ XP", eg. "and on Saturday."; including those would bring our coverage up to 82.6% of the corpus. This result is in the same order of magnitude as what the original grammar achieves on full sentences, and is in the range of what the best wide-coverage grammars that provide semantic representations at present can achieve. However, although extending the grammar in this way is straightforward it would lead to a dramatic rise in the number of average parses, and so for practical purposes we did not carry out such an extension here.

5 Computing the Intended Meaning of Fragments

5.1 Theory

The final problem we have to address is how the underspecified semantic representations that our grammar produces for fragments are resolved contextually. For this, we use a theory of discourse interpretation called SDRT (Asher and Lascarides, 2003). This theory attempts to enrich dynamic semantics with techniques for encoding the contribution of pragmatics. One central notion of dynamic semantics (eg. (Kamp and Reyle, 1993)) is the update of a representation of the context with that of new information; in SDRT, this update is dependent on non-monotonic inferences over linguistic and non-linguistic information. SDRT's update-operation is defined on descriptions like MRSs; it simply adds constraints on the form of logical forms. The inferred information that is most important for us is the *speech act type* (e.g. *QAP*, *Elab*$_{pp}$) that connects the new information to the context; this is what we used to classify fragments in Section 2.

These speech acts are computed via default rules; to give an example, the rule for *IQAP* is shown below. In this rules, $\langle \tau, \alpha, \beta \rangle$ means β is to be attached to α with a rhetorical relation (α and β label bits of content) where α is part of the discourse context τ; α :? means that α is an interrogative, and $A > B$ means *If A then normally B*:[13]

(9)   IQAP: $(\langle \tau, \alpha, \beta \rangle \wedge \alpha :?) > IQAP(\alpha, \beta)$

This rule stipulates that the default contribution of a response to a question is to supply information from which the questioner can infer an answer. Thus inferences about speech acts, and hence about implicit content and goals, can sometime be triggered (by default) purely on the basis of sentence moods. For

---

[13](Asher and Lascarides, 1998) shows that some rules like this can be derived from a precise model of rationality and cooperativity.

other speech acts, information about speaker-goals might be required; however, the general principle is to always minimise the need for such information.

We now address resolving the underspecification indicated by *unknown_rel*. In particular, we argue that there are certain constraints on the *form* of *resolution-via-identity* fragments which do not apply to *resolution-via-inference* fragments. We derive these different constraints from a general discourse coherence principle, but before we come to this, we have to show what these constraints are. We begin with questions and answers like (1).

Intuitively, one can say that there is a 'hole' in questions like (1) or (10), marked syntactically by the *wh*-phrase and semantically by a variable (be that bound by a λ-operator, as in (Groenendijk and Stokhof, 1984) or by a quantifier, as in the ERG).

(10) A: Who came to the party?
      B: Not Sandy.

This initially suggests that to resolve the content of the fragment, one could attempt to do syntactic reconstruction, 'plugging' the syntactic structure of the fragment into the (syntactic) 'hole' in the question (cf. (Morgan, 1973)). Unfortunately, as (Ginzburg and Sag, 2001) (henceforth G&S) attest, such a strategy fails for some cases; eg. for (10) above: "*Not Sandy came to the party."

On the other hand, G&S also attest that a purely semantic reconstruction, where the semantic representation of the fragment is 'plugged into' the (semantic) 'hole' in the question, is also unsatisfactory. Certain grammatical idiosyncrasies seem to persist beyond sentence boundaries. This can be shown with example (11) (from G&S, p.300). Here the fragmental answers must be of the syntactic category required by the verb in the question (VP[*bse*] and VP[*inf*], respectively), even though the semantic objects denoted by these VPs presumably are of the same type.[14]

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

G&S model this constraint by restricting short answers to be syntactically parallel to the question they answer. This however is too strict, as examples like (2) show, which is a short answer as well but does not exhibit such 'parallelism'. We explain the observation in a less direct but more

---

[14]For a further discussion of the exact extent of this parallelism see (Schlangen, 2002).

general way. First of all, our theory is declarative: it describes the form of the preferred resolution (it's the one that satisfies the coherence constraints of the rhetorical relations), but not necessarily how it is generated. We assume as a general discourse principle that resolutions which are (semantic-)structurally very close must satisfy a certain syntactic constraint which says that subcategorization requirements must be satisfied, too. Hence our principle can rule out resolutions even if they satisfy the semantic constraints (eg., it would rule out the 'wrong' combination of questions and answers in (11) above), in case they violate that syntactic constraint. The difference between (1) and (2) now is explained by different contextual requirements. In (2) there is another relation present besides *QAP*, namely *Explanation*. The semantics of this former relation (namely that β explains the propositional content of α) puts additional semantic constraints on the answer; the structural closeness is not required, and hence the fragment is exempt from the syntactic constraint.

For reasons of space, we simply sketch the theory here, but we should stress that, following G&S, we also assume that certain syntactic information persists beyond sentence boundaries.

5.2 Implementation

We have partially implemented the theory described here in a computer program (see also (Schlangen and Lascarides, 2002)). The resolution of *res-via-id* fragments is very straightforward to implement, since for them all possibilities can be generated via simple abstraction and functional application operations over the semantic representations. That of *res-via-inf* fragments is more problematic, and we have only implemented it for a very limited domain, namely that of scheduling dialogues. In this domain, the discourse-plans are particularly simple, and so we can specify the required axioms for reasoning with extra-linguistic information. Again we can't go into details here and only note that even though we minimise the amount of extra-linguistic information that is needed, resolution of *res-via-inf* fragments is a demanding task and can be automated only for very restricted domains.
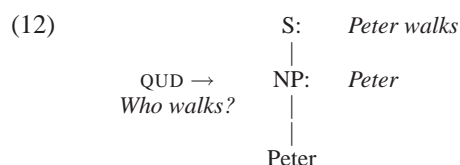
6 Related Work

The idea that content is determined by coherence relations is of course not new, and has been implemented for example in (Hobbs et al., 1993), which also mentions in passing the problem of resolving fragments in context. However, this 'Interpreta-

tion as Abduction'-theory (IAT) differs from our approach in a number of important aspects. First, unlike IAT's weighted abduction where conflict among the clues to interpretation is handled by the extraneous logical machinery of weights, in our theory conflict is resolved automatically by the logical consequence relation itself. Secondly, Hobbs et al. don't consider the syntactic constraints on the resolution of fragments that we discussed above. In fact, they seem to regard fragments as 'syntactically-ill formed utterances', and so do not make a difference between well-formed and ill-formed fragments. In principle, further constraints could be added to the ITA framework, but at the cost of having to re-assign weights so that the results of inference are always as intended, and no princples or regulations are given in (Hobbs et al., 1993) about how to do this.

As mentioned in the introduction, (Ginzburg and Sag, 2001) (henceforth G&S) also offer a non-modular approach to the resolution of short-answers (and some other fragmental speech acts). (12) shows a very schematic representation of their approach.

(12)

$$
\begin{array}{ll}
 & \text{S:} \quad \textit{Peter walks} \\
 & \quad | \\
\text{QUD} \rightarrow & \text{NP:} \quad \textit{Peter} \\
\textit{Who walks?} & \quad | \\
 & \quad | \\
 & \text{Peter}
\end{array}
$$

A grammar rule specific to the *use* made of the fragment (in (12) as an answer) directly projects NPs as sentences, with parts of the sentential content coming from a contextual feature QUD (question under discussion). This grammar rule in one go checks the syntactic constraints and constructs the intended content of the fragment in its discourse context.

In our opinion, our compositional approach has certain advantages. First, the grammatical analysis of fragments is uniform; contextual variation in their meaning is accounted for in the same way as it is for other anaphoric phenomena, via inferences underlying discourse update. This yields the second advantage: resolving fragments is fully integrated with resolving other kinds of underspecification (although we have not shown this here; cf. (Schlangen and Lascarides, 2002)). Third, the interaction between grammar and pragmatics is straightforward: pragmatics enriches information coming from the grammar. In G&S's approach the grammar has to 'decide' on the speech act that has been performed (the grammar-rules are specific for eg. answering, clarification); something that is normally seen to be a defeasible process. Hence, even in

G&S's approach a pragmatic module is required, which then has the task of filtering out unwanted parses. Fourth, we have available a strong theory of contextual interpretation which can explain the reasoning behind the resolution of examples like (2) (although we have not shown here in detail how); the functional application used by G&S seems too weak to do this. Fifth, our compositional approach allowed us to straightforwardly extend an existing wide-coverage grammar. This contrasts with the non-compositional approach which through its demands for making contextual information available entails that standard parsers cannot be used without modifications. Lastly, the separation of the grammar and resolution components means that in theory our grammar can be used with different resolution strategies; however, we have not systematically explored that.

As mentioned in the introduction, (Carberry, 1990) offers an approach that uses plan-recognition techniques to resolve fragments. While such an approach can perhaps model *res-via-inf* cases, it seems to us needlessy powerful for fragment-types like (1), where purely linguistic information is sufficient. Moreover, Carberry does not deal with the syntactic constraints and so overgenerates possible fragments.

7 Conclusions

We draw the following conclusions from the work presented in this paper: fragments occur frequently in dialogues, namely relatively consistently around 10% across dialogue types (but possibly more frequently in question/answer-based informative dialogues). This means that a principled approach to their resolution is important for natural sounding dialogue systems, besides being of theoretical interest. We have offered such an approach, beginning with a comprehensive taxonomy of fragment-types, through to a semantic and syntactic analysis, which we also implemented. In that implementation we identified for future work the sub-type of fragments of the form 'CONJ XP' (eg. "and on Monday.").

Acknowledgements

References

Nicholas Asher and Alex Lascarides. 1998. Questions in dialogue. *Linguistics and Philosophy*,

23(2):237–309.

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

Nicholas Asher. 1993. *Reference to Abstract Objects in Discourse*. Studies in Linguistics and Philosophy. Kluwer Academic Publisher, Dordrecht.

Guy Aston and Lou Burnard. 1998. *The BNC Handbook*. Edinburgh University Press, Edinburgh, UK.

Ellen L. Barton. 1990. *Nonsentential Constituents*. John Benjamins, Amsterdam / Philadelphia.

Sandra Carberry. 1990. *Plan Recognition in Natural Language Dialogue*. MIT Press, Cambridge, Massachusetts.

Ann Copestake and Dan Flickinger. 2000. An open source grammar development environment and broad-coverage english grammar using HPSG. In *Proceedings of the 2nd Linguistic Resources and Evaluation Conference*, pages 591–600, Athens, Greece.

Ann Copestake, Dan Flickinger, Ivan Sag, and Carl Pollard. 1999. Minimal recursion semantics: An introduction. Stanford University, Stanford, CA.

Ann Copestake, Alex Lascarides, and Dan Flickinger. 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)*, pages 132–139, Tolouse, France.

Mary Dalrymple, Stuart Shieber, and Fernando Pereira. 1991. Ellipsis and higher order unification. *Linguistics and Philosophy*, 14:399–452.

Raquel Fernández and Jonathan Ginzburg. 2002. Non-sentential utterances in dialogue: A corpus-based study. In Kristiina Jokinen and Susan McRoy, editors, *Proceedings of the Third SIGdial Workshop on Discourse and Dialogue*, pages 15–26, Philadelphia, USA. ACL Special Interest Group on Dialog.

Jonathan Ginzburg and Robin Cooper. 2001. Resolving ellipsis in clarification. In *Proceedings of the 39th Meeting of the ACL*, Tolouse, France.

Jonathan Ginzburg and Ivan A. Sag. 2001. *Interrogative Investigations: The Form, Meaning, and Use of English Interrogatives*. Number 123 in CSLI Lecture Notes. CSLI Publications, Stanford.

Jeroen Groenendijk and Martin Stokhof. 1984. *Studies on the Semantics of Questions and the Pragmatics of Answers*. Ph.D. thesis, University of Amsterdam, Amsterdam.

J. R. Hobbs, M. Stickel, D. Appelt, and P. Martin. 1993. Interpretation as abduction. *Artificial Intelligence*, 63:69–142.

Hans Kamp and Uwe Reyle. 1993. *From Discourse to Logic*. Kluwer, Dordrecht.

Jerry L. Morgan. 1973. Sentence fragments and the notion 'sentence'. In *Issues in Linguistics: Essays in honour of Henry and Rene Kahane*. UIP, Urbana.

Stephan Oepen and Daniel P. Flickinger. 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.

Uwe Reyle. 1993. Dealing with ambiguities by underspecification. *Journal of Semantics*, 10:123–179.

Ivan A. Sag. 1997. English relative clause constructions. *Journal of Linguistics*, 33(2):431–484.

David Schlangen and Alex Lascarides. 2002. Resolving fragments using discourse information. In Johan Bos, Mary Ellen Foster, and Colin Matheson, editors, *Proceedings of the 6th International Workshop on Formal Semantics and Pragmatics of Dialogue (EDILOG 2002)*, pages 161–168, Edinburgh, September.

David Schlangen. 2002. A compositional approach to short answers in dialogue. In Gideon Mann and Alexander Koller, editors, *Proceedings of the Student Research Workshop at the 40th ACL*, Philadelphia, USA, July.

David Schlangen. 2003. *A Coherence-Based Approach to the Interpretation of Non-Sentential Utterances in Dialogue*. Ph.D. thesis, School of Informatics, University of Edinburgh, Edinburgh, UK.

Bozena H. Thompson. 1980. Linguistic analysis of natural language communication with computers. In *Proceedings of the Eighth International Conference on Computational Linguistics*, pages 190–201, Tokyo.

Wolfgang Wahlster, editor. 2000. *Verbmobil: Foundations of Speech-to-Speech Translation*. Springer, Berlin.