Ellipsis Resolution by Controlled Default Unification for Multi-modal and Speech Dialog Systems

Michael Streit
German Research Center for
Artificial Intelligence - DFKI
Stuhlsatzenhausweg 3
Saarbrücken
Germany
streit@dfki.de

HansUlrich Krieger
German Research Center for
Artificial Intelligence - DFKI
Stuhlsatzenhausweg 3
Saarbrücken
Germany
krieger@dfki.de

Abstract

We present a default-unification-based approach to ellipsis resolution that is based on experience in long running multimodal dialog projects, where it played an essential role in discourse processing. We extend default unification to non-parallel structures, which is important for speech and multimodal dialog systems. We introduce new control mechanisms for ellipsis resolution by considering dialog structure with respect to specification, variation and results of tasks and combine this with the analysis of relations between the information elements contained in antecedent and elliptic structures.

1 Introduction

The application of default unification (Carpenter, 1993) or priority union (Kaplan, 1987; Calder, 1991) to discourse is attractive, because these related concepts meet the intuition that new information extends, corrects, or modifies old information, instead of deleting it, by keeping what is consistent.[1] The use of default unification as a means for ellipsis resolution has been discussed in the first half of the nineties (Prüst et al., 1994; Grover et al., 1994). Later, the discussion silted up, perhaps because the conditions on parallelism that have been imposed occured to be too strong (cf. (Hobbs and Kehler, 1997)).

1.1 Applications in Dialog Systems

Since this discussion, default unification-based ellipsis resolution has been applied in working systems of at least two projects, where it played an essential role in discourse processing.

The first implementations have been provided in the second half of the nineties at Siemens, where the DIAMOD project developed a serial of prototypes for multi-modal human machine dialog (cf. (Streit, 2001)).

[1] Priority union as introduced by Kaplan bears essentially the same idea as Carpenter's default unification.

The DIAMOD project realized applications for appointment management and driver assistance, but also for controlling machines and other technical devices (Streit, 1999). The DIAMOD systems provided the user with a quite natural dialog, including clarification and repair dialogs and even multitasking (e,g,. the user could deal with different appointments in parallel).

Applied under appropriate conditions, default unification turned out to be a robust and efficient method for VP ellipsis and even NP ellipsis. It was also successfully used to inherit information in situations without any syntactically recognizable ellipsis.

Later in the SmartKom project (Wahlster et al., 2001), default unification and its application on ellipsis resolution was reinvented under the label *overlay* (Alexandersson and Becker, 2003). Overlay basically consists of priority unions applied to frame-like semantic representations without considering reentrencies. This inhibits providing sloppy readings of an ellipsis (without impairing the dialog performed in the systems very much).

1.2 Problems with Default Unification

Default unification without additional control shows an inherent tendency to over-accumulate information. Even worse, the method may accumulate information that is semantically inconsistent (but not recognized as such) or at least is practically absurd. Such inconsistencies or absurdities typically arise from dependencies between information elements that are not expressed (or not expressible) in the type hierarchy, domain model, or ontology that is underlying the default unification process. For instance, it may occur that by introducing a name of a new object, the address of an old object of the same kind is inherited, which is pragmatically absurd. Or a numeric date or time specification may be wrongly combined with a deictic reference to another date or time, which is semantically inconsistent. On the other hand, the intrinsic parallelism of default unification does inhibit the handling of fragmentary and other non-parallel ellipsis, which is very common in spoken dialog.

1.3 Problems with Ellipsis Resolution in Information Seeking Dialog

In dialog systems that serve for browsing information sources, it is significant that the user modifies and varies her query, be it spontaneously or systematically. As is discussed mainly in section 5, removing and inheriting old information are equally important for ellipsis resolution in this type of interaction. We want to notice, that the removal of information is independent from the question if information has been grounded (Traum, 1994) or not. Up to now, studies hardly consider these problems.

1.4 Overview

In this paper we will present a revised and fairly extended version of the methods developed in the DIAMOD project.[2] We discuss two main problems. On the one hand we show how default unification can be applied to non-parallel or fragmentary structures, on the other hand we discuss dependencies of ellipsis resolution from the structure of information and tasks, that are rarely addressed in the literature. Especially we discuss the following problems.

- The extension of default unification to fragments.
- Control of ellipsis resolution by considering dialog structure w.r.t. specification, variation and results of tasks.
- Control of ellipsis resolution by considering relations between old and new information.
- Handling of set-valued features in specifying vs. varying dialog phases.

We will couch our approach in terms of typed feature structures. Thereby features correspond to slots, and types correspond to frames.

We assume feature structures being well-typed, but not totally well-typed (Carpenter, 1992). This means that for every type it is defined which features it can have, and which types come into question as values of these features. But it is not required that every possible feature is present or has a value. In order to use feature structures for the purpose of encoding semantic representations, we need set-valued features (which are semantically interpreted as conjunctions of its elements). For instance, a movie may be described by a conjunction of genres (e.g., crime and science fiction), and an appointment usually has more than one participant.[3]

In this paper we will mainly refer to examples taken from DIAMOD, but also make use of material taken from SmartKom user input. We note here that the methods described in this paper are not implemented in SmartKom. We also consider examples as they are discussed in the literature.

2 Default unification

Default unification is a method to inherit *defeasible* (in our case *old*) information which does not contradict *strict* (in our case *new*) information. As already mentioned, the consistency criterion is to weak, but the basic approach is useful. There are two forms of default unification: credulous and skeptical default unification. Credulous default unification tries to maintain as much old information as possible. Due to structure sharing, there are often different alternatives for achieving a maximal amount of old information. Skeptical default unification takes only the information that is common to all credulous solutions. We are interested in getting every maximal solution, which correspond to the strict, sloppy or mixed readings of ellipsis. By mixed readings we mean readings that contain a strict reading in one part, and a sloppy reading in another.

We follow the definition of credulous default unification provided by Carpenter (Carpenter, 1992). But we take the most general type as the top element of the type lattice, while Carpenter takes it as the bottom element.

If O is the old, defeasible information and N is new, strict information, then the credulous default unification of O and N is the unification of O' with N, where O' is a minimal structure that subsumes O such that O' and N unify:

$$O \overset{>}{\sqcap}_c N = \{O' \sqcap N | O' \sqsupseteq O \text{ minimal s.t. } O' \sqcap N \neq \bot\}$$

The following example, shows how default unification can be used in ellipsis resolution.

1. John revises his paper.

$$\begin{bmatrix} \textit{revise} & \\ \text{AGENT} & \boxed{1}\,\textit{john} \\ \text{OBJECT} & \begin{bmatrix} \textit{paper} & \\ \text{AUTHOR} & \boxed{1} \end{bmatrix} \end{bmatrix}$$

[2]The quite simple but effective means for controlling default unification that have been introduced in the DIAMOD systems have not been published yet.

[3]The first case concerns a closed class of values. In principle, this case could technically be solved without set-valued features by introducing a highly differentiated (rather artificial) type system with a type for every value combination. The second case concerns an open class of entities that cannot be compiled in a type system.

2. And Bill does too.

$$\left[\begin{array}{ll} \textit{event-agentive} & \\ \text{AGENT} & \textit{bill} \end{array}\right]$$

The analysis of these utterances is slightly simplified. John would get a more complicated presentation with john being the value of the *NAME* feature of the type *person*. The verb *do* is considered as being the most general verb with an agent. We use here *event-agentive* as a supertype of activities.

In this example the types of the top nodes are on a comparable "level". By *being on a comparable level* we mean that the top node of the one item is a supertype of the top node type of the other item. Notice that due to the well-typing condition, types and features may not be mixed arbitrarily . For instance, the most general type of the type hierarchy (and many others too), cannot be combined with the feature *agent*. Otherwise our level condition would be meaningless.

We find two minimal upper bounds of (1.) that unify with (2.).

$$(1') \left[\begin{array}{ll} \textit{revise} & \\ \text{AGENT} & \textit{john} \\ \text{OBJECT} & \left[\begin{array}{ll} \textit{paper} & \\ \text{AUTHOR} & \textit{john} \end{array}\right] \end{array}\right]$$

$$(1'') \left[\begin{array}{ll} \textit{revise} & \\ \text{AGENT} & \boxed{1} \\ \text{OBJECT} & \left[\begin{array}{ll} \textit{paper} & \\ \text{AUTHOR} & \boxed{1} \end{array}\right] \end{array}\right]$$

We get by unifying (1') with (2) the strict reading (2'), while we get the sloppy reading (2") by using (1").

$$(2') \left[\begin{array}{ll} \textit{revise} & \\ \text{AGENT} & \textit{bill} \\ \text{OBJECT} & \left[\begin{array}{ll} \textit{paper} & \\ \text{AUTHOR} & \textit{john} \end{array}\right] \end{array}\right]$$

$$(2'') \left[\begin{array}{ll} \textit{revise} & \\ \text{AGENT} & \boxed{1}\,\textit{bill} \\ \text{OBJECT} & \left[\begin{array}{ll} \textit{paper} & \\ \text{AUTHOR} & \boxed{1}\,\textit{bill} \end{array}\right] \end{array}\right]$$

3 Default Unification on Substructures

While classical studies focus on parallism, the importance of non-parallel and fragmentary ellipsis is shown by empirical analysis of spoken dialog (cf. (Fernandz and Ginzburg, 2002)). The focus of an elliptic utterance often has no direct counterpart in the antecedent, which makes Rooth's matching condition not directly applicable (cf. (Rooth, 1992), (Hardt and Romero, 2001)). Grammatically required verbs (e.g., the semantically weak verb *do*) may be omitted in dialog ellipsis. In German spoken language, this is also possible in single and sequential utterances of one speaker.

We take an example from TALKY, which is the appointment management multimodal dialog system that was developed in the framework of DIAMOD. The reaction of the system to the first utterance of the user is not necessarily important, because users often proceed with (3) without waiting for the system's answer (i.e., by *barge in*) or without paying much attention to the system's reaction (in case of an experienced user).

1. USER: Ich möchte einen Termin eintragen. (*I want to enter an appointment*)
2. SYSTEM: *presents a new appointment entry*
3. USER: mit Schmid (*with Schmid*)

We achieve the following two representations of the user's utterances.

$$(1) \left[\begin{array}{ll} \textit{want} & \\ \text{AGENT} & \textit{user} \\ \text{TOPIC} & \left[\begin{array}{ll} \textit{enter} & \\ \text{AGENT} & \textit{system} \\ \text{OBJECT} & \textit{appointment} \end{array}\right] \end{array}\right]$$

$$(3) \left[\begin{array}{ll} \textit{thing-with-participant} & \\ \text{PARTICIPANT} & \textit{schmid} \end{array}\right]$$

"Matching" cannot be achieved by assuming that there is a hidden attitude connected to very utterance which could be inserted.

Instead, we search for "matching" nodes with comparable types before normal default unification is applied: *thing-with-participant* unifies with *appointment*, which leads to:

$$\left[\begin{array}{ll} \textit{want} & \\ \text{AGENT} & \textit{user} \\ \text{TOPIC} & \left[\begin{array}{ll} \textit{enter} & \\ \text{AGENT} & \textit{system} \\ \text{OBJECT} & \left[\begin{array}{ll} \textit{appointment} & \\ \text{PARTICIPANT} & \textit{schmid} \end{array}\right] \end{array}\right] \end{array}\right]$$

In principle, it is quite possible that thing-with-participant describes a certain (collective) type of agents. In this case, the processing would produce an ambiguity. In the DIAMOD system as in many other dialog systems, the agent role is usually restricted to the user and to incarnations of the system.

It is not alway posible to find a matching type. In this case we try to find paths that connect terminal nodes of the antecedent structure with the top node of the elliptic structure. It is important, that such connection paths do not introduce new structures corresponding to verbal complements or subordinated sentences.

If no match is achieved we get simply the new structure back, which is the normal result of applying default unification to non-matching structures.

4 Task Completion as a Barrier for Elliptic reference

In the following example (taken from Talky), the user performs her specification in a stepwise manner by extensively using ellipsis.

1. USER: Ich möchte am Montag ein Treffen eintragen. (*I want to enter a meeting at monday*)
2. SYSTEM: *Presents an empty appointment entry*
3. USER: Im Bananensaal (*In the "banana room"*)
4. SYSTEM: *Presents appointment entry with* banana room
5. USER: Ich meine im Raum Leibniz (*I mean in room* Leibniz)
6. SYSTEM: *Presents appointment entry with room* Leibniz
7. USER: um sechs Uhr (*at six o'clock*)
8. SYSTEM: *Presents appointment with room and begin time 6 a.m*
9. USER: abends (*at the evening*)
10. SYSTEM: *Presents appointment with room and begin time 6 p.m*

Some information has been corrected or clarified, but there was no information removed *implicitly*. Locally, most steps could be considerd as a case of fragmentary elaboration of the preceding utterances (cf. (Schlangen and Lascarides, 2003)). But this classification depends on more general properties of the dialog. When the task is finished, the availability of old information has changed:

1 USER: Bitte das Treffen jetzt eintragen. (*Please enter now the meeting*

2 SYSTEM: *Indicates that the meeting is stored*

3 USER

  a Bitte jetzt ein Treffen am Freitag eintragen *please enter a meeting at Friday now*

  b Und am Freitag. *And at Friday!*

  c Am Freitag. *At Friday!*

In case of 3a, the the old information is removed. With 3b we recognize that the activity (entering a meeting in a schedule) is still available for being inherited elliptically, while further information, accumulated before, is no longer relevant. If the user wants to keep the more elements of the old information, she has to use anaphoric references, e.g.,

4 Und dasselbe am Freitag (*And the same at Friday*).

The elliptic reading in (3b) is very clear, (3c) is rather an incomplete utterance that has to be clarified. This is also quite different from the specification phase of the meeting.

Task completion is a barrier for fragmentary elaboration. [4] After task completion, an elliptic relation has to be be marked (e.g., by clue words as *und* (*and*). Even then, ellipsis does not refer to the whole information accumulated before, but rather to the utterance that introduced the specification phase of the preceding task.

5 Information Browsing

Typically, information request are answered after every user input without a lengthly specification phase. As in the case of elliptic specifications, clarification dialog does not affect the elliptic relations between subsequent user queries. If the system actively proposes an action, this will be different. Browsing means to vary requests either because it is not clear in advance which information is relevant, how exactly it can be obtained, or because the user wants to gather broad information in some area.

In browsing dialog, ellipsis is controlled by relations between the informational content of the antecedent and the elliptic utterance. According to our remarks at the beginn of the section, we omit the reactions of the system in the subsequent examples.

By a group, we understand a collection of information that is orthogonal to other information. By

[4] The reader may recognize a certain similarity of the considerations in this section with the approach of (Grosz and Sidner, 1986). An example: We restrict ourself to some remarks: Grosz & Sidner focus on the segmentation of discourse along the hierarchical structure of a task, while we focus on problems concerning repetition (this section) and variation of tasks (next section). Grosz & Sidner are mainly concerned with anaphoric reference while we are concerned with ellipsis and related implicit inheritance of information. In our approach, structural relations between information is as much important as aspects concerning the processing of tasks. Furthermore, we discuss problems in relation to a special resolution mechanism, i.e., default unification.

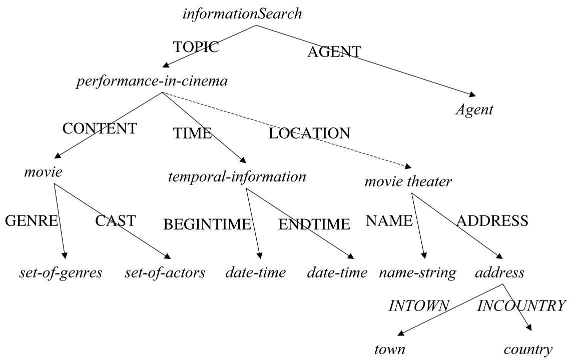

Figure 1: *Ontology for Searching Information about Performances (simplified)*

orthogonal we mean independent and not "competing". For instance, we consider TIME, LOCATION and CONTENT as basic groups of the information that belongs to a performance. Independence is not a sufficient criterion. Actor and genre are independent, but as our examples may show there are considered as competing. We have no formal means to recognize a group. The knowledge about groups has to be provided.

We use the term *information element* (IE) of a feature structure as follows: An IE consists of two parts: a role path and a semantic content. Differing from the usual definition of paths (Carpenter, 1992), a role path is a sequence of alternating types and features (T1F1...TnFn with Types Ti and Features fi). The semantic content is expressed by the substructure which is identified by applying the subsequence of the features of the role path (accordingly to standard definition). Role paths can be translated directly in an obvious way into feature structures.

We speak of an terminal information element (TIE), if the substructure is a type without further specification. A TIE is atomic, if its semantic content is atomic. We represent TIEs as extended role paths by taking the type which expresses their semantic content as last element of the path.

Two TIEs (or IEs) are of the same sort, if their role path has a common prefix. Two TIEs are of the same terminal sort, if their role paths are identical.

TIE1 is more general as TIE2 if TIE1 subsumes TIE2. TIE1 subsumes TIE2 if the subsumption relation holds between their translations to feature structures. It will turn out that this definition is to narrow and does not cover the intuitive meaning of being more general.

The TIEs in elliptic expressions are usually less specific or have a shorter role path than the TIEs in the antecedent. Subsequently we assume that the matching process (as described in former sections) has already been applied and that the TIEs of the elliptic expression are extendend by appending the role path from the root of the antecedent to the matching node. Otherwise we could not correctly determine if an IE is subsumed by another or if they belong to the same group etc.

We only consider readings of elliptic expressions that amount to a new request, ignoring other readings of elliptic expressions, e.g., as positive or negative feedback.

1 USER: Welch filme laufen heute Abend in Saarbrücken? (*Which movies are on today evening at Saarbrücken?*)

2 USER: Welche Krimis kommen? / Krimis! / (?)Und Krimis! (*Which crime (*movies*) are on? / Crime (*movies*) / And crime (*movies*)*)

3 USER:

  a Welche Sience fiction filme laufen? / Und science fiction? / Science fiction! (*Which science fiction movies are on? / And science fiction? / science fiction*)

  b Sind Science fiction filme dabei? (*Are science fiction movies among them?*)

In (2) the general information *movies* (i.e., the TIE informationSearch:TOPIC:-performance−in−cinema:CONTENT:movie) is replaced by the coresponding concrete information *crime movies* (i.e., the TIE informationSearch:-TOPIC:performance−in−cinema:CONTENT:-movie:GENRE:crime). All other information belongs to different groups and is retained. In (3) the information *crime movies* is replaced by information of the same terminal sort. The specification *crime movies* is deleted. GENRE is a set-valued feature. Note that set-valued features act quite differently depending on the context (information browsing vs. task specification). If the information *crime* should be retained, this has to be indicated, e.g. by an anaphorical relation to the result of query (2) as is done in (3b). The reading of (2) and (3) is not affected by the form of the ellipsis, but the strong indication of parallelism that is expressed with "Und Krimis" (*"And crime (movies)"*) seems not acceptable due to the proper subsumption relation between movies and movies with genre crime.

1 USER: Welche Science fiction laufen heute abend in Saarbrücken?. (*Which science fiction*

(movies) are on today evening at Saarbrücken)

4 USER:

a Mit Bruce Willis? (*With Bruce Willis?*)

b Und mit Bruce Willis / Welche filme mit Bruce Willis laufen (*And with Bruce Willis? / Which movies with Bruce Willis are on*)

In (4), the new information element Bruce Willis does not belong to the same terminal sort as any element in the antecedent, but by contributing to the specification of movies it belongs to the same group as science fiction. It is a competing element of 'science fiction', and its effect on the information element 'science fiction' is a mixture of the effect of elements of the same sort and elements of a different group (as may be expected). 4b is an an acceptable utterance in this context and it has the effect of deleting the genre information, while 4a without explicit ellipsis indication could also count as adding a specification.

1 USER: Welche Krimis kommen heute abend in Saarbrücken (*Which crime (movies) are on today evening at saarbrücken?*)

5 USER:

a Und in Saarlouis (*And at Saarlouis?*)

b Welche filme laufen in Saarlouis? (*Which movies are on at Saarlouis?*)

In (5) the information Saarbrücken is replaced by an information of the same terminal sort. 5a has the reading *crime movies in Saarlouis*. In 5b crime movies is replaced by a more general information. This is an indication that the specification *crime* should be removed. But *Welche filme* (*which movies*) has two other (less preferred) readings: an anaphoric reading *which (of those) movies are (also) running at Saarlouis*, or even an elliptic (or E-type) reading *which crime movies are on at Saarlouis*. That *movie* is more general than *crime movie* can directly inferred from examining the ontology, i.e. by subsumption.

1 USER: Welche Krimis kommen heute abend in Saarbrücken (*Which crime (movies) are on?*)

6 USER:

a Und im Scala (*And at the Scala (*movie theater*))

b Welche filme laufen im Scala (*Which movies are on at the Scala movie theater*)

In (6), Saarbrücken is replaced by a more concrete information of the sort location. The Scala movie theater is expected to be in Saarbrücken except for Scala is a aforementioned cinema in another town. The readings are quite similar to (5). But there is one difference: assume (1) gets an empty result. Than (5a) is still appropriate while (6a) is quite odd. (5b remains (slightly) ambiguous, while (6b) has only one reading. The problem with these findings is, that we cannot recognize by subsumption that Scala is more specific than Saarbrücken.

In information browsing, the relations between the information elements contained in the antecedent and the information elements provided by the ellipsis expression are relevant for resolution.

Concrete Information Rule If the elliptic expression contains a more concrete TIE than the antecedent, old specifications that belong to another group are retained.

General Information Rule If the elliptic expression contains more general information than the antecedent, then the general information tends to be understood as deleting the corresponding concrete Information. The more general TIE introduces a choice points for default unification. Default unification has to produce a reading (usually the more likely one) that accepts general information elements as potential barriers for default unification and removes old information which is beyond the barriers.

Same Sort Rule If the elliptic expression contains information of the same terminal sort, the old information is deleted, even if the information elements belong to a set-valued feature, except it is made explicit that the feature should be added.

Competing Information Rule If the elliptic expression indicates parallelism and contains "competing" information of the same group, but not the same terminal sort, the old information is deleted. Otherwise, competing information can be understood as adding a further specification.

Negative Result Condition Ambiguous readings are sensible for the result of the antecedent query. Negative (empty) results excludes readings that make the specification more concrete.

We only consider relations between an antecedent query and a subsequent elliptic query. We do not discuss here relations that come into play if a longer history is considered. The examples show, that default unification has to be controlled by relations between information elements.

6 Conclusion and Problems

We presented an approach for the resolution of non-parallel ellipsis by default uni£cation, which is inherently a parallel method. We discussed the dependence of ellipsis interpretation on the state of the dialog in respect to task processing, but also on relation between the informational content of antecedent structures and elliptic structures, which leads to a removal of inforamation, which is up to now not considered in studies on ellipsis. We also addressed the interplay of these dependencies with indications of parallism that are customarily viewed as the main factors of ellipsis interpretation. and We demonstrated how these insights are realized by using default unification as efficient base processing.

A topic of further research is the relation of general and concrete information. For instance, the ontology shown in figure 1 resembles the ontology used in SmartKom. The location of a cinema is specified by using a common format for addresses, in which country and town are on the same level and the name of the object not directly related to the address. Formally (if groups are already defined) these informations would considered as *competing*. This would prevent the Scala movie theater to be transferred to Saarlouis (in most cases the competition criterion would exclude this possibility, for a certain type of elliptic expression it would recognize an ambiguity). But the criterion would also delete the specification that the Scala cinema is in Saarbrücken if the information element Scala is introduced elliptical after asking a question about Saarbrücken. This kind of problems is not exclusively a problem of locations.

- Wo läuft Matrix? (*Where is Matrix on?*)
- Western? / Wo laufen Western? (*Where are western (movies) on*

The phrase “Western” shows no indications for parallelism, hence the competition criterion would in this case accept the reading that the user is looking for a western named Matrix.

As a practical solution, we introduced a rule that comprises that names and other “identifiers” of individuals are considered as being more concrete than any other information elements, but further exploration of the problem is necessary.

Also, the study of larger pieces of dialog as considered here is an important topic of further research.

At several occasions we noticed that anaphoric relations interact with elliptic relation. The interaction of anaphors and ellipsis is another important topic of research.

References

J. Alexandersson and T. Becker. 2003. The formal foundations underlaying overlay. In *Proceedings of the Fifth International Workshop on Computational Semantics*, Tilburg.

J. Calder. 1991. Some notes on priority unions. In *ACQILEX Workshop on Default Inheritance in the Lexicon*, Cambridge, England.

B. Carpenter, editor. 1992. *The Logic of Typed Feature Structures*. Cambridge University Press.

B. Carpenter. 1993. Skeptical and credulous default uni£cation with applicationns to templates and inheritance. In T. Briscoe, A. Copestake, and V. de Pavia, editors, *Inheritance, Defaults and the Lexicon*, pages 13–37. Cambridge University Press.

R. Fernandz and J. Ginzburg. 2002. A Corpus Study of Non-sentential Utterances in Dialogue. *Tratement Automatique de Languages*, 43(2).

J. Grosz and C.L. Sidner. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12:175–204.

C. Grover, Ch. Brew, S. Manandhar, and M. Moens. 1994. Priority union and generalization in discourse grammars. In *32nd. Annual Meeting of the Association for Computational Linguistics*, pages 17 – 24, Las Cruces, NM. Association for Computational Linguistics.

D. Hardt and Maribel Romero. 2001. Ellipsis and the structure of discourse. In *Sinn und Bedeutung VI*, Osnabrück, Germany.

J.R. Hobbs and A. Kehler. 1997. A Theory of Parallelism and the Case of VP Ellipsis. In P. R. Cohen and W. Wahlster, editors, *Proceedings of the 35th Annual Meeting of the ACL*, pages 394–401. Association for Computational Linguistics.

R. Kaplan. 1987. Three seductions of computational psycholinguistics. In P. Whitelock, H. Somers, P. Bennett, R. Johnson, and M. McGee Wood, editors, *Linguistic Theory and Computer Applications*, pages 149–188. Academic Press.

H. Prüst, R. Scha, and M. van den Berg. 1994. Discourse grammar and verb phrase anaphora. *Linguistics and Philosophy*, 17:261–327.

M. Rooth. 1992. A theory of focus interpreattion. *Natural Language Semantics*, 1:75–116.

D. Schlangen and A. Lascarides. 2003. The Interpretation of Non-Sentential Utterances in Dialogue. In *Proceedings of the 4th SIGdial Workshop on Discourse and Dialogue*, Sapporo, Japan.

M. Streit. 1999. The interaction of speech, deixis, and graphics in the multimodal of£ce agent talky. In P. Dalsgaar, CH. Lee, P. Heisterkamp, and R. Cole, editors, *Proceedings of the ESCA Tutorial and Research Workshop on Interactive Dialog in Multi-Modal Systems*.

M. Streit. 2001. Why are multimodal systems so dif£cult to build. In Harry Bunt and Robbert-Jan Beun, editors, *Cooperative Multimodal Communication*, volume 2155 of *Lecture Notes in Computer Science*. Springer.

D. Traum. 1994. *A Computational Theory of Grounding in Natural Language Conversation*. Ph.D. thesis, Computer Science Dept., U. Rochester.

W. Wahlster, N. Reithinger, and A. Blocher. 2001. Multimodal Communication with a Life-Like Character. In *Proc. of the 7th Proc. European Conf. on Speech Communication and Technology*.