13th European Summer School in Logic, Language and Information
(ESSLLI2001)

Information Structure, Discourse Structure and Discourse Semantics

Workshop Proceedings

Edited by
Ivana Kruijff-Korbayová and Mark Steedman

August 20th–24th 2001
The University of Helsinki, Helsinki, Finland

Workshop Programme Committee

Ivana Kruijff-Korbayová (co-chair; The University of the Saarland)
Mark Steedman (co-chair; The University of Edinburgh)

John Bateman (Universität Bremen)
Daniel Büring (University of California, Los Angeles)
Elisabet Engdahl (Göteborg University)
Bart Geurts (University of Nijmegen and Humboldt University Berlin)
Eva Hajičová (Charles University, Prague)
David Traum (University of Southern California)
Maria Vilkuna (Research Institute for the Languages of Finland)
Bonnie Webber (The University of Edinburgh)

Workshop Programme

Session 1:	**Monday, August 20th**	
17.00–17.35	Opening Mark Steedman and Ivana Kruijff-Korbayová: *Two Dimensions of Information Structure in Relation to Discourse Structure and Discourse Semantics*	(p. 1)
17.35–18.10	Rick Nouwen: *Complement set reference*	(p. 115)
18.10–18:45	Laurent Prévot: *Topic Structure in Route Explanation Dialogues*	(p. 145)
Session 2:	**Tuesday, August 21th**	
17.00–18.10	Invited talk by Enric Vallduví: *Fragments in Information Packaging*	(p. 15)
18.10–18:45	Svetlana McCoy: *Connecting Information and Discourse Structure Levels through "Kontrast": Evidence from Colloquial Russian Particles -TO, ZHE, and VED'*	(p. 85)
Session 3:	**Wednesday, August 22th**	
17.00–17:35	Carla Umbach: *Relating Contrast and Contrastive Topic: a focus-semantic analysis of "but"*	(p. 175)
17.35–18:10	Nobo Komagata: *Entangled Information Structure: Analysis of Complex Sentence Structures*	(p. 53)
18.10–18:45	Ivana Kruijff-Korbayová and Bonnie Webber: *Information Structure and the Semantics of "otherwise"*	(p. 67)
Session 4:	**Thursday, August 23th**	
17.00–18.10	Invited talk by Livia Polanyi: *The Relationship of Discourse Structure to Information Structure*	(p. 11)
18.10–18:45	Massimo Poesio and Barbara di Eugenio: *Discourse Structure and Anaphoric Accessibility*	(p. 129)
Session 5:	**Friday, August 24th**	
17.00–17:35	Katherine Forbes, Eleni Miltsakaki, Rashmi Prasad Anoop Sarkar, Aravind Joshi and Bonnie Webber: *Discourse Parsing with a Lexicalized Tree-Adjoining Grammar (D-LTAG)*	(p. 17)
17.35–18:10	Jennifer Spenader: *Presupposition or Anaphora: Constraints on Choice of Factive Complements in Spoken Discourse*	(p. 161)
18:10–18:45	Closing Discussion	

Preface

The call for papers for this workshop set as its aim to provide a forum for recent research on IS-sensitive discourse processing and to facilitate contact, exchange of ideas and cooperation between researchers in computational and theoretical linguistics and logicians, who are investigating these issues. The mission of the workshop was described as follows:

Information structure (IS) is construed broadly here as comprising structural and semantic properties of utterances relating to the discourse status of their content, the actual and attributed attentional states of the discourse participants, and the participants' prior and changing attitudes (knowledge, beliefs, intentions, expectations, etc.). This broad view of IS is meant to subsume notions like focus, presupposition, given vs. new, theme vs. rheme and the various dichotomies such as topic vs. comment or focus, ground or background vs. focus etc.

While *discourse structure* (DS) is more difficult to define, there is at least agreement that coherent discourse (multi-sentence dialogue or monologic text) is more than a sequence of propositions, just as sentences are more than sequences of words. In discourse, both explicit and implicit devices signify links between sentences, between groups of sentences, and between elements within sentences, and in turn, carry additional elements of discourse semantics. We are thus taking DS broadly, to cover all aspects of the internal organisational structure of a discourse. DS thus subsumes notions such as segmentation, relations between segments (informational and intentional), anaphoric relations, modal subordination, discourse topic, thematic progression, etc.

Understanding IS in light of DS and vice versa is justified on more than just theoretical grounds: experience with applications such as translating telephony and interactive query-answering makes it painfully clear that a theory relating IS and DS is essential for accurate Natural Language Processing. Fortunately, formal accounts addressing these issues have started to emerge and some, to be embodied in computational models of discourse processing. Further development and adaptation into practical systems will require expertise from linguistics, logic and computation. ESSLLI provides an appropriate forum for fostering collaboration between interested researchers across the relevant areas.

We hope that the workshop participants will transcend the difficulties caused by proliferating terminologies, and concentrate instead on investigating the interactions between IS and DS, with the goal of understanding how sentence-level semantic devices that make up IS symbiotically serve the needs of discourse cohesion

Paper presented at the ESSLLI 2001 Workshop on
Information Structure, Discourse Structure and Discourse Semantics
Edited by Ivana Kruijff-Korbayová and Mark Steedman.

and coherence. We seek contributions advancing beyond descriptive frameworks towards an explanatory account of how IS and DS, in whatever framework, interact to refer to and update a dynamically evolving representation of discourse context.

In response to the call for papers, we received 24 submissions, which have been reviewed by the programme committee, and of which we have decided to accept eight for presentation and three as a reserve. An important criterion in selecting among the extremely strong submissions was the extent to which they addressed linkages *between* the areas of information structure, discourse structure and discourse semantics. In addition to the authors of the submitted papers, we have invited two distinguished researchers to give keynote speeches, one from the area of information structure, and one from the area of discourse understanding.

We have aimed to assemble a coherent set of papers, addressing closely related issues from various perspectives, such that the exchange of ideas among the workshop participants could be maximized. The programme of the workshop consists of five sessions, each containing two to three paper presentations followed by discussion. Here, we briefly overview the sessions:

Session1 (Monday, August 20^{th}) In an introductory talk, Mark Steedman attempts to lay out the evolution of IS in various approaches. Then, two papers follow which represent different perspectives:

Rick Nouwen's contribution represents the recent optimality-theoretic strand in formal discourse semantics/pragmatics; he addresses the problem of interpreting plural pronouns with quantified antecedents, and formalizes the resolution of these anaphors in terms of the marked-unmarked opposition between the reference set and the complement set.

Laurent Prévot's contribution is couched in the Segmented Discourse Representation theory approach; based on an investigation of a speech corpus of route explanation dialogues, it is proposed to consider in parallel the dialogue structure (reflecting rhetorical and intentional structuring) and the topic structure (reflecting the construction of a common ground between the participants).

Session 2 (Tuesday, August 21^{th}) *Enric Vallduví* addresses the difference between fragmentary vs. full answers to questions, in terms of combining information structure and a notion of accessibility for propositional descriptions in context.

Svetlana McCoy also discusses the relation between information structure and the cognitive status of information in context; she explains the distribution of several Russian particles in terms of cognitive status and a notion of "kontrast", which she concludes serves to connect the level of clause and discourse structure.

Session 3 (Wednesday, August 22^{th}) Three papers point to a tight relation between the IS-partitioning of an utterance and the relations between utterances in discourse context:

Carla Umbach proposes an analysis of "but" which combines its IS-sensitivity and its discourse function of marking denial with respect to an issue under discussion; she discusses some differences in distribution between "and" and "but" in overinformative responses to questions, and concludes that "but" marks a change of the issue under discussion in the response w.r.t. the question.

Nobo Komagata compares alternative analyses of IS-partitioning in complex sentences, and supports the view that IS is a non-recursive, matrix-level partitioning. With regard to NLP systems, he concludes that information structure and discourse structure processing need to proceed in parallel at some point.

Ivana Kruijff-Korbayová and Bonnie Webber discuss examples where the interpretation of "otherwise" as a discourse anaphor requires access to conditions derived on the basis of the IS-partitioning of the relevant antecedent. It is argued that information structure affects discourse context by influencing what contexts are available for the interpretation of subsequent utterances.

Session 4 (Thursday, August 23th) *Livia Polanyi* discusses discourse structure and its relation to information structure, from both the speaker's and the hearer's perspective, and argues for an integrated approach to discourse continuity.

Massimo Poesio and Barbara di Eugenio address the relationshhip between discourse structure and anaphoric reference: using a corpus of tutorial dialogues, they evaluate the commonly accepted correlation between anaphoric accessibility and discourse structure. Their study confirms an earlier proposal that intentional relations are relevant for modelling attentional state, while informational relations are not. But they also found evidence that attentional state cannot be properly modelled as a stack, but only as a list.

Session 5 (Friday, August 24th) *Katherine Forbes et al.* present an implementation of a parsing system for a lexicalised Tree-Ajoining Grammar for discourse, integrating sentence and discourse level processing. The system is based on the assumption that the compositional aspects of semantics at the discourse-level parallel those at the sentence-level. This coupling is achieved by factoring away inferential semantics and anaphoric features of discourse connectives.

Jennifer Spenader's empirical study of factive verbs in spoken English has revealed that an overwhelming majority of them occur with full propositional complements expressing new information, that needs to be accommodated. She compares presupposed complements with abstract object anaphora, and draws implications for a general theory relating presupposition, anaphora and accommodation.

Finally, in the closing discussion, an attempt is made to summarize the implications of the contributions presented at the workshop.

Ivana Kruijff-Korbayová and Mark Steedman
July 2001

Table of Contents

Introduction: Mark Steedman and Ivana Kruijff-Korbayová: *Two Dimensions of Information Structure in Relation to Discourse Structure and Discourse Semantics* 1

Livia Polanyi, Martin van den Berg and David Ahn: *A Note on the Relationship of Discourse Structure to Information Structure* 11

Enric Vallduví: *Fragments in Information Packaging* 15

Katherine Forbes, Eleni Miltsakaki, Rashmi Prasad Anoop Sarkar, Aravind Joshi and Bonnie Webber: *Discourse Parsing with a Lexicalized Tree-Adjoining Grammar (D-LTAG)* 17

Jeanette Gundel, Michael Hegarty, Kaja Borthen: *Information structure and pronominal reference to clausally introduced entities* 37

Nobo Komagata: *Entangled Information Structure: Analysis of Complex Sentence Structures* 53

Ivana Kruijff-Korbayová and Bonnie Webber: *Information Structure and the Semantics of "otherwise"* 67

Valéria Molnár: *Contrast from a Contrastive Perspective* 99

Svetlana McCoy: *Connecting Information and Discourse Structure Levels through "Kontrast": Evidence from Colloquial Russian Particles -TO, ZHE, and VED'* 85

Rick Nouwen: *Complement set reference* 115

Massimo Poesio and Barbara di Eugenio: *Discourse Structure and Anaphoric Accessibility* 129

Laurent Prévot: *Topic Structure in Route Explanation Dialogues* 145

Jennifer Spenader: *Presupposition or Anaphora: Constraints on Choice of Factive Complements in Spoken Discourse* 161

Carla Umbach: *Relating Contrast and Contrastive Topic: a focus-semantic analysis of "but"* 175

Klaus von Heusinger: *Intonational Phrasing and Discourse Segmentation* 189

Introduction

MARK STEEDMAN
Informatics, University of Edinburgh, 2 Buccleuch Place, Edinburgh EH8 9LW, U.K.
steedman@cogsci.ed.ac.uk

IVANA KRUIJFF-KORBAYOVÁ
Computerlinguistik, Universität des Saarlandes, Saarbrücken, Germany
korbay@coli.uni-sb.de

Two Dimensions of Information Structure in Relation to Discourse Semantics and Discourse Structure

The following introduction is intended to do no more than to establish some basic distinctions and definitions within which we believe all the theoretical positions represented at the workshop can be compared and/or reconciled and their often close relationships to each other be understood.

1 Some Tentative Definitions

The title of the workshop implies a three-way distinction between Information Structure, Discourse Semantics, and Discourse Structure. While on occasion these terms have been used interchangeably or in overlapping senses, we believe that it is useful to distinguish them as follows:

- Information Structure is a **Sentence Internal** partition of the information in an utterance according to its relation to the discourse context under dichotomies such as topic/comment, theme/rheme, given/new, focus/background, etc. Such categories are essentially **Referential** in nature.
- Discourse Semantics centrally concerns the nature of the contextual model, and the entities in it to which Information Structural categories relate, in the form of alternative sets, filecards, presupposed propositions, etc.

Paper presented at the ESSLLI 2001 Workshop on
Information Structure, Discourse Structure and Discourse Semantics
Edited by Ivana Kruijff-Korbayová and Mark Steedman.

- Discourse Structure concerns the **Inter-clausal** relations of explanation, elaboration, exemplification, and illocutionary force that hold between successive utterances of a discourse or dialog, supporting inference about the domain and purposes of the discourse.

2 A Computational Model

We have tried in setting up the workshop to keep computational models of discourse and dialogue in focus. In particular we believe it may be useful to think of these distinctions in terms of some extension to Grosz and Sidner's ((Grosz and Sidner(1986))) computational account. Specifically:

- Information structure seems to belong in what they call "linguistic structure"—essentially, the grammar.
- Its Discourse Semantic extensions seem to concern the "Attentional State" component. (Although the structural representation of this state as a stack of focus spaces seems questionable, and the ontology of types that may inhabit focus spaces must be greatly extended—see Poesio and Di Eugenio, and Forbes et al. below.)
- Other rhetorical and illocutionary aspects of Discourse Structure appear to relate to their "Intentional Structure" component.

3 The Two Dimensions of Information Structure

The terminology that is used to describe Information Structure and its semantics is simultaneously various, and under-formalized. Yet it seems that all definitions have some elements in common. They all draw at least one of the following distinctions: (i) a "topic/comment" or "theme/rheme" distinction between the part of the utterance that relates it to the purpose of the discourse, and the part that advances the discourse; (ii) a "given/new" distinction, between parts of the utterance—actually, words—which contribute to distinguishing the content from other alternatives that the context makes available and those parts that are common to all of them.

There are differences among the theories of course. Some, like Halliday's ((Halliday(1967))), view these two distinctions as orthogonal, applying at independent levels of structure. Others in the Bolingerian tradition, like Vallduví, McCoy, Molnár, and ourselves in the present volume view them as different aspects of a single level of structure. An important issue that further differentiates the Bolingerian theories is that of whether pitch accent corresponds to a single contrastive notion of focus applying to both theme and rheme, or whether "contrastive focus" is a distinct notion, applying to explicitly mentioned entities and associated with topic or theme alone (see Molnár, and Umbach, below). Many of these questions

await fuller empirical and corpus based studies: we welcome the studies of this kind by Spenader, McCoy, Prévot, and Poesio and Di Eugenio.

There are further similarities: while some of the theories leave the associated discourse semantics at an intuitive level, the theories which do address formal semantic issues all tend to use some version of "update" semantics, such as the file-card semantics and segmented DRSs of the Kampo-Heimian synthesis (see Prévot, von Heusinger, and Gundel et al., below), or the Alternative semantics of Rooth and Büring (see Nouwen, Komagata, Umbach, and Kruijff-Korbayová and Webber, in the present collection).

The diagram on the next page displaying our view of the influences and terminological dependencies in theories of Information Structure and the associated Discourse Semantics may help to make these commonalities more obvious.

4 An Example that Exemplifies Both Dimensions

The terminology that we use ourselves may serve to illustrate these two dimensions at work. (We take the terms theme and rheme from Firbas and Bolinger, rather than Halliday, for our topic/comment distinction, and use Dahl's background and focus for the Hallidean given/new distinction, and apply these elements in a single level of information structure.)

(1) Q: I know that Mary likes the man who wrote the musical.
But who does she ADMIRE?

A: (Mary ADMIRES) (the woman who DIRECTED the musical)

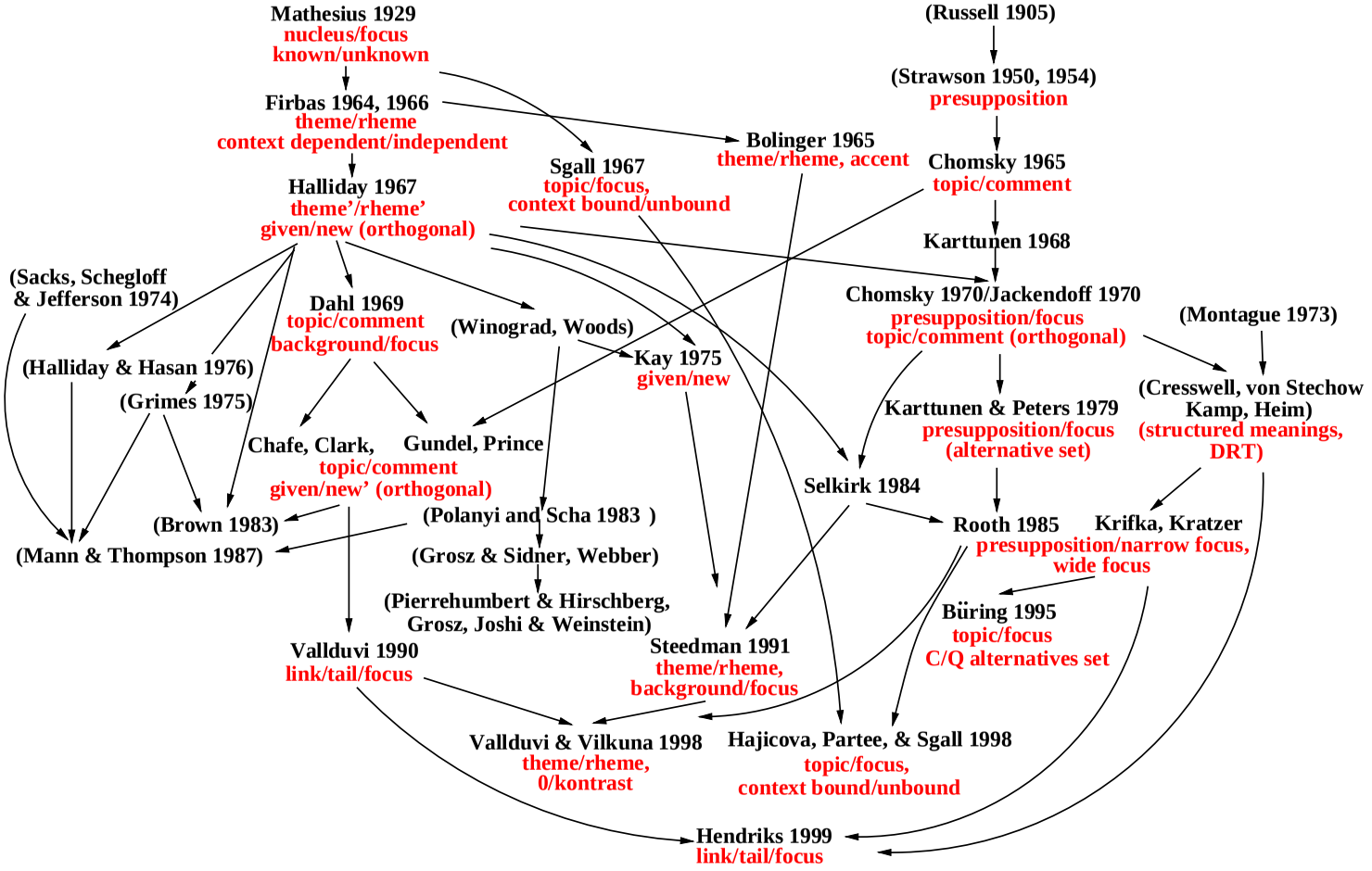
$\underbrace{\qquad}_{\textit{Background}}\ \underbrace{\text{L}+\text{H}*\text{LH\%}}_{\textit{Focus}}\ \underbrace{\qquad}_{\textit{Background}}\ \underbrace{\text{H}*}_{\textit{Focus}}\ \underbrace{\text{LL\%}}_{\textit{Background}}$

$\underbrace{\qquad\qquad}_{\textit{Theme}}\ \underbrace{\qquad\qquad\qquad\qquad}_{\textit{Rheme}}$

In English we have claimed that Information Structure is homomorphic to Intonation Structure, and that the contour described in Pierrehumbert's (Pierrehumbert(1980)) notation as L+H* LH% is one of the "theme tunes" that identify the intonational phrase *Mary admires* as the theme of this utterance, while H* LL% is a "rheme tune". Within both the theme and the rheme, the presence of one or more pitch accents identifies words which contribute to distinguishing that theme and/or that rheme from other themes and rhemes that the context affords. Thus we follow (Bolinger(1965)) and (Lambrecht(1994)) in viewing the role of accent in English as a single undifferentiated contrastive meaning applying to both informational components.

We think of that context as an updatable database including two Roothian alternatives sets, respectively called the "Theme Alternative Set" and the "Rheme Alternative Set"—cf. "Alternative Semantics" (Rooth(1985); Rooth(1992)), (Büring(1995)).

In Combinatory Categorial Grammar (Steedman(2000)), Information Structure and the associated structured meanings are associated directly with surface syntactic derivational structure, uniting Information Structure, Intonation Structure, and Surface Syntax in a single module of Grammar.


Mathesius 1929
nucleus/focus
known/unknown
(Russell 1905)
(Strawson 1950, 1954)
presupposition
Firbas 1964, 1966
theme/rheme
context dependent/independent
Bolinger 1965
theme/rheme, accent
Chomsky 1965
topic/comment
Sgall 1967
topic/focus,
context bound/unbound
Halliday 1967
theme'/rheme'
given/new (orthogonal)
Karttunen 1968
(Sacks, Schegloff
& Jefferson 1974)
Dahl 1969
topic/comment
background/focus
(Winograd, Woods)
Chomsky 1970/Jackendoff 1970
presupposition/focus
topic/comment (orthogonal)
(Montague 1973)
(Halliday & Hasan 1976)
Kay 1975
given/new
(Grimes 1975)
(Cresswell, von Stechow
Kamp, Heim)
(structured meanings,
DRT)
Karttunen & Peters 1979
presupposition/focus
(alternative set)
Chafe, Clark,
topic/comment
given/new' (orthogonal)
Gundel, Prince
Selkirk 1984
(Brown 1983)
(Polanyi and Scha 1983)
Rooth 1985
Krifka, Kratzer
presupposition/narrow focus,
wide focus
(Mann & Thompson 1987)
(Grosz & Sidner, Webber)
(Pierrehumbert & Hirschberg,
Grosz, Joshi & Weinstein)
Büring 1995
topic/focus
C/Q alternatives set
Vallduvi 1990
link/tail/focus
Steedman 1991
theme/rheme,
background/focus
Vallduvi & Vilkuna 1998
theme/rheme,
0/kontrast
Hajicova, Partee, & Sgall 1998
topic/focus,
context bound/unbound
Hendriks 1999
link/tail/focus

5 Some Questions for the Workshop

In terms of this we hope uncontentious systematization of the terms of reference of the workshop, the following questions can be offered for consideration among the many others raised by the presenters at the workshop:

- What is the Discourse Semantics of Information Structure?
- What part does Information Structure play in Discourse Structure?
- How does Information Structure relate to Intonation Structure in English and related languages?
- How does Information Structure relate to Syntax and Semantics in the languages of the world?
- *Your question here*

6 Some Answers that Emerged

- *What is the Discourse Semantics of Information Structure?*
  There seemed to be two flavors of semantics: Forbes et al. , Korbayová & Webber, and McCoy used forms of Alternative Semantics. Prévot; Vallduví, and Poesio were in the Kampo-Heimian Structured Meaning tradition. This appears to be a mainly notational difference and the systems seem in some sense equivalent.
- *What part does Information Structure play in Discourse Structure?*
  Many Discourse structural and Discourse Semantic phenomena discussed at the workshop appeared to depend at least in part on on information-structural representations in the context, including anaphora, presupposition, and the meaning of connectives like *but*, *otherwise* and *although*, and discourse particles like the Russian *zhe, ved'*. (Polanyi; Spenader; Poesio, Nouwen; Korbayová & Webber; Umbach)
- *How does Information Structure relate to Intonation Structure in English and related languages?*
  Several papers claimed or were compatible with the view that Intonation Structure in English is identical to an underspecified form of Infomation Structure, and in particular that intonational phrases correspond to information units. (Vallduví; Umbach; Prévot; McCoy; Komagata; Steedman.) There were interesting questions raised about dialect variation, the consistent specification of information structural role by prosodic tones, and the status of All-Theme utterances (Kamp; Gundel.)

- *How does Information Structure relate to Syntax and Semantics in the languages of the world?*
  There was general agreement that the same pair of orthogonal distinctions– in Vallduvíand Vilkuna's terms, theme/rheme and Kontrast/ground—were at work in other languages, marked by other devices such as morphology, verbal particles, word-order etc. (Prévot; Komagata; McCoy; Molnár; von Heusinger). There was an open question how does Umbach's *quæstio* fit in—is it a theme?

7 Further Questions that Emerged

- *How are Contextual Representations to be Managed?*
  Contexts, whether viewed as collections of alternative sets or of DRSs or filecards have a complex structure. Alternative representations of Attentional State were discussed, including those related to computational dialog management, such as that proposed in the TRINDI framework, which maintains lists of QUDs. Themes and QUDs seem closely related. The question of representing changing salience of discourse objects was also discussed.

- *The Role of Accomodation*
  Accomodation was central to a number of proposals. What types of dicourse entity can accomodate, and when is accomodation blocked?

- *The Role of Information Structure Partitioning*
  What is the linguistic unit of discourses structure? It seemed clear the themes and rhemes could be discontinuous within a single utterance. The question of whether any such units were obligatory remains unresolved. The question of how to reliably identify informational structural units in real texts remained problematic, but a number of tests were proposed including looking at effects of systematically alterning intonation (and equivalent parameters in other languages) and the effect on the interpretation of IS-dependent discourse connectives like *although,* focus particles, etc.

- *What is the best empirical agreement for investigating these questions?* There was general agreement that dialog, rather than extended exposition, was the domain that really stretched the theories.

Bibliography

Bolinger, Dwight. 1965. *Forms of English*. Cambridge, MA: Harvard University Press.

Brown, Gillian. 1983. "Prosodic Structure and the Given/New Distinction." In Anne Cutler, D. Robert Ladd, and Gillian Brown, eds., *Prosody: Models and Measurements*, 67–77. Berlin: Springer-Verlag.

Büring, Daniel. 1995. *The 59th Street Bridge Accent: On the Meaning of Topic and Focus*. Ph.D. thesis, Universität Tübingen. Publ. as (**?**).

Chafe, Wallace. 1976. "Givenness, Contrastiveness, Definiteness, Subjects, Topics and Points of View." In Charles Li, ed., *Subject and Topic*, 25–55. New Tork: Academic Press.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1971. "Deep Structure, Surface Structure, and Semantic Interpretation." In Danny Steinberg and Leon Jakobovits, eds., *Semantics*, 183–216. Cambridge: Cambridge University Press.

Clark, Herbert, and Susan Haviland. 1977. "Comprehension and the Given-New Contract." In R. Freedle, ed., *Discourse production and comprehensio*, 1–40. Norwood, NJ: Ablex.

Cresswell, M.J. 1973. *Logics and Languages*. London: Methuen.

Dahl, Östen. 1969. *Topic and Focus: a Study in Russian and General Transformational Grammar*. Göteborg: Elandres Botryckeri.

Firbas, Jan. 1964. "On Defining the Theme in Functional Sentence Analysis." *Travaux Linguistiques de Prague*, 1, 267–280.

Firbas, Jan. 1966. "Non-thematic Subjects in Contemporary English." *Travaux Linguistiques de Prague*, 2, 229–236.

Grimes, Joseph. 1975. *The Thread of Discourse*. Mouton.

Grosz, Barbara, Aravind Joshi, and Scott Weinstein. 1995. "Centering: A Framework for Modeling the Local Coherence of Discourse." *Computational Linguistics*, 2, 203–225.

Grosz, Barbara, and Candace Sidner. 1986. "Attention, Intention and the Structure of Discourse." *Computational Linguistics*, 12, 175–204.

Gundel, Janet. 1974. *The Role of Topic and Comment in Linguistic Theory*. Ph.D. thesis, University of Texas, Austin.

Hajičová, Eva, Barbara H. Partee, and Petr Sgall. 1998. *Topic-Focus Articulation, Tripartite Structures, and Semantic Content*. Dordrecht: Kluwer.

Halliday, Michael. 1967. "Notes on Transitivity and Theme in English, Part II." *Journal of Linguistics*, 3, 199–244.

Halliday, Michael, and Ruqaiya Hasan. 1976. *Cohesion in English*. Longman.

Heim, Irene. 1982. *The semantics of definites and indefinite noun phrases in English.*. Ph.D. thesis, University of Massachussetts, Amherst.

Hendriks, H. 1998. "Links without Locations." In F. Hamm and E. Zimmermann, eds., *Linguistische Berichte, Sonderheft 9/1998: SEMANTIK*. Opladen: Westdeutscher Verlag.

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Kamp, Hans. 1984. "A Theory of Truth and Semantic Representation." In Jeroen Groenendijk, Theo Janssen, and Martin Stokhof, eds., *Truth, Interpretation, and Information*, 1–41. Dordrecht: Foris.

Karttunen, Lauri. 1968. "What Makes Definite Noun Phrases Definite?" Tech. rep., Rand Corporation.

Karttunen, Lauri, and Stanley Peters. 1979. "Conventional Implicature." In Choon-Kyu Oh and David Dinneen, eds., *Syntax and Semantics 11: Presupposition*, 1–56. New York: Academic Press.

Kay, Martin. 1975. "Syntactic Processing and Functional Sentence Perspective." In Roger Schank and Bonnie Nash-Webber, eds., *Theoretical Issues in Natural Language Processing, (Supplement)*, 12–15. MIT, June.

Kratzer, Angelika. 1991. "The Representation of Focus." In Arnim von Stechow and Dieter Wunderlich, eds., *Semantics: An International Handbook of Contemporary Research*, 825–834. Berlin: de Gruyter.

Krifka, Manfred. 1991. "A Compositional Semantics for Multiple Focus Constructions." *Cornell Working Papers in Linguistics*, 10, 127–158.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.

Mann, William, and Sandra Thompson. 1987. "Rhetorical structure theory: A framework for the analysis of texts." Tech. Rep. RS-87-185, Information Science Institute, Marina del Rey, CA.

Mathesius, Vilem. 1929. "Functional Linguistics." In Josef Vachek, ed., *Praguiana: Some Basic and Less Well-known Aspects of the Prague Linguistics School*, 121–142. Amsterdam: John Benjamins. Publ. 1983.

Montague, Richard. 1973. "The Proper Treatment of Quantification in Ordinary English." In Jaakko Hintikka, J. M. E. Moravcsik, and Patrick Suppes, eds., *Approaches to Natural Language: Proceedings of the 1970 Stanford Workshop on Grammar and Semantics*, 221–242. Dordrecht: Reidel. Reprinted in Montague 1974, 247-279.

Pierrehumbert, Janet. 1980. *The Phonology and Phonetics of English Intonation*. Ph.D. thesis, MIT. Distributed by Indiana University Linguistics Club, Bloomington.

Pierrehumbert, Janet, and Julia Hirschberg. 1990. "The Meaning of Intonational Contours in the Interpretation of Discourse." In Philip Cohen, Jerry Morgan, and Martha Pollack, eds., *Intentions in Communication*, 271–312. Cambridge, MA: MIT Press.

Polanyi, Livia, and Remko Scha. 1983. “The Syntax of Discourse.” *Text*, 3, 261–270.

Prince, Ellen. 1981. “Towards a Taxonomy of Given-New Information.” In Peter Cole, ed., *Radical Pragmatics*, 223–256. New York: Academic Press.

Rooth, Mats. 1985. *Association with Focus*. Ph.D. thesis, University of Massachusetts, Amherst.

Rooth, Mats. 1992. “A Theory of Focus Interpretation.” *Natural Language Semantics*, 1, 75–116.

Russell, Bertrand. 1905. “On Denoting.” *Mind*, 14, 479–493.

Sacks, Harvey, Emanuel Schegloff, and Gail Jefferson. 1974. “A Simplest Semantics for the Organization of Turn–Taking for Conversation.” *Language*, 50, 696–735.

Selkirk, Elisabeth. 1984. *Phonology and Syntax*. Cambridge, MA: MIT Press.

Sgall, Petr. 1967. “Functional Sentence Perspective in a Generative Description.” *Prague Studies in Mathematical Linguistics*, 2, 203–225.

von Stechow, Arnim. 1991. “Focusing and Backgrounding Operators.” In Werner Abraham, ed., *Discourse Particles: Descriptive and Theoretical Investigations on the Logical, Syntactic and Pragmatic Properties of Discourse Particles in German*, 37–84. Amsterdam: John Benjamins.

Steedman, Mark. 1991. “Structure and Intonation.” *Language*, 67, 262–296.

Steedman, Mark. 2000. “Information Structure and the Syntax-Phonology Interface.” *Linguistic Inquiry*, 34, 649–689.

Strawson, Peter. 1950. “On Referring.” *Mind*, 59, 320–344.

Strawson, Peter. 1954. “A Reply to Mr. Sellars.” *Philosophical Review*, 63, 216–231.

Vallduví, Enric. 1990. *The Information Component*. Ph.D. thesis, University of Pennsylvania.

Vallduví, Enric, and Maria Vilkuna. 1998. “On Rheme and Kontrast.” In Peter Culicover and Louise McNally, eds., *Syntax and Semantics, Vol. 29: The Limits of Syntax*, 79–108. San Diego, CA: Academic Press.

Webber, Bonnie. 1988. “Tense as Discourse Anaphor.” *Computational Linguistics*, 14, 61–73.

Winograd, Terry. 1972. *Understanding Natural Language*. Edinburgh: Edinburgh University Press.

Invited Talk:
A Note on the relationship of discourse structure to information structure

LIVIA POLANYI
FXPAL, 3400 Hillview, Palo Alto, CA 94304
polanyi@pal.xerox.com

MARTIN VAN DEN BERG
FXPAL, 3400 Hillview, Palo Alto, CA 94304
vdberg@pal.xerox.com

DAVID AHN
Department of Computer Science, University of Rochester, Rochester, NY 14627
davidahn@cs.rochester.edu

1 Introduction

Although it is generally accepted that a sentences information structure (IS) is determined by its relationship to previous text, the question of how to establish the appropriate discourse context for IS assignment is never raised. Analyses of IS normally assume that that sentence is a question and the target sentence is the answer to that question [1]. The assumption is always that the prior context for a target sentence is the immediatly preceding sentence. While this assumption may be a convenient convention for investigating the subtler aspects of intrasential IS, serious problems arise in analyzing actual texts when the critical context for assigning IS is not part of the immediately preceding sentence. For these cases, it is necessary to have a theory of how discourse structure (DS) constrains the choice of possible contexts for determinining sentential IS. This note addresses how discourse structure and information structure (IS) are related and sketches an integrated approach to the phenomena of discourse continuity.

In brief, our claim is that DS constrains the set of possible constituents in a discourse that can provide the relevant context for structuring information in a target sentence, while IS critically constrains DS ambiguity. To develop an argument

Paper presented at the ESSLLI 2001 Workshop on
Information Structure, Discourse Structure and Discourse Semantics
Edited by Ivana Kruijff-Korbayová and Mark Steedman.

to support this claim, we rely upon the Linguistic Discourse Model (LDM) as a theory of DS [2], and follow Steedman in assuming that sentence IS consists of a theme and rheme, each further subdivided into background and focus [3, 4].

2 Discourse Structure and Information Structure

Under the LDM, the representation of a discourse is constructed incrementally using information in the surface structure of incoming utterances together with discourse construction rules and inference over the meaning of the utterances to recursively construct an open-right tree of discourse constituent units (DCUs). This tree indicates which units are accessible for continuation and anaphora resolution. All nodes on an LDM tree are first class objects containing structural and semantic information; terminal nodes correspond to the strings of the discourse, while non-terminals are constructed nodes labeled with a discourse relation (coordination, subordination, or n-ary). IS is represented at terminals and non-terminals as well. A C-node inherits the generalization of the themes of its constituent nodes and of their rhemes. An S-node inherits the IS of its subordinating daughter directly.

Our initial hypothesis, illustrated by the example text and accompanying chart[1], is that the attachment is (1) a C-node if the theme of the main clause of the new sentence matches thematic information available at the attachment point or (2) an S-node if the theme of the main clause of the new sentence matches rhematic information available at the attachment point[2].

In analyzing a discourse, discourse syntax assigns each incoming sentence its place in the emerging discourse tree. In current approaches, lexical information, syntactic and semantic structure, tense and aspect, and world knowledge are used to infer the attachment point and relation (c.f. [5]). However, after exploiting these resources, attachment ambiguities often still remain. Given that normal language users seldom experience discourse attachment ambiguities, additional sources of information must be used in attachment decisions. We believe that the IS of both the incoming sentence and accessible DCUs provides information critical for disambiguation. The problem of identifying the target DCU that provides the context for IS assignment for an incoming sentence is analogous to anaphora resolution: the target unit must be along the right edge of the tree and therefore accessible [6].

From a discourse perspective, the IS of an incoming sentence divides it into a theme, which must be linked back to the preceding discourse, and a rheme, which need not be. Establishing a link between the theme of the main clause of a new sentence and information available at an accessible node in the tree determines

[1] Here, for the sake of presentational simplicity, constituents are assumed to be sentences. however, under the LDM, the much more finely-grained DCU segmentation conventions enable subordinate clauses to serve as attachment points for the main clauses of subsequent sentences.

[2] Here we will not discuss n-ary nodes that are used to represent the structure of dicsourse genres as well as conversational adjacency structures and logical relations further. It is enough for our purposes to mention that they follow more ad-hoc, though well-defined, rules.

the sentences attachment point. The type of attachment, coordination, subordination, or n-ary, reflects the themes relation to the IS of the DCU represented at the attachment node.

3 Analysis of an example text

In the example, themes are marked with a θ; rhemes are unmarked. Words receiving stress are in SMALL CAPS.

> 1. (Japanese people occasionally choose to eat)$_\theta$ NOODLES. 2. (Noodles are USUALLY eaten)$_\theta$ for LUNCH or a light SNACK. 3. Depending on the SEASON, (noodles might be served)$_\theta$ in a HOT SOUP or COLD like a salad. 4. (When noodles are served in a hot SOUP,)$_\theta$ VEGETABLES, TOFU, and MEAT are ALSO found within the soup. 5. Several TYPES of noodles (are eaten IN JAPAN.)$_\theta$ 6. (UDON)$_\theta$ are THICK, WHITE noodles made fresh from wheat flour and are USUALLY served with a hot SOUP. 7. (SOBA)$_\theta$ are THIN BUCKWHEAT noodles which are FIRMER than udon. 8. (They can be served in a SOUP like UDON,)$_\theta$ but are USUALLY served as a COOL dish in the SUMMER. 9. (RAMEN)$_\theta$ are very thin, CURLY wheat noodles served as a QUICK meal or a LATE night SNACK. 10. (Noodles are eaten)$_\theta$ as a VARIATION for the daily MEAL.

Sentence #	1	2	3	4	5
Theme	Japanese people eat	Noodles ... eaten	Noodles ... served	Noodles ... hot soup	Eaten IN JAPAN
Rheme	NOODLES	LUNCH SNACK	SEASON ... HOT SOUP COLD	VEGETABLES TOFU MEAT ALSO	TYPES
Attachment	N/A	1	2	3	(S2 – (S 3–4))
Relation	N/A	S	S	S	C

Sentence #	6	7	8	9	10
Theme	UDON	SOBA	The ... SOUP ... [UDON]	RAMEN	Noodles ...
Rheme	THICK WHITE USUALLY SOUP	THIN BUCK-WHEAT	USUALLY ... COOL SUMMER	VERY CURLY QUICK LATE	VARIATION ... MEAL
Attachment	5	6	7	(C 6–(S 7–8)) (S 5 ...))	(C (S 2 ...)– (S 5 ...))
Relation	S	C	S	C	C

As the chart indicates, (1)-(4) exhibit theme-rheme chaining, resulting in nested subordinations. For (5), the appropriate context for IS assignment is provided by (2), with a theme-theme link resulting in a coordination. The rheme of (5) intensionally introduces a set of *types of noodles* picked up as the theme alternative set

for (6), (7), and (9). The theme focus for each of these sentences (*udon, soba, ramen*) is presupposed to belong to this set. These sentences are therefore coordinated to each other and subordinated to (5). Processing (8) demonstrates that both DS and IS may operate autonomously. The IS of (8) is determined primarily by the conjunction but which acts with the possibility modal in its first conjunct (which provides an accessible set of possible worlds as the rheme alternative set) to construct a theme- rheme pair, while discourse attachment of (8) fulfills anaphora resolution requirements rather than IS. For (10), (5) provides the appropriate context for the IS assignment. The theme-theme link results in a coordination that pops the state of the discourse several levels.

4 Conclusion

Although the assignment of IS to a sentence depends on the DS, and the construction of the DS may depend on the IS of the units involved, the dependency between IS and DS is complementary and not circular. For the speaker, the DS provides a set of possible contexts for continuation while IS assignment is independent of DS. For the hearer, the IS of a sentence together with DS instructs dynamic semantics how rhematic information should be used to update the meaning representation of the discourse (c.f. [7]). Thus, the relationship between DS and IS reflects the different but deeply related tasks of speaker and hearer in a communicative situation.

References

[1] Vallduví, Enric. (1994) Information Packaging: A Survey, Word Order , Prosody and Information Structure Project Report. University of Edinburgh.
[2] Polanyi, L. and R. Scha (1984) A syntactic approach to discourse semantics, COLING'84.
[3] Steedman, M. (1991) Structure and intonation, Language, 68.
[4] Steedman, M. (to appear) Information structure and the syntax-phonology interface, Linguistic Inquiry, 31:4.
[5] Lascarides, A. and N. Asher (1993) Temporal interpretation, discourse relations and Commonsense Entailment, Linguistics and Philosophy, 16:5.
[6] Polanyi, L. and M. v. d. Berg (1999) Logical structure and discourse anaphora resolution, ACL99 Workshop on Discourse/Dialogue Structure and Reference.
[7] Polanyi, L. and M. v. d. Berg (1996) Discourse structure and discourse interpretation, 10th Amsterdam Colloquium.
A number of text types are distinguished by being about one person. In an obituary or letter of recommendation, for example, no other person may even be

Invited Talk:
Fragments in Information Packaging

ENRIC VALLDUVÍ
Dept de Traduccio, Universitat Pompeu Fabra
enric.vallduvi@trad.upf.es

Papers on information structure often make use of question-answer pairs like (1) and (2) to illustrate the sentential partition between rheme (focus, new info) and ground (theme, presupposition, old info, topic):

(1) a. Who likes beer?
b. [$_R$ JOHN] [$_G$ likes beer.]

(2) a. What does John like?
b. [$_G$ John likes] [$_R$ BEER.]

The idea is that the question is the exponent of some underspecified description in a context or information state. In (1a), for instance, an instantiation for the experiencer parameter is needed. The rheme in (1b) is the structural element that ”answers the question” in the sense that it expresses a value for this uninstantiated parameter, while the ground simply expresses the pieces of meaning in this description that are already available from context. Let us call the pieces of meaning that make a contribution to context the ”update (potential)” and the pieces of meaning that are already available in context the ”base”.

The fact is often neglected that (1b) and (2b) are not discourse-natural answers to (1a) and (2a). In fact, the most natural answers to (1-2a) are those in (3) (and also those in (4), see below):

(3) a. JOHN. (answer to (1a))
b. BEER. (answer to (2a))

(4) a. JOHN does. (answer to (1a))
b. He likes BEER (answer to (2a))

A full account of information packaging should take this dialogical evidence into account. The difference between (1-2b) and (3) on the other is clearly not in their update potential. Still, we need to account for the fact that (3) is a more natural answer to (1-2a) than (1-2b).

Paper presented at the ESSLLI 2001 Workshop on
Information Structure, Discourse Structure and Discourse Semantics
Edited by Ivana Kruijff-Korbayová and Mark Steedman.

The proposal defended here is that the difference between (1-2b) and (3) is in the nature of the base, namely, in the accessibility status of the base. In (1-2) the immediate, explicit questions make the base fully accessible at the time of utterance, so a ground is not needed. This follows from a view of grounds as context-anchoring elements. The ground is the structural expression of a base, but such expression is only needed when the anchoring of the update to a given base is not self-evident at the time of utterance. In question-answer exchanges, where the base is fully accessible from context, grounds are superfluous; this is why rheme-only fragment utterances, which express only the update, are so common as direct answers to questions. Of course, the ground may be redundantly used as an anchor in contexts where it is not needed, but this often gives rise to secondary interpretive effects, sometimes triggered by the violation of Gricean maxims.

What we need is a theory that allows us to establish some notion of accessibility for propositional descriptions in context, akin to the notion of accessibility assumed in accounts of anaphora resolution. Grounds are unnecessary/disfavored when the base is within some local attentional domain: the link between the update potential expressed as rheme and the base in the context proceeds as the result of certain default dynamics of dialogue. However, grounds become necessary when the base of a given update is outside this local attentional domain: the unification of update and base cannot take place without the explicit anchoring effected by the ground. In a sense, then, rheme-only fragments are to rheme-ground sentences what pronouns are to full definite descriptions in a theory of referential accessibility.

The answer sentences in (4) above behave more like sentence fragments than full answers with respect to information packaging (their dialogic distribution is closer to (3) than (1-2b)). Clearly, the examples in (4) are not syntactic fragments, but their behavior is accounted for if they are treated as informational fragments; in other words, the answers in (4) do not have a ground. The weak pronoun and the verbal elements that appear in (4) are there exclusively for morphosyntactic reasons and do not carry out the anchoring function that grounds carry out. This leads us to propose too that (5b) is as much of an informational fragment as (6b) is (although (5b) is obviously not a syntactic fragment):

(5) a. How does he feel about Bill?
    b. He LOVES him.

(6) a. Who does John love?
    b. BILL.

Both (5b) and (6b) are answers that do not need a ground to make the connection explicit between their update potential and the pertinent base in context.

D-LTAG System - Discourse Parsing with a Lexicalized Tree Adjoining Grammar

KATHERINE FORBES
University of Pennsylvania
forbesk@linc.cis.upenn.edu

ELENI MILTSAKAKI
University of Pennsylvania
elenimi@linc.cis.upenn.edu

RASHMI PRASAD
University of Pennsylvania
rjprasad@linc.cis.upenn.edu

ANOOP SARKAR
University of Pennsylvania
anoop@linc.cis.upenn.edu

ARAVIND JOSHI
University of Pennsylvania
joshi@linc.cis.upenn.edu

BONNIE WEBBER
University of Edinburgh
bonnie@dai.ed.ac.uk

ABSTRACT.

We present an implementation of a discourse parsing system for a lexicalized Tree-Ajoining Grammar for discourse, specifying the integration of sentence and discourse level processing. Our system is based on the assumption that the compositional aspects of semantics at the discourse-level parallel those at the sentence-level. This coupling is achieved by factoring away inferential semantics and anaphoric features of discourse connectives. Computationally, this parallelism is achieved because both the sentence and discourse grammar are LTAG-based and the same parser works at both levels. The approach to a LTAG for discourse has been developed by (Webber & Joshi 1998; Webber et al. 1999b) (among others) in some recent papers. Our system takes a discourse as input, parses the sentences individually, extracts the basic discourse units from the sentence derivations, and reparses the discourse with reference to the discourse grammar.

1 Introduction

All work on discourse starts from the premise that discourse meaning is more than the sum of its parts (i.e., its constituent sentences or clauses). The question is how to get there. Work in the tradition of *Rhetorical Structure Theory* (RST) (Mann & Thompson 1988) – both in interpretation (Marcu 2000) and generation (Mellish et al. 1998) – views the additional meaning solely in terms of discourse relations that hold between adjacent text spans, treating discourse connectives as signalling types of discourse relations. How the basic text spans are assigned an interpretation, and how that interpretation might contribute to discourse meaning apart from discourse relations, is largely ignored.

Not so in more formal work on discourse semantics (Gardent 1997; Polanyi & van den Berg 1996; Scha & Polanyi 1988; Schilder 1997; van den Berg 1996), which takes seriously a compositional process of deriving discourse meaning from

Paper presented at the ESSLLI 2001 Workshop on
Information Structure, Discourse Structure and Discourse Semantics
Edited by Ivana Kruijff-Korbayová and Mark Steedman.

the meaning of its parts. However, this work (1) only makes use of two mechanisms for deriving discourse meaning from the meaning of its parts – compositional semantics and inference – and (2) treats the process by which discourse derives compositional aspects of meaning as being entirely separate from how clauses do so. Both of these are the focus of the approach developed in (Webber & Joshi 1998; Webber et al. 1999b). In this approach, it is argued that certain aspects of discourse meaning are better seen as deriving from anaphoric and presuppositional properties of lexical items, and that this is facilitated through a uniform lexicalised treatment of both clausal syntax and semantics and discourse syntax and semantics. This paper presents an initial implementation of a discourse parsing system (D-LTAG) that draws on the insights of this latter approach.

Our motivation for using this approach is to explore the hypothesis that the boundary between sentence level structure and discourse level structure is not a sharp one. Sentence level structure supports compositional semantics even though there are other semantic aspects, such as anaphoric relations (e.g., intrasentential links for pronoun reference) and inferential interpretation (e.g., interpretation of compound nouns) that need to be accounted for. In the same way, discourse level structure is also viewed as supporting compositional aspects of semantics, while allowing for other interpretive components to be added on for a complete semantics for discourse – e.g., for determining anaphoric and inferential interpretation. Thus, we pursue the idea that the formal device used for deriving the structural descriptions at both levels is the same, while noting that at the discourse level, the device may have less generative power. In addition, we also illustrate that the described architecture for discourse parsing allows for a smooth transition from sentence level to discourse level processing and for the use of a single parser at both levels.

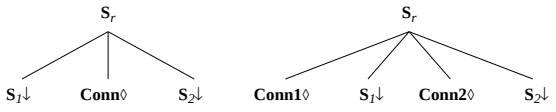
In Section 2, we discuss the LTAG framework for discourse description, as outlined in (Webber & Joshi 1998). Section 3 presents a discussion of our methodology for determining the structure and semantics of discourse connectives, accompanied with a case study of the discourse connective *however*. In Section 4, we describe the architecture of our system, and discuss various issues that arose during the implementation. Section 5 discusses some of the advantages of our system, in particular, with respect to the close link between sentence level and discourse level semantics. In Section 6, we compare our system with some other approaches, in particular with those that use some variant of TAG for describing discourse structure, such as (Gardent 1997) and (Schilder 1997), and those that attempt to automate the derivation of discourse structure, such as (Marcu 2000)s.

2 The Framework: A Lexicalized Tree Adjoining Grammar for Discourse

The D-LTAG system is based on the approach to a lexicalized TAG for discourse, as described in (Webber & Joshi 1998). A LTAG for discourse posits two kinds

of elementary trees: *initial* trees, which encode predicate-argument dependencies, and *auxiliary* trees, which are recursive and modify and/or elaborate elementary trees. All structural composition is achieved with two operations, *substitution* and *adjunction.* Clauses connected by a subordinating conjunction form an initial tree whose compositional semantics is determined by the semantic requirements of the subordinate conjunction (the predicate) on its arguments (the clauses). Auxiliary trees are used for providing further information through adjunction. They can be anchored by adverbials, by conjunctions like *and*, or may have no lexical realization. Furthermore, a discourse predicate may take all its arguments structurally, as in the case of subordinating conjunctions, or anaphorically, by making use of events or situations available from the previous discourse, as in the case of *then.*[1] This division between the compositional part of discourse meaning (projected by the tree structures) and the non-compositional contributions due to general inferencing and anaphora is a key insight of the approach to an LTAG for discourse. It simplifies the structure of discourse and extends compositional semantic representations from the sentence level to the discourse.

Figure (7.1a) shows one initial tree in the grammar.[2] We treat connectives anchoring this tree as discourse predicates which require two clausal arguments. In general, such trees are anchored by subordinating conjunctions such as *because*, *when* etc. A corollary of the structure of elementary trees in the discourse grammar is that discourse connectives are allowed discourse initially only if they anchor an initial tree. A second initial tree is shown in Figure (7.1b). As suggested in previous work (Webber & Joshi 1998), this tree reflects dependencies in parallel constructions and is projected by pairs of connectives such as *on the one hand ... on the other hand....* (As noted in this previous work, both members of the pair need not be realized in the surface string.)



(a) (b)

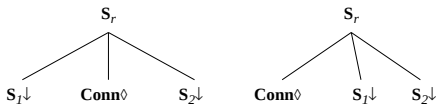
Figure 7.1: Initial Trees in D-LTAG

As in the sentence-level grammar, trees in the D-LTAG grammar are grouped into *tree families*, which are taken to reflect surface clause order variation realized with preposed and postposed subordinate clauses. Trees belonging to the same

[1]Our use of the term *anaphora* does not include anaphoric relations such as those established by pronouns and definite descriptions. Accounts of these relations have been actively pursued in other discourse-oriented semantic theories such as DRT (Kamp 1981) and Dynamic semantics (Groenendijk & Stokhof 1991). Obviously, a full account of the phenomenon of anaphora in discourse will have to take these into account. But they are not our present concern.

[2]In all the elementary trees shown in the paper, "$\diamond$" marks the anchor of the tree, "$\downarrow$" marks the substitution nodes, and "$*$" marks the adjunction nodes. Subscripts are used to distinguish non-terminal nodes with the same label.

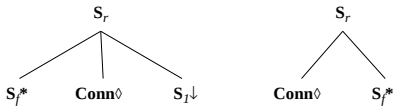
family share the same predicate-argument dependencies. One such tree family is shown in Figure 7.2, anchored by connectives like *because.*



(a) (b)

Figure 7.2: Tree Family in D-LTAG

The second type of elementary trees consist of *auxiliary trees,* which introduce recursion and serve to extend or modify a description of the previous discourse. There are two kinds of auxiliary trees, shown in Figure 7.3.



(a) (b)

Figure 7.3: Auxiliary Trees in D-LTAG

The tree in Figure 7.3(a) is anchored by connectives that simply continue the description conveyed by the structure to which the tree adjoins. Other aspects of meaning that relate the two arguments are derived anaphorically or inferentially (e.g., based on the relationship between the tense/aspect of the two arguments (Hitzeman et al. 1995; Kehler 1994; Kehler 2000; r Lascarides & Asher1993; Webber 1988). The anchor of this tree can also remain lexically unrealized, when it is used to connect adjacent clauses without overt connectives, such as "*Mary walked towards the car. The door was open*". The tree in Figure 7.3(b) is selected by connectives whose first argument is resolved anaphorically and the second argument is the interpretation of the clause they adjoin to. We say more about this in the next section.

3 Determining Tree Structures for Discourse Connectives

In the previous section, we defined the elementary trees included in the D-LTAG grammar. The next crucial step in lexicalizing a Tree-Adjoining Grammar is determining which trees or family of trees are selected by a discourse connective. In previous work (Webber et al. 1999b,Webber et al. 1999c), it was shown that the connectives *then, for example* and *otherwise* are best treated as anaphoric, anchoring trees of type 7.3(b). But for some other connectives, such as *however*, it was less clear whether they are structural or anaphoric.

In what follows, we take the view that the lexicalization of trees is an empirical question and we describe the methodology we adopt to determine the structures

lexicalized by connectives. When in doubt about the structure of a certain connective, we start with the hypothesis that the arguments of the connective are realized structurally. This is because, from a computational point of view, it would be less interesting to start with the assumption that arguments are resolved anaphorically. Assuming that all connectives find their arguments anaphorically would probably be adequate to characterize all predicate-argument relationships on the discourse level. However, it would not shed much light on those aspects of structural organization that are relevant to language structure and presumably contribute to the efficiency of the inferencing processes required in the interpretation of the discourse. (This motivation is inspired by (Joshi & Kuhn 1979)). Predicates which find their arguments structurally define a domain of locality for structural dependencies and constrain the interpretation of discourse in a computationally efficient way, as is the case for verb predicates at the sentence level syntax.

On empirical grounds, the diagnostic we use to test if a connective is structural is crossed structural dependencies. The current XTAG grammar for English does not lead to crossed dependencies as they seem to be unnecessary at the sentence level for English. We make a similar assumption for the discourse level and conclude that a connective defines a domain of structural locality only when such domains do not cross tree nodes.[3]

A Case Study: *However* For the connective *however*, our first assumption is that *however* anchors the structural auxiliary tree, shown in Figure 7.3(a). Regarding its semantic contribution, we follow (Knott 1996) and (Lagerwerf 1998) in assuming that *however* presupposes a defeasible rule holding between a generalization of its first argument and a generalization of the negation of its second argument, and asserts that the rule fails to hold in this case (see (Webber et al. 1999a) for a formalization of the rule). To investigate if both arguments are realized structurally, we conducted a corpus study of the connective. We identified 71 tokens of *however* from the Brown corpus and located the two arguments of the connective for each token. In 69 out of the 71 instances, the position of both arguments in the discourse was consistent with the structural hypothesis: one argument was realized in the sentence containing the connective, and the second argument was realized either in the immediately preceding sentence (58 instances) or in an immediately preceding chunk of text (11 instances). In both cases, attachment to the previous discourse did not yield crossing of tree nodes.

The remaining 2 cases were of two kinds. One, exemplified in (1), involved an argument that was not directly realized in the previous discourse. Rather, the presupposed defeasible rule could only be seen as holding between rather complicated

[3]However, until we have accumulated ample empirical evidence, such conclusions are tentative and subject to revision. Also, it would be very interesting to investigate languages which allow crossed dependencies at the sentence level (e.g. Dutch) and examine whether in those languages crossed dependencies are also permissible on the discourse level. Our conjecture is that this will not be the case.

generalizations which would have to be *inferred* from the two arguments. Here we take the defeasible rule to be something like "If the speaker/writer makes an apparently negative comment about a book, then his/her attitude is negative towards it."

(1) a. If this new Bible does not increase in significance by repeated readings throughout the years, it will not survive the ages as has the King James Version.

b. *However*, an initial perusal and comparison of some of the famous passages with the same parts of other versions seems to speak well of the efforts of the British Biblical scholars.

In the other case, *however* appeared to make no semantic contribution to the discourse, other than simple continuation. This is shown in (2).

(2) a. It is in this spirit which explains some of the anomalies of American Catholic higher education, in particular the wasteful duplication apparent in some areas.

b. I think for example of three women's colleges with pitiful enrollments, clustered within a few miles of a major Catholic university, which is also co-educational.

c. This is not an isolated example;

d. this aspect of the total picture has been commented upon often enough.

e. It would seem to represent esprit de corps run riot.

f. Apart, *however*, from the question of wasteful duplication, there is another aspect of the "family business" spirit of Catholic higher education that deserves closer scrutiny.

While it is clear that (2f) attaches higher up to the structure containing (2b)-(2e), it is less clear what the semantic contribution of *however* is to the interpretation of the discourse. *However* here seems to be acting similar to the discourse marker *now* (e.g., "Now, apart from the question of wasteful duplication...") (Hirschberg & Litman 1987), reinforcing the IRU cue (i.e., "apart from the question of wasteful duplication") as a signal of returning to (2a) after a conceptually embedded segment was closed off at (2e).[4]

To summarize, the corpus-based study for the connective *however* provides considerable support for the hypothesis that it finds its arguments structurally. However, as indicated by the more complex examples (1) and (2), further empirical studies will be required before a definitive conclusion is reached.

4 System Description and Implementation

In this section, we describe our initial implementation of a discourse parsing system based on a lexicalized Tree-Adjoining Grammar for discourse. Discourse structure

[4]Informationally Redundant Utterances (IRUs) are characterized as repetitions of propositions already available in the discourse. (Grosz & Sidner 1986) have shown that IRUs correspond to embedded segments. (Walker 1993) argues that, with respect to a well defined task, IRUs are used as a discourse strategy to improve the efficiency of completing a task. The distribution of IRUs in Walker's corpus indicates that IRUs function as markers of returning to a superior segment. See also (r Forbes & Miltsakaki2001) for a discussion on the collaboration of IRUs with other cues derived from Centering Theory to signal the boundaries of embedded segments.

is derived in two passes of parsing. In the first pass, the sentences in the discourse are parsed, whereas discourse parsing is done in the second pass. Without losing sight of the key ideas of the theory of an LTAG for discourse, this two pass implementation achieves a considerable simplification over a single pass of parsing, especially in terms of the parsing time and space requirements that would result from using both the sentence-level and the discourse-level grammar at once.

Figure 7.4: D-LTAG: System Description

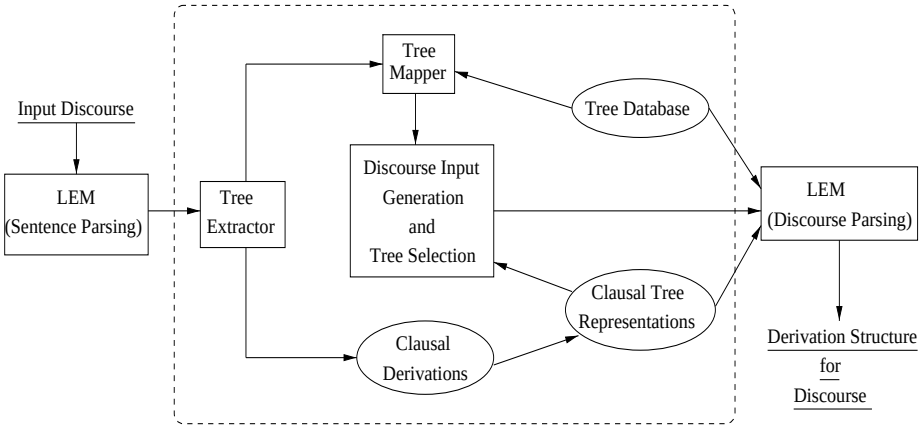


Figure 7.4 shows the overall architecture of the system. The input discourse is submitted to LEM, the **Parser**, which parses each sentence in the discourse with reference to the sentence grammar. The output derivations (one derivation each for each of the sentences) are then submitted to the **Tree Extractor**, which extracts the basic discourse constituent units from each sentence derivation. The basic discourse units constitute the elementary trees lexicalized by discourse connectives in the sentence-level grammar, and the derivation (and derived) structures associated with the clausal units.[5] In the next step, the sentence-level elementary trees anchored by the connectives are mapped by the **Tree Mapper** to their corresponding elementary trees in the discourse grammar. The discourse grammar, as specified in Section 2, is contained in the **Tree Database**. The output of the Tree Mapper, together with the clausal units and the input discourse, is then used to construct a **discourse input representation** that consists of a sequence of lexicalized trees (tree selection), with the extracted connectives and clausal units as the lexicalizing elements.[6] Finally, the discourse input, the Tree Database, and the clausal tree representations are submitted to the same **Parser**, which provides derivations for the given discourse.

[5]In this paper, we assume that clausal units correspond to the minimal tensed clause. The tensed clause is further taken to include all sentential complements, relative clauses and participial clauses. In some other discourse works, such as (Polanyi 1996), a greater range of propositional elements are regarded as the minimal units of discourse.

[6]Each extracted clause derivation is taken to be an *atomic* unit in the discourse grammar, much like a single lexical item.

In the rest of this section, we describe the different components of the system in greater detail, and discuss various issues that arose during the implementation.

PARSER (LEM).The parser is a chart-based head-corner parser (Sarkar 2000). The sentence-level grammar used by the parser is the XTAG grammar (XTAG-Group 2001), a wide-coverage grammar of English developed at the University of Pennsylvania.[7] For each sentence, each subsequent phase of the system assumes that there is exactly one derivation per sentence. Since, in general, there can be many ambiguities for each sentence in the discourse, the parser picks one derivation per sentence to pass on to subsequent processing. In the system described in this paper, the parser produces a single parse for each sentence by using heuristics that (a) decide which elementary tree to assign to each word, and (b) pick the lowest attachment between these trees. In future work, we plan to experiment with two other methods to deal with ambiguity: (1) using the parser as a statistical parser (2001) where it reports the most probable parse based on training the parser on the Penn Treebank, and (2) representing the many parses associated with each clausal unit in the sentence in a compact form (a parse forest) and representing these as the elementary units in the discourse.

TREE EXTRACTOR. The task of this component is to extract, from each sentence derivation, the clausal derivations and any elementary tree(s) anchored by discourse connectives. The Extractor first does a top-down traversal of the sentence derivation, and identifies the part of the derivation associated with any connectives. Identification of the connectives is done against a database containing a list of possible discourse connectives as well as the elementary tree(s) anchored by each of them in the sentence grammar.

The use of both lexical and structural information is necessary to correctly identify the discourse usages of connectives in the sentential derivations. That is, neither kind of information by itself is sufficient for identification. On the one hand, many elements that function as discourse connectives can also have other functions: *and* functions as a discourse connective when it conjoins clauses, as in "The dog barked *and* Mary smiled", but it can also conjoin noun phrases (among other phrasal categories), as in "Lana ate cheese *and* crackers". As a result, if we used only the lexical appearance of the elements as the identification criterion, then the *and* which conjoins non-sentential categories would be incorrectly treated as a discourse connective. Knowledge about the elementary trees associated with the discourse usage of *and* is therefore necessary to rule this out. On the other hand, it is not sufficient to only use structural information to identify discourse connectives. For example, the sentence-level grammar does not make a structural distinction between sentential adverbs that *are* discourse connectives and those that are *not*: the elementary tree in Figure 7.5 can be lexicalized both by *however*, which *is* a connective, as well as by *always*, which is not a connective. Identification of the former - and not the latter - structure can therefore be done only if the lexicalizing elements are also used.

[7] For a recent evaluation of the XTAG grammar, see (Prasad & Sarkar 2000).

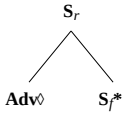


Figure 7.5: Elementary tree anchored by an adverb in the sentence-level grammar

After the identification of the connectives, the clausal derivations are detached in the sentence derivation at the substitution and/or adjunction nodes of the connective elementary tree. The result of this procedure is shown in Figure 7.6 for the derivation of the sentence *while she was eating lunch, she saw a dog*.[8]

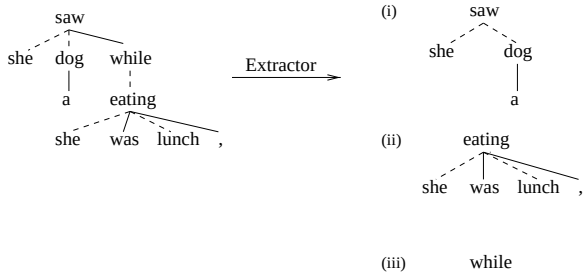


Figure 7.6: Result of Tree Extractor applied to derivation of *while she was eating lunch, she saw a dog.*

The above procedure can be shown to work on all derivations in which connectives take clauses as their arguments in the sentence-level grammar. In the surface string, this corresponds to connectives appearing at clause boundaries. However, connectives can also appear in clause-medial positions, as in Example 3. The connective *then* adjoins to the verb phrase (VP node) in the clause.

(3) Susan will *then* take dancing lessons.

Though such clause-medial connectives are posited as taking clauses as their arguments in the discourse-level grammar, we believe that their clause-internal syntax should be visible at the discourse-level description, as it is probably an indicator of Information Structure (IS).[9] The Extractor achieves this by only making a *copy* of the derivations for these connectives, and by replacing - in the derivation of the clause - the lexical occurrence of the connective by an index, to indicate its clause-internal position. The result of this procedure for example (3) is given in Figure 7.7. $\{$*then*$\}$ in 7.7(i) represents the clause-medial connective index left by the Extractor.

[8] In derivation structures, dotted lines indicate substitution and solid lines indicate adjunction. Also, note that each node is labeled by the lexical items, but these only serve as labels for the elementary tree with which they are associated.

[9] The hypothesis we are pursuing is that a clause-medial connective flags material to its left as being a contrastive theme (Steedman 2000a) – cf. Section 5.

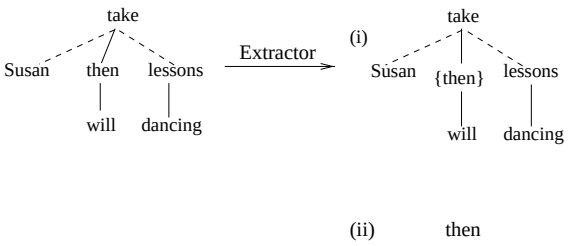


Figure 7.7: Result of Extractor applied to derivation of *Susan will then take dancing lessons*

Thus, the output of the Extractor is, for each sentence, a set of elementary trees anchored by connectives in the sentence grammar, and a set of clausal derivations. For the example discourse given in (4), (5) shows the parts of the discourse input corresponding to the extracted discourse constituent units. (b), (f), (i) and (k) are the extracted connectives, and the rest are the clausal derivations.

(4) a. Mary was amazed.
b. While she was eating lunch, she saw a dog.
c. She'd seen a lot of dogs, but this dog was amazing.
d. The dog barked and Mary smiled.
e. Then she gave it a sandwich.

(5) a. mary was amazed
b. while
c. she was eating lunch
d. she saw a dog
e. she'd seen a lot of dogs
f. but
g. this dog was amazing
h. the dog barked
i. and
j. mary smiled
k. then
l. she then gave it a sandwich

TREE MAPPER. The connective-lexicalized elementary trees that are extracted from the sentence derivations are submitted to the Tree mapper, which maps their sentence-level structural descriptions to their discourse-level structural descriptions (taken from the Tree Database). This is a crucial step in the discourse derivation because it is involved with determining what kinds of contribution(s) a given connective makes to discourse meaning, that is, what it contributes through compositional semantics, through anaphora and through inference. Furthermore, as has been pointed out in Section 3, determining the discourse structures anchored by connectives is an empirical matter. A major part of the future work in this project

is to fully determine this mapping with corpus based work on the behavior of connectives. We continue here by assuming the mappings shown in Figure 7.8 for the example discourse (4).

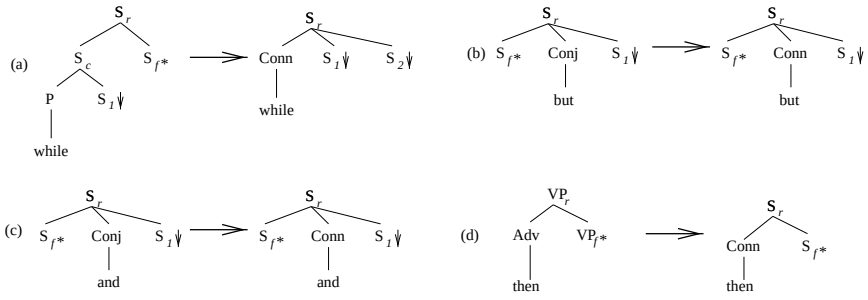


Figure 7.8: Elementary tree mappings for connectives in discourse (4)

DISCOURSE INPUT GENERATION. In the next phase of the system, the clausal derivations are first converted into elementary tree representations, which are treated as singular atomic units that can serve as the arguments of the discourse connectives. These clausal units, the input discourse, and the connective elementary trees generated by the Tree Mapper are then used to generate a discourse input representation that is essentially a sequence of lexicalized trees, where the lexicalizing elements are the connectives and the clausal units. Because of the extraction of the discourse units from the sentence derivations, and the tree mapping of the structures of connectives, tree selection ambiguities at the discourse level are minimized, and discourse parsing thus simplified.

The sequence of lexicalized trees is ordered with reference to the surface order of the input discourse (compare (4) and (5)), except for the clause-medial connectives. These are placed before the clause from within which they are copied out. This does not, however, disrupt the surface string order: the clause-internal index of these connectives, left by the Tree Extractor, succeeds in preserving the sentential surface string representation (see Figure (5i)).

This phase also includes an insertion algorithm to insert trees with an empty lexical anchor (which may still carry some feature information) into the input representation. Recall from Section 2 that the grammar contains an auxiliary tree that is used to continue the description by adjoining to the previous discourse (henceforth, continuation auxiliary trees) (Figure 7.3a). This auxiliary tree may be anchored by connectives like *and* and *or*, or remain lexically unrealized. In the extracted units shown in (5), there are only 2 overt connectives that can anchor this auxiliary tree: *and* and *but*. This means that the lexically empty trees need to be inserted at the appropriate positions in the input representation. The insertion algorithm does this by referring to the tree labels for each of the units in the (thus far created) input representation and by following a few simple insertion rules. We use the label "E" to indicate a null anchor. Alternatively, these trees can be taken to be lexicalized by the sentence-final punctuation markers.

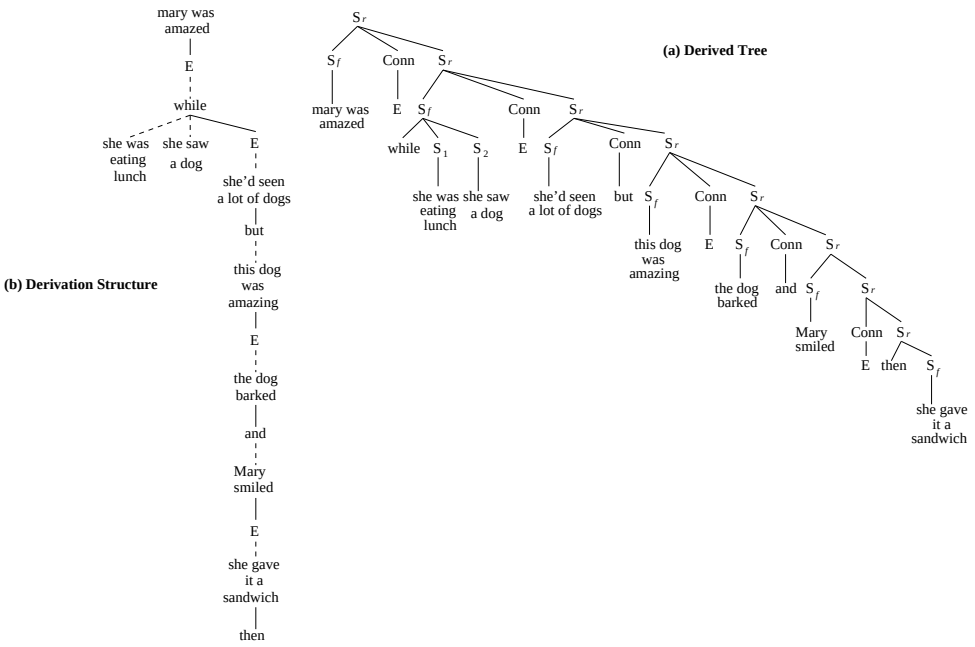


Figure 7.9: Derived Tree and Derivation Structure for Example Discourse in (4)

The sequence of lexicalized trees after insertion of the lexically empty trees is then parsed with the same parser (LEM) that we used to parse each sentence in the discourse. Since the trees are uniquely selected by the connectives and the clausal units, the resulting discourse parse contains no ambiguities that are caused by tree selection.[10] However, the system does contain attachment ambiguities caused by the continuation auxiliary trees. In the current approach, these may be resolved with an inferential component, or by statistical methods. For present purposes, we pick a unique derivation out of all the parses which satisfies the following two criteria: (a) adjunction in initial trees is only allowed at the root node; and (b) for all other permissible adjunctions, only lowest adjunction is allowed. Given the simple grammar posited in the system, these two criteria are sufficient to yield a unique derivation.

The derived tree and derivation structure for the example discourse in (4) after discourse parsing are shown in Figure 7.9.

We have also tested our system on connective rich sections of the Wall Street Journal (WSJ) from the Penn Treebank (Marcus et al. 1993). In order to avoid the problem of getting too many sentential derivations for the long and complex sentences typically found in this corpus, we used the single derivations produced by LEXTRACT (Xia et al. 2000), which takes the Treebank and Treebank-specific information and produces derivation trees for the sentences annotated in the Tree-

[10]This result obtains because the discourse grammar assumed here is quite simple, with discourse connectives projecting a single elementary tree. We note that upon further empirical investigation of the behavior of individual connectives, this may not turn out to be the case.

bank. For the WSJ discourse segment (taken from Section 21 of the WSJ corpus) given in Example 6, the derived tree and derivation structure are shown in Figure 7.10. The discourse connectives in the text are shown in bold.

(6) a. The pilots could play hardball by noting they are crucial to any sale or restructuring because they can refuse to fly the airplanes.[11]

b. **If** they were to insist on a low bid of, say $200 a share, the board mighn't be able to obtain a higher offer from other bidders **because** banks might hesitate to finance a transaction the pilots oppose.

c. **Also**, **because** UAL chairman Stephen Wolf and other UAL executives have joined the pilots' bid, the board might be able to exclude him from its deliberations in order to be fair to other bidders.

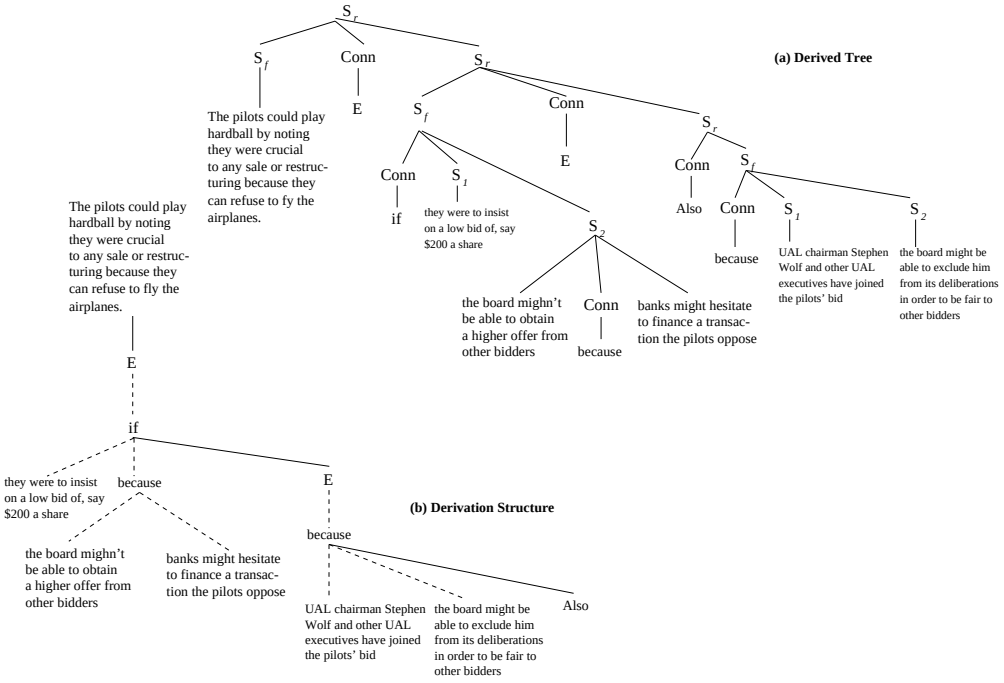


Figure 7.10: Derived Tree and Derivation Structure for WSJ Discourse in (6)

5 Discussion

5.1 Deriving Discourse Semantics

In (Joshi & Vijay-Shanker 1999) and (Kallmeyer & Joshi 2001), an approach to compositional semantics was provided for the LTAG grammar. The compositional

[11] Note that "because" in this sentence has not been treated as a connective because, initially, we have ignored sententially embedded connectives. How they contribute to discourse structure and meaning remains an important topic for future work. See also fn 5.

semantics was defined with respect to the derivation tree structure and not the derived tree. The derivation tree is a record of the history of composition of the elementary trees. Semantic representations are associated with the elementary trees and these are composed via unification. In D-LTAG, the intuition is that a similar process will be carried out at the discourse structure level using the derivation tree of the D-LTAG grammar. The details of this compositional semantics for D-LTAG have not been worked out yet. However, in general, the final representation will be essentially a *flat* structure, as is the case for the semantics of LTAG.

While each elementary LTAG tree is associated with a semantic representation, this representation does not have to reproduce the hierarchical structure of the elementary tree. The elementary tree is thus considered as a "semantic unit". This view of representing semantics directly corresponds to the so-called flat representation, which is a conjunction of formulas, where hierarchical structure appears only when needed, for example, for operators on predicates. Such a "flat" representation is also motivated in the context of generation, where one wants to start with a representation of the input which makes the minimal commitment to structure. Details on associating a flat semantics with a derived TAG tree can also be found in (1997) and (Stone et al. 2001). In (7) below, we roughly illustrate the semantic formula associated with the example discourse in (4). We refer to the semantics of the connective trees by the names of the connectives, and use "&" to represent the semantics associated with the auxiliary tree associated with continuation/elaboration. The numbered arguments of these trees are labeled either "S" for states or "E" for events, depending on their semantic content.

(7) S1 & while(S2,E3) & but(S4,S5) & (E6 & E7) & after(E7, E8)

5.2 Discourse Connectives, Information Structure, and Discourse Semantics

While the current study does not directly address interactions between information structure (IS) and discourse structure/semantics, we note that a *lexicalised grammar* for both sentences and discourse allows semantic material from both the lexicon and constructed phrases to project into both sentence-level meaning and discourse meaning. In particular, this allows sentence-level IS distinctions to be projected into discourse-level descriptions. We illustrate this by considering clause-medial adverbial discourse connectives.

Many adverbial connectives display a wide variability with respect to the position they are found in the sentence. This variability, while usually not altering the compositional semantics of the sentence, appears to correlate with IS. In particular, we are pursuing the hypothesis that a clause-medial connective indicates that material to its left serves as a *contrastive theme*. The simplest case is given in Examples 8 and 9.

(8) Mary smiled. *However*, John frowned.

(9) Mary smiled. John, *however*, frowned.

In (8), the clause-initial position of *however* is, by itself, neutral about the partition of the sentence into *theme* or *rheme* and about whether or not the theme is contrastive. In (9), on the other hand, the clause-medial position of *however* correlates with stress on "John" and appears to convey that John and Mary are elements of an alternative set (in the sense of (Rooth 1992)) – that is, that John serves as a contrastive theme.

This comes out more strongly in Examples (10) and (11). (10) is infelicitous because medial *however* flags the subject as contrastive theme, but this subject is a coreferential (unstressed) pronoun and cannot serve as a contrastive theme. Example (11), on the other hand, is fine, as the position of *however* flags the adverbial *then* as the contrastive theme (presumably in an alternative set with the time of Mary smiling).

(10) $*$ Mary smiled. Then she, *however,* frowned.

(11) Mary smiled. Then, *however,* she frowned.

Our claim here is just that, by having elements lexicalised both with respect to sentence and discourse, we can represent in the same way their contributions to both, as well as inter-relations between them. For example, in (11), not only does the clause-medial position of *however,* flag *then* as a contrastive theme (in contrast with alternatives provided in the discourse or the speech situation), but the defeasible rule presupposed (or conventionally implicated) by *however* (Knott 1996; Lagerwerf 1998) involves that specific "inertial" property – i.e., if someone smiles, they will continue to do so. *However* asserts that it fails to hold, and what happened *then* is the source of the failure. While we have not yet explored this with respect to LTAG and D-LTAG, (Bierner & Webber 2000) and (Bierner 2001) illustrate how another lexicalised grammar, Combinatory Categorial Grammar (CCG) (Steedman 2000b), can be used to express both assertional and presuppositional components of meaning associated with the sentence and with discourse, and (Steedman 2000a) shows how one can compute both IS-partitioning, its prosody and its semantics in lockstep with other aspects of meaning.

6 Comparison with Related Approaches

Recently, (Marcu 2000) developed a system for identifying rhetorical relations on unrestricted text. His system trains on a corpus annotated with rhetorical relations and utilizes correlations of surface-based features with RST relations to assign rhetorical structure to unseen text. Our system is a clear departure from this approach in two significant ways: a) we develop a system that actually parses discourse allowing the semantics to be built compositionally from the sentence to the discourse level, and b) discourse connectives are not viewed as names of relations, instead the semantics of the connectives form only a part of the compositional

derivation of discourse relations. [12]

(Gardent 1997) uses a variant of Feature-based Tree Adjoining Grammars to constuct the structure of discourse and the semantics derived from it. (Schilder 1997) extends Gardent's formalism to handle world and contextual knowledge, proposing a non-monotonic reasoning system to achieve that. Despite this similarity of the above works with our approach, both systems differ significantly from ours in the following way. Gardent's system (also Schilder's) builds the semantics of discourse compositionally but only after the semantics of the input segments and the rhetorical relation connecting every two segments is identified. However, it is not clear how the semantics of the input segments are computed since, apparently, the size of the input segment ranges from tensed clauses ('We were going to take John as a lawyer'), to complex sentences ('As we found out, either he is on sick leave') or even fragments ('Too honest for his own good, in fact').[13] In our approach, we do not assume pre-processing or segmentation of the textual input. The output from the sentence level parser is the input to the discourse parser, building up the semantics compositionally from the sentence level to the discourse level. Likewise, rhetorical relations are not assumed nor picked out from a previously defined set of relations. We are interested in those aspects of discourse interpretation that are *derived* compositionally, factoring away non-compositional semantic contributions, i.e. inferencing based on world-knowledge and anaphoric presuppositions.

7 Conclusions

Building on earlier work, we have developed and implemented a system for discourse parsing based on a lexicalized Tree-adjoining Grammar for discourse, in which the discourse connectives are the predicates, and the clauses are the arguments of these connectives. The system takes a discourse as its input, parses the sentences independently, extracts "discourse" connectives and clausal units from the output derivations of the sentences, and reparses the discourse input by submitting fully lexicalized trees to the same parser.

We have motivated a corpus study of discourse connectives in order to fully determine the semantic contribution they make to discourse, and thus, to also determine the elementary tree type(s) they lexicalize in the discourse grammar. The grammar thus developed will serve as a crucial component of the implemented system which uses this information after extracting the connectives from the sentence derivations, in order to create lexicalized elementary trees at the discourse-level.

[12] In other words, in our view, the 'name' of a rhetorical relation is ultimately derived from the compositional semantics of our system, and other non-compositional aspects of discourse meaning, i.e. the inferential component. The use of 'rhetorical relations' in discourse intepretation seems to conflate those two distinct aspects of meaning, namely compositional and inferential. In our system, we tease the two apart and derive the compositional part.

[13] The examples are from (Gardent 1997), pp.7.

The submission of the lexicalized trees as the input for discourse level parsing simplifies the parsing process considerably, and this simplification is achieved because the system integrates sentence-level processing with discourse-level processing.

Acknowledgements

We thank Miriam Eckert, Alexandra Kinyon, Alistair Knott, Bangalore Srinivas and Fei Xia for helpful discussions during the different stages in the preparation of this paper. We also thank three anonymous reviewers for their comments, which have helped in improving the content and presentation of this paper.

Bibliography

Bierner, Gann (2001). *Alternative Phrases and Natural Language Information Retrieval*. In Proc. of the 39^{th} ACL. Toulouse, France.

Bierner, Gann & Bonnie Webber (2000). *Inference through Alternative Set Semantics*. Journal of Language and Computation, 1(2):259–274.

Forbes, Katherine & Eleni Miltsakaki (to appear). *Automated Identification of Embedded Structure in Discourse Segmentation.* In Penn Working Papers in Linguistics, 2001.

Gardent, Claire (1997). *Discourse Tree Adjoining Grammars*. Claus Report 89, Saarbrücken: Univeristät des Saarlandes.

Groenendijk, Jeroen & Martin Stokhof (1991). *Dynamic Predicate Logic*. Linguistics and Philosophy, 14:39–100.

Grosz, Barbara J. & Candace L. Sidner (1986). *Attention, Intentions, and the Structure of Discourse*. Computational Linguistics, 12(3):175–204.

Hirschberg, Julia & Diane. J. Litman (1987). *Now Let's Talk About Now: Identifying Cue Phrases Intonationally.* In Proc. of the 25th ACL, pp. 163–171.

Hitzeman, Janet, Marc Moens & Claire Grover (1995). *Algorithms for Analysing the Temporal Structure of Discourse*. In Proc. of EACL, pp. 253–260. Dublin, Ireland.

Joshi, Aravind & Steve Kuhn (1979). *Centered Logic: The Role of Entity Centered Sentence Representation in Natural Language Inferencing.* In Proc. of the 6th IJCAI, pp. 435–439.

Joshi, Aravind & K. Vijay-Shanker (1999). *Compositional Semantics with Lexicalized Tree-Adjoining Grammar (LTAG): How much Underspecication is Necessary?* In H. C. Blunt & E. G. C. Thijsse (Eds.), Proc. of the Third International Workshop on Computational Semantics, Tilburg, Netherlands, pp. 131–145.

Kallmeyer, Laura & Aravind Joshi (to appear). *Factoring Predicate Argument and Scope Semantics: Underspecified Semantics with LTAG*. In Journal of Language and Computation, 2001.

Kamp, Hans (1981). *A Theory of Truth and Semantic Representation.* In J. Groenendijk, Th. Janssen & M. Stokhof (Eds.), Formal Methods in the Study of Language, pp. 277–322. Amsterdam: Mathematisch Centrum Tracts.

Kehler, Andrew (1994). *Temporal Relations: Reference or Discourse Coherence*. In Proc. of the 32nd ACL, Student Session, pp. 319–321. Las Cruces NM.

Kehler, Andrew (2000). *Resolving Temporal Relations using Tense Meaning and Discourse Interpretation.* In Martina Faller, Stefan Kaufmann & Marc Pauly (Eds.), Formalizing the Dynamics of Information. CSLI publications.

Knott, Alistair (1996). *A Data-Driven Methodology for Motivating a Set of Coherence Relations*, (Ph.D. thesis). Edinburgh: University of Edinburgh.

Lagerwerf, Luuk (1998). *Causal Connectives have Presuppositions.* The Hague, The Netherlands: Holland Academic Graphics. PhD Thesis, Catholic University of Brabant.

Lascarides, Alex & Nicholas Asher (1993). *Temporal Interpretation, Discourse Relations and Commonsense Entailment.* Linguistics and Philosophy, 16(5):437–493.

Mann, William C. & Sandra A. Thompson (1988). *Rhetorical Structure Theory. Toward a Functional Theory of Text Organization.* Text, 8(3):243–281.

Marcu, Daniel (2000). *The Rhetorical Parsing of Unrestricted Texts: A Surface-Based Approach*. Computational Linguistics, 26(3):395–448.

Marcus, Mitchell P., Beatrice Santorini & Mary Ann Marcinkiewicz (1993). *Building a Large Annotated Corpus of English: The Penn Treebank.* Computational Linguistics, 19.

Mellish, Chris, Mick O'Donnell, Jon Oberlander & Alistair Knott (1998). *An Architecture for Opportunistic Text Generation.* In Proc. of the 9th Intl. Workshop on NLG, pp. 28–37. Ontario, CA.

Polanyi, Livia (1996). *The Linguistic Structure of Discourse*. Technical Report CSLI-96-200: CSLI.

Polanyi, Livia & Martin H. van den Berg (1996). *Discourse Structure and Discourse Interpretation.* In P. Dekker & M. Stokhof (Eds.), Proc. of the 10th Amsterdam Colloquium, pp. 113–131. University of Amsterdam.

Prasad, Rashmi & Anoop Sarkar (2000). *Comparing Test-Suite Based Evaluation and Corpus-Based Evaluation of a Wide-Coverage Grammar For English.* In Using Evaluation within HLT Programs: Results and Trends LREC'2000 Satellite Workshop. Greece, pp. 7–12.

Rooth, Mats (1992). *A Theory of Focus Interpretation.* Natural Language Semantics 1, pp. 75–116.

Sarkar, Anoop (2000). *Practical Experiments in Parsing using Tree Adjoining Grammars*. In Proc. of TAG+5, France, May 25–27.

Sarkar, Anoop (2001). *Applying Cotraining Methods to Statistical Parsing*. In Proc. of the 2nd NAACL. Pittsburgh, PA.

XTAG-Group, The (2001). *A Lexicalized Tree Adjoining Grammar for English*. Technical Report IRCS 01-03: University of Pennsylvania.

Scha, Remko & Livia Polanyi (1988). *An Augmented Context Free Grammar for Discourse*. In Proc. of the COLING'88, pp. 573–577. Hungary.

Schilder, Frank (1997). *Tree Discourse Grammar, or How to Get Attached to a Discourse*. In Proc. of the Tilburg Conference on Formal Semantics, Netherlands, January 1997.

Steedman, Mark (2000a). *Information Structure and the Syntax-Phonology Interface*. Linguistic Inquiry, 34:649–689.

Steedman, Mark (2000b). *The Syntactic Process*. Cambridge MA: MIT Press.

Stone, Matthew & Christine Doran (1997). *Sentence Planning as Description using Tree Adjoining Grammar*. In Proc. of ACL, pp. 198–205.

Stone, Matthew, Christine Doran, Bonnie Webber, Tonia Bleam & Martha Palmer (2001). *Microplanning from Communicative Intentions: Sentence Planning using Descriptions (SPUD)*. Submitted to Computational Intelligence.

van den Berg, Martin H. (1996). *Discourse Grammar and Dynamic Logic*. In P. Dekker & M. Stokhof (Eds.), Proc. of the 10th Amsterdam Colloquium, pp. 93–111. University of Amsterdam.

Walker, Marilyn A. (1993). *Informational Redundancy and Resource Bounds in Dialogue*, (Ph.D. thesis). University of Pennsylvania, CIS

Webber, Bonnie (1988). *Tense as Discourse Anaphor*. Computational Linguistics, 14(2):61–73.

Webber, Bonnie & Aravind Joshi (1998). *Anchoring a Lexicalized Tree-Adjoining Grammar for Discourse*. In Manfred Stede, Leo Wanner & Eduard Hovy (Eds.), Discourse Relations and Discourse Markers: Proceedings of the Conference, pp. 86–92. Somerset, New Jersey: ACL.

Webber, Bonnie, Alistair Knott & Aravind Joshi (1999a). *Multiple Discourse Connectives in a Lexicalized Grammar for Discourse*. In 3rd IWCS, Tilburg, The Netherlands.

Webber, Bonnie, Alistair Knott, Mathew Stone & Aravind Joshi (1999c). *Discourse Relations: A Structural and Presuppositional Account using Lexicalised TAG*. In Proc. of the 36th ACL, College Park MD., pp. 41–48.

Webber, Bonnie, Alistair Knott, Mathew Stone & Aravind Joshi (1999b). *What are Little Trees made of: A Structural and Presuppositional Account using Lexicalized TAG*. In Proc. of the 36th ACL, College Park, MD, pp. 151–156.

Xia, Fei, Martha Palmer & Aravind Joshi (2000). *A Uniform Method of Grammar Extraction and its Applications.* In Proc. of the Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora (EMNLP/VLC). Hong Kong.

Information structure and pronominal reference to clausally introduced entities

JEANETTE GUNDEL
University of Minnesota
gunde003@umn.edu

MICHAEL HEGARTY
Louisiana State University
mhegar1@lsu.edu

KAJA BORTHEN
Norvegian University of Science and Technology
kaja.borthen@hf.ntnu.no

ABSTRACT. Clausally introduced entities in English are more frequently accessible to reference with a demonstrative pronoun than with the personal pronoun *it*. This fact can be explained on the assumption that such entities are typically activated, but not brought into focus, immediately subsequent to their introduction into a discourse. However, clausally introduced entities are, in fact, sometimes referenced with *it* immediately subsequent to their introduction. An examination of the discourse environments in which this is possible provides insights into the various factors, including information structure, which can boost the salience of an entity and bring it into focus.

1 Introduction

When entities are introduced into a discourse by a clause (or other non-nominal expression), they are accessible to immediate subsequent reference with demonstrative pronouns, but comparatively less accessible to reference with personal pronouns, as noted by Webber (1988) and others since[1]. Thus, Webber (1991) found that of 96 pronominal references referring to the interpretation of one or more previous clauses in written English texts, only 15 used the personal pronoun *it*, while the rest were demonstrative *this* or *that*. Personal pronouns tend to favor reference

[1]Our examples here will be from English, although similar restrictions on pronominal reference to clausally introduced entities can be found in other languages.

Paper presented at the ESSLLI 2001 Workshop on
Information Structure, Discourse Structure and Discourse Semantics
Edited by Ivana Kruijff-Korbayová and Mark Steedman.

to nominally introduced entities, and reference to a clausally introduced entity with *it* is often impossible, or at least highly infelicitous, as illustrated in the following examples.

(1) a. There was a snake on my desk. **That** scared me.

b. There was a snake on my desk. **It** scared me. (Borthen, Thorstein, and Gundel 1997)

(2) a. Max destroyed his leaf collection last night. **That** was dumb.

b. Max destroyed his leaf collection last night. **It** was dumb

(3) a. Simplified English disallows the use of passive, progressive, and perfective auxiliary verbs, among other things. **This** requires engineers ...

b. Simplified English disallows the use of passive, progressive, and perfective auxiliary verbs, among other things. **It** requires engineers ... (Gundel, Hedberg, and Zacharski 1993)

(4) a. "We believe her, the court does not, and **that** resolves the matter," Mr. Montanarelli said today of Ms. Lewinsky's testimony that she had an independent recollection of the date.

b. "We believe her, the court does not, and **it** resolves the matter," Mr. Montanarelli said today of Ms. Lewinsky's testimony that she had an independent recollection of the date. (*New York Times*, May 24, 2000)

(5) a. Cloned humans might show higher rates of cancer or other diseases, but we'd only find out by cloning them and waiting to see if disaster strikes. None of **this** means, however, that ... even that human cloning isn't going on right now. (Talbot, Margaret. February 4, 2001. *New York Times Magazine*, Section 6, p.45.)

b. .... # None of **it** means, however, ...

(6) A: I read somewhere that the poodle is one of the most intelligent dogs around.

B: well uhm..I definitely wouldn't dispute **that**. (Switchboard Corpus, Dialog 2019)

B': ?? well uhm..I definitely wouldn't dispute **it**.

(7) A1: So you fired her?

B: We're going to do a lot more than just fire her.

A2: What does **that** mean? (from the soap opera "The Bold and the Beautiful")

A2': # What does **it** mean?[2]

[2]Note that *it* in this example, as well as the preceding one, would be infelicitous even if stress falls on the verb.

The demonstrative pronoun in (1)-(7) refers to some entity (a situation, fact, act, etc.) introduced by a previous clause. By contrast, the pronoun *it* is either most naturally interpreted as referring to an entity introduced by a nominal, as in (1)-(4), or it renders the sentence infelicitous in the context when there is no appropriate nominal antecedent, as in (5)-(7). In (2), for example, the pronoun *it* is most easily interpreted as referring to Max's leaf collection, not his act of destroying it. And in (7) , A2' is infelicitous because *it* , unlike *that*, cannot refer to B's statement "we are going to do a lot more than just fire her."

In this paper, we argue that facts regarding the distribution and interpretation of *this/that* vs. *it* referring to clausally introduced entities can be explained within the theory of reference and cognitive status proposed by Gundel, Hedberg, and Zacharski (1993 and earlier work). Approached in the context of this framework, these facts also provide insights into the more general question of how various linguistic factors, including information structure, promote the salience of discourse entities and bring them into the addressee's focus of attention.

2 The Givenness Hierarchy

Gundel *et al* (1988, 1993) propose that determiners and pronouns constrain possible interpretations of nominal forms by conventionally signaling the memory or attention status that the intended referent is assumed to have in the mind of the addressee. Gundel et al identify six cognitive statuses. The array of statuses, called the Givenness Hierarchy, is presented in Figure 1.

in focus	⊃	activated	⊃	familiar	⊃	uniquely identifiable	⊃	referential	⊃	type identifiable
(*it*)		{ *that* / *this* / *this* N }		(*that* N)		(*the* N)		(indefinite *this* N)		(*a* N)

Figure 1. The Givenness Hierarchy (GH) and associated forms in English.

Statuses on the hierarchy correspond to memory and attention states, ranging from most restrictive, "in focus", to least restrictive, "type identifiable." An empirical claim of the theory is that all languages have ways of coding cognitive statuses with individual determiners and pronouns, and that such forms will be used appropriately only if the status they conventionally code within the language is satisfied in the given context of use. The forms thus serve as processing signals which assist the addressee in restricting possible interpretations.

The statuses are in a unidirectional entailment relation. If something is in focus (center of attention), it is necessarily activated (in working memory); if it is activated, it is necessarily familiar (in memory); if it is familiar, then the addressee can associate a unique representation; if the addressee can associate a unique representation, then it is referential; and if it is referential, it must be type identifiable.

The theory thus correctly predicts that a given cognitive status can be appropriately coded by a form which explicitly signals that status, but also by forms whose meanings are entailed by that status. In the latter case (e.g. use of a definite article for a referent that is in focus) the form is simply underspecified for cognitive status of the intended referent.

The use of underspecified forms has limits, however, because of interaction of the Givenness Hierarchy with general pragmatic principles involved in language production and understanding (see Grice, 1975, Sperber and Wilson, 1986/95). The implicational nature of the GH gives rise to "scalar implicatures", in the sense of Horn (1972), which further restrict the distribution and interpretation of referring forms (see Gundel *et al.*, 1993, Gundel and Mulkern, 1998). For example, in English, the indefinite article is rarely used if the status is higher than referential, resulting in association of the indefinite article with unfamiliarity. Use of the indefinite article typically implicates by the first part of the Quantity Maxim (make your contribution as informative as possible) that conditions for using a more restrictive form are not met since the addressee is not able to uniquely identify an intended referent. Another result of interaction of the Givenness Hierarchy with the Quantity Maxim is that most in-focus referents are not coded with demonstratives, even though they could be; and demonstratives often implicate a focus shift.

As seen in Figure 1, Gundel et al propose that demonstrative pronouns in English code the status "activated", whereas the pronoun *it* codes the more restrictive status "in focus". This permits an explanation of facts like those in (1)-(7), if the clausally introduced entiities in these examples have been activated but not brought into focus. In the following section, we examine factors that contribute to bringing an entity into focus, including the role that information structure plays in determining the cognitive statuses of referents introduced by clauses, and thus the nominal forms which can be used to refer to these entities.

3 What brings an entity into focus of attention?

3.1 Syntactic structure

The framework outlined above makes predictions about the appropriateness of different pronominal forms depending on whether or not the intended referent can be assumed to be in focus for the addressee. Although the theory itself does not predict what brings an entity into focus, Gundel et al. (1993) suggest that "the entities in focus at a given point in the discourse will be that partially-ordered subset of activated entities which are likely to be continued as topics of subsequent utterances." Membership in this set is partly, though not wholly, determined by syntactic structure. For example, subjects and direct objects of matrix sentences are more likely to bring an entity into focus than elements in subordinate clauses and prepositional phrases (cf. the Centering Algorithms of Grosz, Joshi, and Weinstein, 1995a,b). These assumptions make it possible to explain facts about the distribution and interpretation of demonstratives and unstressed personal pronouns

(including *it*) such as those illustrated in (8) and (9).

(8) a. My neighbor's Bull Mastiff bit a girl on a bike.
    b. **It's/That's** the same dog that bit Mary Ben last summer.

(9) a. Sears delivered new siding to my neighbors with the Bull Mastiff.
    b. **#It's/That's** the same dog that bit Mary Ben last summer.

Since the Bull Mastiff is introduced in matrix subject position (and is most likely also the topic) in (8a), it is brought into focus, and can therefore be appropriately referred to with either *that* or *it* in (8b). The pronoun *it* is possible in (8b) because the intended referent is in focus. The pronoun *that* is possible because anything in focus is also activated, i.e. in working memory. But in (9), where the Bull Mastiff has been introduced in a more peripheral position, it is activated but not brought into focus. Therefore, only reference with *that* is possible.

This account can be naturally extended to facts like those in (1)-(7) if we make the relatively uncontroversial assumption that entities (indirectly) introduced by a whole clause, or sequence of clauses, will be activated, but are much less likely to be brought into focus than entities introduced by major thematic arguments of the verb. For example, in (2), at the conclusion of A's utterance, the act of destroying the leaf collection can be assumed to be activated, since it was just introduced in the preceding sentence, but not in-focus; the focus of attention after the utterance is processed is on the referents of the major arguments in (2A), specifically, Max and the leaf collection. Similarly, in (5), the complex situation consisting of potential drawbacks to human cloning is rendered activated by the first paragraph, but we can assume that it is not rendered in focus given the higher salience conferred by this passage on cloned humans, rates of cancer, and other referents of main clause arguments.

A fact or proposition introduced by an NP within a clause is also more likely to be brought into focus than one which is introduced by the whole clause. Compare (10) with the examples (1)-(7) above, for example.

(10) a. Then, Maria brought up another fact. **It** sent shivers down my spine.
     b. Max then introduced a new proposition. But **it** was rejected.

3.2 Semantic and pragmatic factors

Conditions which appear to boost the salience of entities also include less overt factors such as covert arguments, presuppositions and prior beliefs, and even inquisitive looks, all of which can cause an entity to be "reprocessed", and thus brought into focus, even when it is overtly mentioned only once (see Borthen, 1997, Gundel, Borthen, and Fretheim, 1999).

In (11), a baseline case for comparison, the speaker, upon clausally introducing the fact that linguists earn less than computer scientists, can assume that this fact is rendered activated, but not in-focus, for the hearer, leading to a preference for *that* over *it* in the follow-up reference to this fact.

(11) a. I hear linguists earn less than computer scientists, and **that**'s terrible.

b. ?? I hear linguists earn less than computer scientists, and **it**'s terrible.

In (12), in contrast, the follow-up reference is made by another speaker, which results in somewhat more complicated inferences regarding the cognitive status of the fact at issue.

(12) A: I just read that linguists earn less than computer scientists.

B: (i) **That**'s terrible! (ii) **It**'s terrible!

At the completion of A's utterance, B can assume that the fact that linguists earn less than computer scientists is at least activated for A. In response B(i), B's use of *that* signals the assumption that this fact has been activated, but possibly not brought into focus, by A's utterance, thereby inviting A to infer that the fact is news to B. In response B(ii), B signals the assumption that the fact is in focus for A, or ought to be, consistent with it being accepted background information for discourse in the relevant social circle; this invites A to infer that B already knew the fact.

In (13) below, the proposition that B has a dental appointment is clausally introduced by A's utterance. If the mere utterance of a sentence does not bring the expressed proposition into focus, this would explain why (13)B' sounds unnatural, given that *it* requires the referent to be in focus, whereas *that* merely requires activation.

(13) A: You have a dental appointment at noon.

B: That's true.

B': ??It's true.

B": It's true, then.

But (13) B" is noticeably more acceptable than (13)B'. Following Gundel et al. (1999), we suggest an explanation of this fact, drawing on a relevance-theoretic approach to the pragmatics of language understanding (Sperber and Wilson, 1986/95). The word *then* in B" functions as an interpretive particle which conveys the meaning that the content of the sentence it is appended to follows by way of inference from something the addressee just said. The response by B in (13)B" means essentially, "Given your assertion that I have a dental appointment at noon, then I can take it as confirmed that I have a dental appointment at noon." The only way the utterance in B" can yield contextual effects for A is if A's utterance confirmed the truth of a proposition that B had been questioning, and B knows that A is aware of this. Thus, the fact that B had a dental appointment at noon was not activated for the first time by A; rather, A's utterance brought into focus a fact that was already mutually manifest to both A and B beforehand, thereby licensing the use of *it* in B".

Salience can also be boosted non-linguistically. For example, the exchange in (14) below is fully natural if A gives B a skeptical look during the indicated pause.

(14) A: Why didn't you come to the rehearsal yesterday?

B: I thought I told you. I had to help Peter move. (Pause) It's true!

The skeptical look communicates A's skepticism about the truth of the proposition just expressed by B, thus causing the proposition that B has to help Peter move to be reprocessed (by both A and B) and assuring that it is mutually in focus, making it accessible to reference with *it*.

Salience of an entity in the environment also suffices for pronominal reference with *it*. If A and B are in a room together with a baby who suddenly begins to walk, A can produce the utterance in (15), or, if A sees B watching the baby walk, the utterance in (16).

(15) Will you look at **that**! The baby's walking. (Jackendoff 2001)

(16) Isn't **it** great? [it = the fact that the baby is walking]

4 The role of information structure

The cognitive status, and therefore the accessibility to pronominal reference, of a clausally introduced entity is partly constrained by the information structure of the utterance in which it is introduced into a discourse[3]

In particular, information structure yields some striking effects, but also a surprising asymmetry, when higher order entities are introduced by (or within) clausal complements.

Entities introduced by clausal complements to bridge verbs, such as *think*, *believe*, and *say*, exhibit the familiar pattern of being rendered activated, but not in-focus, through mention by a clause. This is shown by the naturally occurring example in (17) below, as well as by the constructed data in (18), tested on a small survey of English speakers[4]

(17) Ising reportedly believed that his negative results would hold in higher dimensions as well.
In bf this conjecture he was wrong. (*American Scientist* 88:385)
In **this/ #it**, he was wrong.

(18) What does Alex think?

[3]By information structure, we mean a bifurcation of material in an utterance into what has been called focus versus ground, comment versus topic, or rheme versus theme. This notion is not to be identified with contrastive focus or with the more general distinction between new versus old information. Information structural focus is also distinct from the cognitive status "in focus". See Vallduví (1990) and Gundel (1999a) for more detailed discussion of related terminological and conceptual issues. We will indicate information structural focus by the subscript F.

[4]The use of *it* in (17) would be just as infelicitous if the PP were not preposed. Thus, the infelicity of *it* in (17) cannot be attributed to its incompatibility with the secondary focal stress it bears in this position.

A: Alex believes [$_F$ that the company destroyed the FILE].

B: **That**'s false; the file has been submitted to the district judge.

B': # **It**'s false; the file has been submitted to the district judge.

When (18A) is used with the focus-structure shown, to introduce the proposition that the company destroyed the file, the response by B using *that* is much more felicitous than the response with *it*. However, *it* and *that* are equally good when the complement clause is in the ground (theme/topic) of A's utterance, as in (19).

(19) A: Alex [$_F$ INSISTS/BELIEVES] that the company destroyed the file.

B: But **that**'s/**it**'s false; the file has been submitted to the district judge.

Since an entity associated with the ground (theme;topic) is already at least familiar to the addressee prior to the utterance (see Gundel 1988) inter alia), its mention within the utterance suffices to bring it into the focus of attention, if it does not already have that status.

In (17)-(19), relational givenness/newness and referential givenness/newness (in the sense of Gundel, 1988, Gundel 1999a,b) are coextensive. For example, the information structural focus in (18) represents a proposition that is not only new in relation to the topic (what Alex believes), but also referentially new to the hearer; and the clausal complement in (19A) (that the company destroyed the file) represents a proposition which is not only given in relation to the informational structural focus; it is also referentially given in the sense of being at least familiar, and probably also activated. But material in the informational focus doesn't have to be referentially new (see Gundel 1980, Gundel 1999a, Gundel 1999b, Vallduví 1990, Lambrecht 1994). So when we have a bridge verb complement which is an information structural focus, but is already activated in the discourse, which factor wins out? Is an entity expressed by such a complement rendered in-focus or does it remain merely activated? Is it accessible to reference with *it*, or only with *that*? Consider (20).

(20) A1: I believe that the company destroyed the file, but not everybody does.

B1: What does Alex believe?

A2: Alex believes [$_F$ that the company destroyed the file].

B2: But **it**'s/**that**'s false; the file has been submitted to the district judge.

(20B2) suggests that it is referential givenness (cognitive status of a discourse entity) , and not relational givenness (topic-focus structure) that determines whether the complement of a bridge verb will be brought into focus.

But now flip the problem around. Content in the topic/ground of an utterance does not always have a high degree of referential givenness. Its cognitive status may be merely familiar, but not necessarily activated. So when we have a bridge verb complement which is ground material, but new to the discourse, which factor wins out? Is an entity introduced by such a complement rendered in-focus, because

it is in the ground, or merely activated, because it is new to the discourse? Is it accessible to reference with *it*, or only with *that*? Consider (21) [secondary stress on *murdered*]:

(21) a. Alex is hopeless.
b. He [$_F$ INSISTS] that Tom was murdered, for example,
c. –even though there's not a shred of evidence for *that*.
–even though there's not a shred of evidence for ıit.

Use of *it* is as felicitous as *that* in (21c). The information structure of (21) forces an interpretation where the content of the complement clause is already familiar, so that (21b) renders it in-focus, making it available to reference using *it*. Thus, presentation of a clausally introduced entity in the ground of an utterance is another way to promote salience, and bring the entity into focus, even if it is, in fact, new to the discourse.

With bridge verb complements, we thus appear to have an asymmetric situation: bifurcation into focus/ground has no effect on the cognitive status of an entity introduced within the information structural focus[5] But it **can** have an effect when an entity is mentioned (even introduced) within ground material, because mention within the ground necessarily signals a higher cognitive status for the entity. This conclusion is preliminary, however, in that the judgments are subtle, and naturally occurring data that would bear directly on the issue is sparse.

5 Lexical structure versus information structure

When the bridge verb in an example like (18) is replaced with a factive verb, demonstrative and personal pronouns can both be used to immediately refer to the proposition expressed by the complement clause, regardless of the information structure of A's utterance, as seen in (22).

(22) A: Alex verified that the company destroyed the file.
B: That's false; the file has been submitted to the district judge.
B': It's false; the file has been submitted to the district judge.

Thus, the contrast in (18) between subsequent reference with *it* versus *that* is not exhibited in (22). The lexical semantics of the factive verb enforces the condition that the entity expressed by the complement clause be already familiar (or at least capable of being accommodated as familiar) to the addressee, so that its further mention in A's utterance renders this entity in-focus.

In order to understand this fully, it is useful to note that this pattern is not confined to complements of factive verbs. It is also obtained in complements to

[5]Gundel (1999a) makes a similar observation, concluding that mention within the information structural focus (her "semantic focus") doesn't necessarily bring an entity into focus of attention.

certain non-factive (and non-bridge) verbs, including *agree*, *emphasize*, *deny*, and *doubt*, and in complements to the non-factive adjectival predicate *be certain*.[6]

(23) a. Alex and Susan agree that the company destroyed the file. I'm surprised that they believe it.

b. Alex and Susan agree that the company destroyed the file. I'm surprised that they believe that.

(24) A: Alex is certain that the company destroyed the file.

B: That's false: the file has been submitted to the district judge.

B': It's false: the file has been submitted to the district judge.

As with factive predicates, the pattern in (23)-(24) is one in which **it** is at least as felicitous as **that** in referring to the content of the complement clause, and, in some cases, more so.

The predicates in (23)-(24) are not factive (in the sense made clear by Kiparsky and Kiparsky (1971)) since they don't commit the speaker of the ascription in which they occur to the truth of their complement clauses. However, they share with factives a slightly more subtle semantic property: they are felicitous when the proposition, fact, or situation expressed by the complement clause is already accepted as given or familiar in the discourse (see Hegarty, 2001). Using a situation variable in the semantics, in the context of Discourse Representation Theory (Kamp and Reyle, 1993), the interpretation of the factive ascription in (22) can be expressed by the Discourse Representation Structure (DRS) shown in (25) below[7]

The ascriptions with ıagree and *certain* in (23)-(24), though non-factive, would have identical DRS's, with trivial substitution of the verb denotations.

(25)

u, v, z, s
Alex(u)
Company (v)
File (z)
destroy (v, z, s) (w_o)
verify ($u, \lambda w$ [destroy (v, z, s)(w)])

In contrast, a belief ascription such as that in (18A), using a bridge verb, is interpreted semantically as just a relation between Alex and the proposition expressed by the complement clause. A DRS for (16A) is presented in (26).

[6]Cattell (1978) noticed that these non-factives pattern with factives in wh-extraction from their complements. See also Melvold (1991), Hegarty (1992), and Schulz (1999) for discussion of this class of predicates.

[7]Subordinate DRSs are abbreviated as formulas here to save space. For semantic representations using a situation variable, see Ginzburg (1995a,b), and, for similar structures with an event variable, (Higginbotham 1985, Higginbotham 1989).

(26)

u, v, z
Alex (u)
Company (v)
File (z)
believe ($u, \lambda w$ [$\exists s$ [destroy (v, z, s)(w)]])

Of course, the ascription made by A in (18) could express a proposition which is already familiar to the hearer. The property distinguishing bridge verbs from the factive and other predicates discussed here is not that the content of the bridge verb complement **must** be unfamiliar, but only that it **can** be. Bridge verbs, unlike other predicates discussed here, do not assume the familiarity of the content of the complement.

Interrogatives pattern with factive complements with regard to the status of abstract entities mentioned by or within them. A naturally occurring example is shown in (27).

(27) One common attribute of a scientist is an unusually acute sense of numbers and their implications.
A sense of numbers - why do I dwell on this observation? Perhaps **it's** because we who come from a background of engineering...(*American Scientist* 88:378)

(28) A: Alex wonders whether the company destroyed the file.

B: **It**'s not likely. The file contained no incriminating information.

B: **That**'s not likely. The file contained no incriminating information. [it/that = that the company destroyed the file]

(29) a. Max wonders who destroyed the file; **it** has impeded the investigation.

b. Max wonders who destroyed the file; **that** has impeded the investigation. [it/that = that someone destroyed the file]

The possibility of immediate subsequent reference with a personal pronoun in (28)-(29) follows from the presuppositional nature of questions. To simplify, within DRT, the *wonder*-ascription in (28A) should be represented with a DRS of the form shown in (30), where ϕ is an appropriate relation between Alex and the proposition p specified on the penultimate line of the DRS[8]

[8]To unsimplify, questions are, in fact, constrained not only by the formal semantic condition captured here, but by rich contextual conditions on what would count as a suitable answer to a question in a given context. See Ginzburg (1995a), Ginzburg (1995b), and Asher and Lascarides (1998). The important point, for present purposes, is that these accounts would incorporate, and add to, the presuppositional condition given here. The proposals sketched here would therefore be a part of an account given according to these richer theories of the interpretation of questions.

(30)

u, v, z, p
Alex (u)
Company (v)
File (z)
$\mathrm{p}(w) = \lambda w \exists s$ [destroy $(v, z, s)(w)$]
ϕ

Interpreted as in (30), the *wonder*-ascription in (28A) is a question about the proposition that the company destroyed the file. This should be the form of any semantic account of the *wonder*-ascription which captures the presuppositionality of the embedded question: the proposition that the company destroyed the file must be an established discourse entity prior to the utterance of (28A), or it must be accommodated in the sense of Heim (1982). The assertive content of (28A) should be captured in the last line of the DRS, ϕ. On one realization of ϕ, given in Hegarty (2001), (28A) asserts that Alex is in the state of wonder with respect to the proposition that p holds of the actual world, w_o.

Thus, the presuppositionality involved in the lexical structure of a factive (or related) predicate, and the semantic presuppositionality of embedded questions, are additional factors which can bring an entity into focus. In these cases, information structure has no bearing on the cognitive status of the clausally introduced entity.

6 Conclusion

In this paper, we addressed the fact that clausally introduced entities, immediately subsequent to their introduction into a discourse, are typically accessible to reference with a demonstrative pronoun, but not with the personal pronoun *it*. We found that this fact can be explained on the basis of the observation that such entities are typically activated, but not brought into focus, upon their introduction to a discourse. However, clausally introduced entities are, in fact, sometimes referenced with *it* immediately subsequent to their introduction. An examination of the discourse environments in which this is possible provides important insights into the various syntactic, semantic, and pragmatic factors that can boost the salience of an entity and bring it into focus.

We've shown that information structure, in the sense of a focus-ground bifurcation, is one such factor when an entity is mentioned with a bridge verb complement, but only in a way which is asymmetric, depending on whether the entity is mentioned within focal or non-focal material. When the complement is focal, there is no effect: the cognitive status of an entity expressed by a focal complement depends entirely on the referential givenness/newness (i.e. the cognitive status) of the entity. But when the complement is part of the ground (topic/theme), the entity is brought into focus.

In factive complements and embedded questions, the lexical nature of the embedding predicate and the semantic nature of the construction require an entity

mentioned with the subordinate clause to be treated as referentially given independently of the information-structure of the utterance. This suggests that the salience promoting effect of information structure is indirect. It is the referential givenness of the ground, i.e. the fact that topics are at least familiar, and not information structure per se, which contributes to bringing an entity into focus.

Bibliography

Asher, N. and A. Lascarides (1998). Questions in dialogue. *Linguistics and Philosophy 21*, 237–309.

Borthen, K., F. Thorstein, and J. K. Gundel (1997). What brings a higher-order entity into focus of attention? In R. Mitkov and B. Boguraev (Eds.), *Operational factors in practical, robust anaphora resolution*, pp. 88–93.

Cattell, R. (1978). On the source of interrogative adverbs. *Language 54*, 61–77.

Ginzburg, J. (1995a). Interrogatives I. *Linguistics and Philosophy 18*, 459–527.

Ginzburg, J. (1995b). Interrogatives II. *Linguistics and Philosophy 18*, 567–609.

Grice, H. P. (1975). Logic and conversation. In P. Cole and J. Morgan (Eds.), *Syntax and Semantics 3, Speech Acts*, pp. 41–58. Academic Press.

Grosz, B. J., A. K. Joshi, and S. Weinstein (1995a). Providing a unified account of definite noun phrases in discourse. In *Proceedings for the 21st Annual Meeting of the Association for Computational Linguistics*, pp. 44–50.

Grosz, B. J., A. K. Joshi, and S. Weinstein (1995b). Towards a computational theory of discourse interpretation. *Computational Linguistics 21*, 203–25.

Gundel, J. K. (1980). Zero NP-anaphora in Russian: a case of topic-prominence. In *Proceedings from the Parasession on Anaphora*, pp. 139–146. Chicago Linguistic Society.

Gundel, J. K. (1988). Universals of topic-comment structure. In M. Hammond et al. (Eds.), *Studies in syntactic typology*, pp. 209–239. John Benjamins.

Gundel, J. K. (1999a). On three kinds of focus. In P. Bosch and R. van der Sandt (Eds.), *Focus. Linguistic, Cognitive and Computational Perspectives*, pp. 293–305. Cambridge University Press.

Gundel, J. K. (1999b). Topic, focus and the grammar pragmatics interface. In N. J. Alexander and M. Minnick (Eds.), *Proceedings of the 23rd Annual Penn Linguistics Colloquium*, Volume 6.1, pp. 185–200. Penn Working Papers in Linguistics.

Gundel, J. K., K. Borthen, and T. Fretheim (1999). The role of context in pronominal reference to higher order entitites in English and Norwegian. In P. Bouquet et al. (Eds.), *Modeling and Using Context. Proceedings from the Second International and Interdisciplinary Conference, CONTEXT '99*, Lecture Notes in Artificial Intelligence 1688, pp. 475–478. Springer.

Gundel, J. K., N. Hedberg, and R. Zacharski (1988). The generation and interpretation of demonstrative expressions. In *Proceedings of the XIIth International Conference on Computational Linguistics*, pp. 216–221. John Von Neumann Society for the Computing Sciences.

Gundel, J. K., N. Hedberg, and R. Zacharski (1993). Cognitive status and the form of referring expressions in discourse. *Language 69*, 274–307.

Gundel, J. K. and A. Mulkern (1998). Quantity implicatures in reference understanding. *Pragmatics and cognition 6: Special issue on the concept of reference in the cognitive sciences*, 21–45.

Hegarty, M. (1992). Adjunct extraction without traces. In D. Bates (Ed.), *Proceedings of the Tenth West Coast Conference on Formal Linguistics*, pp. 209–223. Stanford: CSLI.

Hegarty, M. (2001). Reference to abstract entities within clausal complements. In *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*.

Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. Ph. D. thesis, University of Massachusetts, Amherst.

Higginbotham, J. (1985). On semantics. *Linguistic Inquiry 16*, 547–593.

Higginbotham, J. (1989). Elucidations of meaning. *Linguistics and Philosophy 12*, 465–517.

Horn, L. R. (1972). *On the semantic properties of logical operators in English*. Ph. D. thesis, UCLA.

Jackendoff, R. (2001). Reference and truth. Plenary Talk, Annual Meeting of the Linguistic Society of America, Washington, D.C.

Kamp, H. and U. Reyle (1993). *From Discourse to Logic*. Kluwer Academic Publishers.

Kiparsky, P. and C. Kiparsky (1971). Fact. In D. Steinberg and L. Jacobovits (Eds.), *Semantics*, pp. 345–369. Cambridge University Press.

Lambrecht, K. (1994). *Information structure and sentence form: topic, focus, and the mental representations of discourse referents*. Cambridge University Press.

Melvold, J. (1991). Factivity and definiteness. In L. Cheng and H. Demirdash (Eds.), *More Papers on WH-Movement*, Volume 15, pp. 97–117. MIT Working Papers in Linguistics.

Schulz, P. (1999). *Getting the facts, finite complements, factive verbs and their acquisition*. Ph. D. thesis, Tübingen.

Sperber, D. and D. Wilson (1986/95). *Relevance: Communication and cognition*. Blackwell. 2nd Edition.

Vallduví, E. (1990). *The Informational Component*. Ph. D. thesis, University of Pennsylvania.

Webber, B. L. (1988). Discourse deixis and discourse processing. Technical report, University of Pennsylvania.

Webber, B. L. (1991). Structure and ostension in the interpretation of discourse deixis. *Language and Cognitive Processes* 6(2), 107–135.

Entangled Information Structure: Analysis of Complex Sentence Structures

NOBO KOMAGATA

Department of Computer and Information Science, University of Pennsylvania

komagata@linc.cis.upenn.edu

ABSTRACT. While information structure has traditionally been viewed as a single partition of information within an utterance, there are opposing views that identify multiple such partitions in an utterance. The existence of alternative proposals raises questions about the notion of information structure itself and also its relation to discourse structure. This paper supports the traditional view by observing contextual requirements and linguistic phenomena associated with information structure for each alternative.

1 Introduction

Traditionally, information structure (IS) has been viewed as non-recursive, matrix-level organization of information within an utterance (e.g., Mathesius 1975). According to this approach, even a complex utterance has only one IS partition as can be seen below (adapted from Lambrecht 1994).

(1) *Q*: Why did you hit him?

A: [I hit him]$_{Theme}$ [because he insulted me]$_{Rheme}$.

Here and throughout this paper, the IS labels 'theme' and 'rheme' (T and R, respectively, in later examples) are used instead of more overloaded terms, e.g., 'topic' and 'focus', most closely following Steedman (2000). Although the notions associated with various terms may differ in many respects, we try to limit our discussion to the essential properties of theme and rheme involving binary informational contrast between them (cf. Communicative Dynamism of Firbas 1964).

In contrast to the traditional view, some researchers observe multiple IS partitions within an utterance. For example, Kruijff-Korbayová and Webber (2001) propose the following analysis.

(2) Although [Clyde married]$_T$[BERTHA]$_R$, [he]$_T$[did not inherit a PENNY]$_R$.

SMALL CAPITALS are used to indicate pitch accents. In addition, Partee (1996) considers even a recursive IS such as the following.

Paper presented at the ESSLLI 2001 Workshop on
Information Structure, Discourse Structure and Discourse Semantics
Edited by Ivana Kruijff-Korbayová and Mark Steedman.

(3) [What convinced Susan that [our arrest]$_T$[was caused by HARRY]$_R$]$_T$[was a rumor that [someone]$_T$[had witnessed Harry's confession]$_R$]$_R$.

Naturally, the existence of three competing views poses a challenge to the definition of IS.[1] In addition, this issue is also relevant to the analysis of the relation between IS and discourse structure (DS). To see this point, let us first classify the above-mentioned three approaches by referring to the span of a theme-rheme pair as 'domain of IS'.

(4) *a*. Traditional view: Domain of IS = utterance

b. Kruijff-Korbayová and Webber: Domain of IS = clause

c. Partee: Domain of IS = utterance *and* clause (recursive)

We then observe the idea of 'thematic progression' studied by Daneš (1974), which is schematically shown below. T_i and R_i refer to the theme and rheme, respectively, of the *i*th utterance in a discourse.

(5)
$$\begin{array}{lllllll} T_1 & \longrightarrow & R_1 & & & & \\ & & \downarrow & & & & \\ & & T_2\ (=R_1) & \longrightarrow & R_2 & & \\ & & & & \downarrow & & \\ & & & & T_3\ (=R_2) & \longrightarrow & R_3 \end{array}$$

If we assume that DS is the organization of discourse units *corresponding to clauses* (e.g., Grosz and Sidner 1986), thematic progression seems to be able to characterize the IS-DS relation quite well, especially if all sentences are simple. Daneš's idea can also be interpreted in the following way:

(6) The DS of a discourse can be determined by the DS prior to the current (last) utterance and the IS of the current (last) utterance.

I would like to see this as the primary contact point between the notions of IS and DS. However, once complex sentences are involved, the situation appears more complicated. For example, the analysis of Kruijff-Korbayová and Webber (2001) involving two IS's in a single utterance might be represented as follows, where the example (2) would correspond to the lower level:

(7)
$$\begin{array}{llllllllll} \mathrm{T}_1 & \longrightarrow & \mathrm{R}_1 & & & & & & & \\ & & \downarrow ? & & & & ? \curvearrowleft & & & \\ & [& \mathrm{T}_{2a} & \longrightarrow & \mathrm{R}_{2a} & & \mathrm{T}_{2b} & \longrightarrow & \mathrm{R}_{2b} &] \\ & & & \mathrm{IS}_1 & & & & \mathrm{IS}_2 & & \end{array}$$

[1] As pointed out by one of the reviewers, these different views may reflect the black-and-white situation involving IS. But pursuing one position against others without accepting the mixed view seems essential for a deeper understanding of the subject. Although one of the reviewer states that it is "quite plausible that clauses have IS", I do not think such 'plausibility' has ever been demonstrated.

With multiple IS's in a single utterance, the description (6) would no longer hold in its given form. We may need to ask questions such as the following. Would the DS analysis proceed in two steps, i.e., IS_1 is used to form the DS up to that point *not* including IS_2? Or, would some form of complex IS (IS_1 and IS_2) be used to form the DS up to and including IS_2 all at once? The former analysis would raise a question about the role of the subordinator (at the beginning of the first clause), which is supposed to connect the two clauses. The latter analysis would raise a question about the relation between the two IS's in connection to DS. The situation would be even more complicated with the analysis of Partee.

Between the three alternatives, this paper supports the traditional view of IS through an equivalent proposition: there is exactly one IS partition even for a complex utterance. The present position is also related to the idea: linguistic marking of information structure is a matrix-level phenomenon (Komagata 1999, p. 37). One of the consequences of this position is that the domain of IS is not fixed to the unit of DS (i.e., clause). In other words, the IS domain would appear *entangled* around clause boundaries. However, this potential complication seems to be inevitable for developing a DS out of both simple and complex utterances in a way consistent with the view (6).

The constructions we focus in this paper are complex structures involving subordinators such as *although* and *because.* The (sentence) coordinate structure is not discussed in this paper as it is fundamentally different from the subordinate structure (e.g., Quirk et al. 1985, pp. 920). It can be considered as a sequence of utterances, each of which may contain its own IS. Furthermore, if a multiple-clause structure is considered as coordinate structure, e.g., nonrestrictive sentential relative clause, there can be a separate IS for each utterance.

The second qualification is that we do not discuss a special case of IS partition within an embedded clause such as the following.

(8) *Q*: What did you think Marcel proved?
A: [I thought Marcel proved]$_T$[completeness]$_R$.

While this type of 'non-traditional' constituents are fairly common (Steedman 2000), they do not appear in the type of complex structures discussed in this paper.

The organization of this paper is as follows. Section 2 discusses problems with (2). Section 3 points out problems with (3). In Section 4, we support the traditional view of IS, mainly by discussing potential counterexamples.

2 IS Partition within the Subordinate Clause

2.1 Semantic Motivation

The main point of Kruijff-Korbayová and Webber (2001) is that we can explain the semantics of *although* if we consider an IS partition for each clause as in (2).

Roughly, their idea is that the conventional implicatures for *although* can be specified in terms of the alternative sets associated with the themes and the rhemes, for the two readings of *although*, i.e., denial of expectation and concessive opposition.

Their approach adopts the framework of Steedman (2000), which is based on alternative semantics (Rooth 1985), and does clarify the interpretation of *although.* But the question here is whether the effect is due to IS. To see this point, let us now examine the following example.

(9) The marriage of Clyde to BERTHA did not let him inherit a PENNY.

This sentence seems to be substitutable for (2) in virtually any context. In particular, both sentences contain the same contrastive situation involving Bertha in relation to the proposition that Clyde is married to Bertha. In (9), the subject-predicate relation is no longer based on *although*, but depends on the contrastive semantics associated with *Bertha* as in (2), which occurs within a noun phrase (with no embedded clauses). I do not know of any proposal for an IS partition within this type of noun phrases. Although one may contest this assumption, it seems that such a move would lead to a major revision of the standard view of IS. Thus, the contrast observed in (9) and (2) must be represented semantically regardless of the presence of an IS.

Such a semantic effect can actually be accounted for by the analysis of Steedman (2000), which distinguishes two levels between IS and focus-background. Note that the notion of 'focus' here is as in (Rooth 1985) and is *not* the same as 'rheme'. While 'rheme' is a component of IS, 'focus' is a notion tightly connected with a phonological prominence. Further, a focus can appear in a theme or a rheme as can be seen below.

(10) *Q*: I know that Marcel likes the man who wrote the musical.
But who does he ADMIRE?

A: [Marcel $\underset{\textit{theme-focus}}{\text{ADMIRES}}]_T$[the woman who $\underset{\textit{rheme-focus}}{\text{DIRECTED}}$ the musical]$_R$.

With appropriate semantic operations, both (2) and (9) could be analyzed in a similar way in terms of the contrastive situation. This paper does not show how this can be done as it is not our point. But it must be similar to the approach of Kruijff-Korbayová and Webber (2001), replacing the theme-rheme distinction with the focus-background distinction.

2.2 Availability of Contexts

In their paper, Kruijff-Korbayová and Webber (2001) consider a context for (2) such as a question "Is Clyde HAPPY?" (for the concessive-opposition interpretation). But it is not clear whether this or other questions can actually provide the right context for the proposed IS, which is assumed for both the denial-of-expectation and concessive-opposition interpretations. Before proceeding, we need

a few notes. The question test is useful in many cases. But it is always possible to respond to a question indirectly. In such a response, the IS cannot necessarily be identified based on the question. For a more precise discussion of IS, we may need to consider a more formal approach such as the one developed by Steedman (2000). But this paper remains open in this respect.

One point we can still make is that an isolated question can be used in support of an IS in a *direct* response to the question, as in (1). The following example (the denial-of-expectation interpretation), in conjunction with an additional utterance prior to the *wh*-question, can show a certain IS.[2]

(11) *Q*: I know Clyde married one of those rich women. But what happened to him after the woman died?

A: [Although Clyde married BERTHA]$_{T1}$, [he]$_{T2}$ [did not inherit a PENNY]$_{R}$.

Note that it is not crucial that there are two, discontiguous themes above.

However, it seems difficult to demonstrate an IS partition within the subordinate clause even with a direct question intended to single out a rheme in such an environment. As an attempt, let us consider Japanese in which a *wh*-word can be placed in a subordinate clause freely as in the following example (grammatical labels: TOP: topic/thematic, NOM: nominative, ACC: accusative, COP: copula, Q: question).

(12) Dare-ga Ken-o tasuketa-kara Naomi-ga koreta-no?
who-NOM Ken-ACC helped-because Naomi-NOM could.come-Q
"Naomi was able to come because Ken is helped by whom?"

An interesting point is that even this type of question cannot give rise to an IS partition within the subordinate clause. To see this, we use the property that no part of the rheme can be omitted in a response.

(13) A_1: #Erika.

A_2: Erika-ga tasuketa-kara.
Erika-NOM helped-because
"Because Erika helped (him)."

It is not possible to respond to such a question only with the constituent corresponding to the *wh*-word; the entire subordinate clause is needed in the response.

[2]One of the reviewers pointed out that the concessive-opposition may lead to separate utterances with their own IS's. Then, it would be analogous to the following type of parallel structure with two utterances and two IS's.

(1) *Q*: What's happened to John and Mary?

A: [John]$_{T}$[went to the ZOO]$_{R}$and [Mary]$_{T}$[went to the MUSEUM]$_{R}$.

If this can be shown, the IS analysis of (2) may be possible for concessive opposition. However, we note that (Kruijff-Korbayová and Webber) assume the same IS analysis for the both interpretations.

Note that it is perfectly fine to respond to certain *wh*-questions with just a noun. While this demonstration is not a proof, it appears consistent with our position.

Next, we look at intonation in English as an additional support for the argument that the IS analysis in (2) is hard to come by. As a non-native speaker of English, I cannot judge the appropriateness of intonation in English.[3] But I conjecture that the examples below can be used in support of the current position.

First, Steedman (2000) presents an example where the response to a question introduces an implication.

(14) *Q*: Does Marcel love opera?
A$_1$: Marcel likes MUSICALS. (H*)
A$_2$: Marcel likes MUSICALS. (L+H*)

Here, (*A*$_1$) is analyzed as a rheme (with the rheme tune), a complete utterance; (*A*$_2$) is analyzed as a theme (with the theme tune), with an implied rheme. For the case of (*A*$_2$), if the respondent thinks that the inquirer would not understand her implied rheme, she would have uttered a more explicit response depending on the implication, as in the following.

(15) *A*$_{2a}$: As Marcel likes MUSICALS (L+H*), he loves OPERA (H*).
A$_{2b}$: Although Marcel likes MUSICALS (L+H*), he doesn't love OPERA (H*).

Structurally, (*A*$_{2b}$) is parallel to (2). Now, let us suppose that some context allows the IS in (2). The intonation pattern of (2) would be as follows (applying Steedman 2000):

(16) Although [Clyde married]$_T$[BERTHA (H*)]$_R$, [he]$_T$[did not inherit a PENNY (H*)]$_R$.

My conjecture is that this intonation pattern is inconsistent with the hypothetical context (at least for denial of expectation). On the other hand, the intonation pattern of (15A$_{2b}$) seems possible for (2) for some context such as (11), but is inconsistent with the proposed IS.

In summary, lack of a convincing context for the IS analysis in (2) is a problem for Kruijff-Korbayová and Webber (2001). In addition, the proposed semantic motivation is not sufficient because it is needed independent of IS. One additional question concerning example (2) is whether there is another level of IS connecting the two clauses leading to a recursive IS, which is not discussed in their paper.

3 Recursive IS and Tripartite Structure

This section discusses the recursive analysis of Partee (1996) as seen in (3). Her motivation for this move is to mediate two analyses involving presupposition. The

[3] It would be possible to evaluate native speakers' intuition using synthesized speech/intonation applying the ideas discussed in Prevost (1995).

first is an analysis of Hajičová (1984), who argues that the distinction between presupposition and 'allegation' (potential presupposition, see below) is affected by IS. The second is that of Heim (1982), who adopts 'tripartite structure', a type of semantic structure closely associated with quantifier scopes, for her analysis of presupposition.

In order to discuss the issue at hand, let us observe an example from Hajičová (1984) (the possibility of the presupposition relevant to the current discussion is also shown to the right of each sentence).

(17) *a*. This time John's COUSIN caused our victory. ⇒ We won.
Negation: This time John's COUSIN didn't cause our victory. ⇒ We won.

b. This time John's cousin caused our VICTORY. ⇒ We won.
Negation: this time John's cousin didn't cause our VICTORY. ⇏ We won.

The proposition "we won" is a presupposition of (*a*) but only an 'allegation' of (*b*) because its negation cannot entail the proposition. Hajičová's argument is that when the presupposition-triggering material is in the theme (rheme), it results in a presupposition (allegation). Partee (1996) attempts to connect this analysis to Heim's analysis of presupposition along the recursive tripartite structure. According to Partee, the presupposition/allegation distinction can be observed recursively just as the tripartite structure is.

But the presupposition/allegation distinction can be observed within a noun phrase as well, as shown below.

(18) *a*. the RECORD of our arrest ⇒ We were arrested.
Negation: no RECORD of our arrest ⇒ We were arrested.

b. the record of our ARREST ⇒ We were arrested.
Negation: no record of our ARREST ⇏ We were arrested.

Thus, the distinction must be analyzed independent of IS as in the previous section. As pointed out by one of the reviewers, this distinction can be made in terms of the notions of CB (contextually-bound) vs. NB (non-bound) (e.g., Sgall et al. 1986). This situation seems to correspond to the point that the analysis of Kruijff-Korbayová and Webber (2001) could be cast within the foreground-background distinction of Steedman (2000).

The same recursive idea is discussed even further in Hajičová et al. (1998). Although they state that an IS can be associated with each embedded clause, this point was never emphasized in their earlier work (e.g., Sgall et al. 1986). In fact, Partee writes that she learned this point only after 1991 (Hajičová et al. 1998, p. 95). In addition, another paper of Hajičová et al. (1995) on a computational analysis of IS completely leaves out complex structures. This seems to imply that

the notion of recursive IS might not be as general or essential as argued in Partee (1996) and Hajičová et al. (1998)

Further, the growing trend in 'interpreting' IS is to adopt an informal, procedural view (Vallduví 1990) or a formal, dynamic view (Steedman 2000). The recursive approaches to IS have not discussed this aspect of IS analysis yet. The same comment also applies to another recursive approach of Hoffman (1995).

As in the case of (2), it hardly seems possible to come up with a question that directly confirm such an IS. Again, this does not reject Partee's proposal. But as before, the motivation cannot be the presupposition/allegation distinction because it is independent of IS. From the discussion in this and the previous sections, I would like to present two conjectures. First, the semantic/pragmatic effect that can be observed entirely within a noun phrase cannot be an IS effect. Second, IS only provides a bound on the domain of tripartite structure, a much weaker view of the relation between IS and tripartite structure.

4 Subordinate vs. Coordinate Structures

We begin this section with potential counterexamples to the main point of this paper. Then, I will discuss examples in Japanese, German, and in English in support of our position.

4.1 Potential Counterexamples

Although the written form of English is not very rich in marking IS, topicalization and focus movement can be considered to mark IS (Prince 1984). Bonnie Weber [p.c.] points out that the presence of one of these constructions within an *although*-clause can be a counterexample to the proposed position, and provided with the following examples from the British National Corpus (BNC).

(19) *a*. The shape seemed to be looking through a book, although what the book was Henry could not tell. [ASS 676]

b. His mother was always telling him that it was important for teachers to give, although what they were supposed to give she did not say. [HR 831]

c. Although what that could possibly be, I have no idea, Melissa thought to herself as she put down the phone. [GVP 1963]

First, I agree that topicalization/focus movement is a weak form of IS markers (Komagata 1999), but will question the status of the *although*-clauses observed in these examples. Next, in most of the potential counterexamples including (19 *a*, *b*), the *although*-clause follows the main clause. The only exception known to me is (19*c*). I would not provide a detailed explanation for this example at this point except for pointing out the following: the presence of topicalization/focus movement in the *although*-clause suggests that there is an IS division within the

although-clause, but there is no further, direct evidence of double IS in this example; I conjecture that the the *main* clause in (19*c*) is an afterthought (and thus a part of the theme). The position that sentence-initial *although*-clauses are thematic is also related to the following. While not categorical, the first components in an utterance tends to be thematic, as Halliday (1967) pointed out. In addition, according to Quirk et al. (1985, p. 919), one of the semantic characterizations of subordination is that the subordinate clause presents information as if it is presupposed as given rather than asserted as new. This description seems to apply to the majority of *although*-clauses, especially when the subordinate clause precedes the main clause (57% of all the instances involving an *although*-clause in the ACLDCI corpus from LDC). As a consequence, this paper will still be left with one potential counterexample (but not a number of counterexamples as commented by a reviewer). The remaining wide range of potential counterexamples will be accounted for in a fairly systematic manner in the following subsections. Note that the above potential counterexamples do not directly support the particular IS analysis in (2).

There are other related constructions, which are occasionally misunderstood as IS markers. In particular, both *it*-clefts and pseudoclefts do not necessarily mark an IS partition, esp. in embedded environments. According to the findings of Prince (1978), Collins (1991), and Delin (1995), *it*-clefts serve heterogeneous functions of marking IS, contrastiveness, and referential status. Also due to Prince (1978) and Collins (1991), the free relative part of a pseudocleft is either 'evoked' or 'inferrable', which is analogous to the referential status of the definite expression. As definite expressions can appear in themes and rhemes not marking information structure, pseudocleft cannot be a direct IS marker.

4.2 Morphology in Japanese

Next, we explore potential counterexamples in Japanese. This language has an explicit theme marker, i.e., the particle *wa* (Kuno 1973), which is useful for analyzing information structure. A caveat is that the same morpheme also serves as the contrastiveness marker. However, the contrastive case requires that there be a phonological prominence within the noun phrase that is suffixed with *wa* (summarized in Komagata 1999). Thus, the instances of *wa* suffixed to a non-prominent noun phrase can be considered thematic.

One of the properties of the thematic *wa* observed by several Japanese linguists including Kuno (1973) is that thematic *wa* cannot occur in the embedded environment. This is in accordance with our proposition. But Noda (1996) discusses a classification of subordinate clauses into the following categories, which include potential counterexamples.

(20)

Type	Examples	*ga*	*wa*
Strong	*(re)ba* "*if*", *toki* "*when*", *koto* "*that*" *node* (focal) "*because*"	yes	no
Weak	*node* "*since*", *ga* "*but*"	yes	yes

Although Noda says that the thematic *wa* cannot be used in strong subordinate clauses, he also lists several examples that he considers exceptional. But the *wa*-marked subjects in these examples are either the subject of the main clause or contrastive. Thus, they are not exceptions at all.

In addition, Noda lists several examples of weak subordinate clauses that involve a thematic *wa*. One case involves the conjunctive particle *ga* "*but*" as the sentence connector. But these examples are better classified as coordinate structures. The other case involves a 'subordinator' particle, but lacks the main clause as follows:

(21) memorii-wa ... mottomo anzen-na basyo-dakara-da.
memory-TOP most safe place-because-COP
"Because the memory is the safest place (for viruses to stay)."

The subject NP suffixed with *wa* is considered as the theme of this utterance. Even though this sentence contains a subordinator-like particle, it is an independent utterance. Then, it is no surprise that there is an IS partition within it, and thus, it is not a counterexample to the current position. In this case, the subordinator-like particle *dakara* "*because*" is better considered a discourse connector.

4.3 Syntax in German

We next investigate the *obwohl* "*although*"-clause in German, which is particularly interesting because only the matrix clause exhibits the verb-second (V2) phenomenon.

While the *obwhol*-clause typically has the verb-final pattern (i.e., subordinate clause), Günthner (1996) observes the growing tendency of the V2 configuration in the *obwhol*-clause in spoken colloquial German as shown below.

(22) A: DU ich brauch en kleinen STIFT
"hey I need a small pencil"
B: moment mal
"just a second"
A: obwhol NE eigntlich weiß ichs auch AUSwendig glaub ich
"although no actually I do know it by heart"

Günthner argues that this type of *obwhol*-clauses form a coordinate structure rather than a subordinate structure. Then, *obwhol* can be considered as a discourse connector. Günthner also analyzes the condition for using subordinate *obwhol*-clauses as follows: the relation between the main clause and the *obwhol*-clause is very loose (or independent illocutionary force for the *obwhol*-clause). According to Günthner, the availability of the two patterns and the condition for the *weil* "*because*"-clause are analogous to the *obwhol*-clause.

¿From the examples in Japanese (previous subsection) and German, we may infer the following. First, subordinate(-like) clauses can exist on their own (without the main clause) or weakly connected to the main clause. Second, this type of clauses can actually be considered as utterances and thus they may contain IS partitions.

4.4 *Although*-clause as a Coordinate Structure

The use of subordinate(-like) clauses as an independent utterance can be observed in English as well (Quirk et al. 1985, p. 564). For example, the word *because* here can be analyzed as a discourse connector because it connect the proposition "he did it" with the response as the reason for the proposition.

(23) *Q*: Why did he do it?
A: Because he was angry.

Similarly, the following example seems possible.

(24) A: I heard that you went to the park yesterday.
B: Although it was raining.

This example can be considered completely in parallel to (23). Such an example might be found in spoken corpora, but I have not been able to check this possibility. The word *although* is a concessive, discourse connector. For this type of utterance, it is natural to consider an IS that is felicitous to the context. While the distinction between coordinate and subordinate structures is not necessarily clear in English (e.g., Quirk et al. 1985, p. 927), the analogous distinction is clearly seen in the German examples because of the V2 phenomenon. While IS-marking too is not necessarily clear in English, this aspect is observed in the Japanese example.

Let us now turn to the case where the subordinate clause follows the main clause with particular placements of phonological prominence such as the following (Quirk et al. 1985, p. 1077).

(25) *a*. Raven didn't leave the party early because CAROL was there.
b. Raven didn't leave the party EARLY, because CAROL was there.

In fact, the same pattern seems possible with *although* as well (replacing *because* with *although* above). While the scope of the negation includes the subordinate clause in (*a*), it is not the case in (*b*). This suggests that there is a difference between (*a*) and (*b*) with respect to the strength of the connection. Analogous to the observation of Günthner (1996) in German, we may consider the entire sentence (*a*) as a single utterance, but the sentence (*b*) possibly as a coordinate structure consisting of two utterances. Although this analysis depends on the degree of connectedness between the two clauses, such an analysis would be possible as in Günthner (1996).

In summary, the presence of independent IS in the subordinate(-like) clause in the potential counterexamples is actually not inconsistent with the current position, and thus is not considered as counterexamples. I suggest that the analysis of complex structures proposed here is not specific to the *although* and *because*-clauses but applicable to subordinate clauses headed by various subordinators. In addition, we can make a related prediction based on the IS-related distinction between *since* (only thematic) and *because* (Quirk et al. 1985; Lambrecht 1994): the *since*-clause would not give rise to an independent utterance as the *because*-clause does.

5 Conclusion

This paper supports the traditional view of IS as a non-recursive, matrix-level phenomenon, and argues that the alternative views are not sufficiently motivated and that potential counterexamples to the traditional view can be analyzed in a systematic manner.

Structurally speaking, the domain of IS based on the traditional view is not fixed with respect to the DS unit (i.e., clause). As a consequence, IS appears *entangled* around clauses. Nevertheless, this type of entanglement seems inevitable to maintain the simple view of IS-DS relation (6) in the spirit of Daneš (1974). One implication of the current position with respect to NLP system design is as follows: due to the entanglement of IS's with clauses, we cannot arrange IS and DS processing sequentially. In order to deal with complex sentences, the IS and DS processing must proceed in parallel at some point.

In this paper, we discuss examples in English, German, and Japanese. Although the data are fairly consistent in my view, examples from other languages may reveal more about the issues under investigation. Thus, it would be very interesting to compare a larger number of languages along the current line.

Acknowledgements

I would like to thank Bonnie Webber for discussion and potential counterexamples; Mark Steedman for discussion about relevant topics at various points; Claire Gardent, Aravind Joshi, and Martha Palmer for comments on my dissertation, on which parts of this paper is based; and the reviewers of the workshop for stimulating comments.

Bibliography

Collins, P. C. (1991). *Cleft and pseudo-cleft constructions in English.* Routledge.

Daneš, F. (1974). Functional sentence perspective and the organization of the text. In F. Daneš (Ed.), *Papers on functional sentence perspective*, pp. 106–128. Prague: Academia and The Hague: Mouton.

Delin, J. (1995). Presupposition and shared knowledge in *it*-clefts. *Language and Cognitive Process 10*(2), 97–120.

Firbas, J. (1964). On defining the theme in functional sentence analysis. *Travaux Linguistiques de Prague 1*, 267–280.

Grosz, B. J. and C. L. Sidner (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics 12*(3), 175–204.

Günthner, S. (1996). From subordination to coordination? verb-second position in German causal and concessive constructions. *Pragmatics 6*(3), 323–356.

Hajičová, E. (1984). Presupposition and allegation revisited. *Journal of Pragmatics 8*, 155–167.

Hajičová, E., B. H. Partee, and P. Sgall (1998). *Topic-Focus Articulation, Tripart Structures, and Semantic Content*. Kluwer.

Hajičová, E., H. Skoumalová, and P. Sgall (1995). An automatic procedure for topic-focus identification. *Computational Linguistics 21*(1), 81–94.

Halliday, M. A. K. (1967). Notes on transitivity and theme in English (Part II). *Journal of Linguistics 3*, 199–244.

Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. Ph. D. thesis, University of Massachusetts at Amherst.

Hoffman, B. (1995). *The Computational Analysis of the Syntax and Interpretation of "Free" Word Order in Turkish*. Ph. D. thesis, University of Pennsylvania.

Komagata, N. N. (1999). *A Computational Analysis of Information Structure Using Parallel Expository Texts in English and Japanese*. Ph. D. thesis, University of Pennsylvania.

Kruijff-Korbayová, I. and B. L. Webber (2001). Concession, implicature and alternative sets. In *Proceedings of the International Workshop on Computational Semantics (IWCS-4), Tilburg, Jan., 2001*.

Kuno, S. (1973). *The Structure of the Japanese Language*. MIT Press.

Lambrecht, K. (1994). *Information Structure and Sentence Form: Topic, focus, and the mental representations of discourse referents*. Cambridge University Press.

Mathesius, V. (1975). *A Functional Analysis of Present Day English on a General Linguistic Basis, edited by Josef Vachek*. The Hague: Mouton.

Noda, H. (1996). *WA-to GA (WA and GA)*. Kuroshio.

Partee, B. H. (1996). Allegation and local accommodation. In B. H. Partee and P. Sgall (Eds.), *Discourse and meaning: papers in honor of Eva Hajičová*, pp. 65–86. John Benjamins.

Prevost, S. (1995). *A Semantics of Contrast and Information Structure for Specifying Intonation in Spoken Language Generation.* Ph. D. thesis, University of Pennsylvania.

Prince, E. F. (1978). A comparison of wh-clefts and *it*-clefts in discourse. *Language 54*(4), 883–906.

Prince, E. F. (1984). Topicalization and left-dislocation: A functional analysis. In S. J. White and V. Teller (Eds.), *Annals of New York Academy of Sciences, Vol. 433: Discourses in Reading and Linguistics*, pp. 213–225. New York: The New York Academy of Sciences.

Quirk, R., S. Greenbaum, G. Leech, and J. Svartvik (1985). *A Comprehensive Grammar of the English Language*. Longman.

Rooth, M. E. (1985). *Association with Focus.* Ph. D. thesis, University of Massachusetts at Amherst.

Sgall, P., E. Hajičová, and J. Panevova (1986). *The meaning of the sentence in its semantic and pragmatic aspects.* D. Reidel.

Steedman, M. (2000). Information structure and the syntax-phonology interface. *Linguistic Inquiry 31*(4), 649–689.

Vallduví, E. (1990). *The informational component.* Ph. D. thesis, University of Pennsylvania.

Information Structure and the Interpretation of "otherwise"

IVANA KRUIJFF-KORBAYOVÁ
Computerlinguistik, Universität des Saarlandes, Saarbrücken, Germany
korbay@coli.uni-sb.de

BONNIE LYNN WEBBER
Informatics, University of Edinburgh, 80 South Bridge, Edinburgh EH1 1HW, U.K.
bonnie@dai.ed.ac.uk

ABSTRACT. We have been investigating whether and how the interpretation of discourse connectives is sensitive to the Information Structure (IS) of the clauses or sentences they relate. Here we focus on the anaphoric connective "otherwise" and show how the IS of its antecedent affects what condition it can be "otherwise" to. This work is part of a larger enterprise aimed at understanding what role(s) sentence-level IS plays in the interpretation of larger units of discourse.

1 Introduction

It is well-known that *Information Structure* (**IS**) influences the interpretation of individual sentences. Of the famous sign in the London Underground, "Dogs must be carried", Halliday (1970) observes that this text can be pronounced with different intonation patterns, e.g., (1) vs. (2) reflecting different IS. Thereby, different instructions (here, paraphrased in italics) are conveyed to passengers. One supposes that (2) was not the intention of the London Transport Authority.

(1) Dogs must be CARRIED.
H* LL%
If there is a dog, carry it.

(2) DOGS must be carried.
H* LL%
Carry a dog.

In English, IS is most often conveyed by intonation. In languages with freer word order, differences in IS are most often conveyed by different word ordering. For example, the Czech counterparts of (1) and (2), conveying the same instructions to the hearer, are (3) and (4), respectively:

Paper presented at the ESSLLI 2001 Workshop on
Information Structure, Discourse Structure and Discourse Semantics
Edited by Ivana Kruijff-Korbayová and Mark Steedman.

(3) Psi se musí NÉST.
Dogs$_{nom}$ refl must$_{3pl}$ carry$_{inf}$

(4) Musí se nést PSI.
Must$_{3pl}$ refl carry$_{inf}$ dogs$_{nom}$

Over the past decade, the understanding of IS within the sentence has been enriched by intensive research in formal semantics. It is now widely accepted that IS affects both interpretation and realization, even though there is no uniform account. However, much less is known about what, if any, use is made of IS beyond clause and sentence boundaries and how IS interacts with other aspects of discourse structure and semantics. Our work extends the repertoire of IS-sensitive accounts in this direction. In this paper, we concentrate on how the IS of a previous sentence or clause can affect the meaning projected through the subsequent adverbial discourse connective "otherwise" ("jinak", in Czech). We show that an IS-based account of its meaning provides access to contextually appropriate interpretations that are unavailable to accounts that ignore IS.

Webber *et al.* (1999) have argued that "otherwise" contributes meaning to the discourse in part through structure, in part through anaphora: roughly, they say that the *complement* of the anaphorically-derived argument of "otherwise" serves as a *condition* under which the interpretation of its structural matrix holds.[1] As might be expected, different ways of resolving the anaphoric argument lead to different interpretations, as in (5a) vs. (5b):

(5) If you have brought a dog, you must pay 50p.
   a. Otherwise you will not be allowed to enter.
   b. Otherwise you can come in for free.

which can be paraphrased by resolving the anaphor and making the anaphorically-derived condition explicit:

(6) a. *If you have brought a dog and you do not pay 50p, you will not be allowed to enter.*
   b. *If you have not brought a dog, you can come in for free.*

Here, the antecedent used in (6a) is the preceding main clause, while that for (6b) is the preceding "if"-clause.[2]

[1] (Webber et al. 2001) present as evidence for this, *inter alia*, the fact that the first argument of "otherwise" may not be explicit, but rather have to be derived by inference from the previous discourse, and the fact that it can behave like a 'donkey' pronoun, deriving its first argument from a relative clause — e.g., "Farmers who beat their donkeys would *otherwise* be beating their wives."

[2] As with anaphoric pronouns, an automated procedure for resolving anaphoric "otherwise" must be able to reject contextually inappropriate ways of resolving it as in

(6a′) *If you have not brought a dog, you will not be allowed to enter.*

(6b′) *If you have brought a dog and you do not pay 50p, you can come in for free.*

But we do not consider this aspect of the problem any further in this paper.

That IS can affect what conditions can be derived can be seen by considering "otherwise" in the context of two different single-clause utterances, which differ only in their IS — here (7) vs. (8) in both English and Czech.[3]

(7) You must CARRY a dog. Otherwise you might get HURT.
H* LL% H*LL%

Psa musíte NÉST. Jinak byste mohli přijít k úrazu.
Dog$_{acc}$ must$_{2pl}$ carry. Otherwise be$_{2pl}$ could$_{pl}$ come to injury$_{dat}$.

(8) You must carry a DOG. Otherwise you might get HURT.
H*LL% H*LL%

Musíte nést PSA. Jinak byste mohli přijít k úrazu.
Must$_{2pl}$ carry dog$_{acc}$. Otherwise be$_{2pl}$ could$_{pl}$ come to injury$_{dat}$.

The "otherwise" clause in (7) will be interpreted as warning the hearer (H) that H might get hurt if s/he has a dog but isn't carrying it (e.g., H might get tangled up in the dog's lead). On the other hand, the "otherwise" clause in (8) warns H that s/he might get hurt if not carrying a dog, period (e.g., H might be walking past fanatical members of the Royal Kennel Club).

If the IS of one sentence or clause can affect how another is interpreted, then IS must be incorporated into an account of discourse interpretation and discourse updating. This we do in terms of Rooth's notion of an *alternative set* (Rooth 1985; Rooth 1992) and the alternative-set semantics of information structure worked out in (Steedman 2000a; Steedman 2000b), and refining our earlier presentations in (Kruijff-Korbayová and Webber 2000a; Kruijff-Korbayová and Webber 2000b).

The paper is organized as follows: In Section 2 we present the approach to IS and IS-sensitive context updating we are employing. In Section 3 we describe our IS-sensitive analysis of α *otherwise* β where α is a simple sentence. In Section 4 we describe our IS-sensitive analysis of α *otherwise* β where α is a complex sentence, making more options available. Section 5 concludes the paper and delineates the future directions of this work.

2 Information Structure and Context Updating

The notion of IS we are employing originates in the work of Mathesius (1975), and has been elaborated in subsequent work within the Prague School (Sgall et al. 1986) and by others, e.g., (Firbas 1992, Halliday 1985, Steedman 2000b). Specifically, we adopt the formal account presented in (Steedman 1996; Steedman 2000a; Steedman 2000b) which (1) provides a well worked out compositional semantics of English intonation in IS terms; (2) interprets the elements of IS in terms of alternative sets, and (3) assumes a general IS-sentence notion of discourse context update. Leaving terminological differences aside, Steedman's account is by and

[3]Throughout the paper, SMALL CAPITALS indicate intonation centers (pitch accents), thereby distinguishing Focus$_{is}$ from Background$_{is}$ within both Theme$_{is}$ and Rheme$_{is}$.

large straight-forwardly compatible with the Prague School approach, and thus when analyzing Czech examples, we can combine Steedman's account with Sgall and Hajičová's ideas relating IS and word order (Hajičová and Sgall 1987; Sgall et al. 1986).

Building on the findings originating in the Prague School (Firbas 1992; Mathesius 1975; Sgall et al. 1986), Steedman recognizes two dimensions of IS: The first defines a partitioning at the sentence-level into *Theme*$_{is}$ and *Rheme*$_{is}$; the second is a further partitioning of both into *Background*$_{is}$ and *Focus*$_{is}$.[4] The latter partitioning is related to Halliday's *Given-New* dichotomy (Halliday 1970; Halliday 1985) and concerns distinguishing the Theme$_{is}$ and the Rheme$_{is}$ from other alternatives that the context makes available.

In English, Czech and many other languages, IS is established as a result of an interplay of intonation, word order and grammatical structure. Below we give three of the possible IS partitions into Theme$_{is}$-Rheme$_{is}$ that Steedman's approach provides for the string "You should carry the dog".[5] The situation is one in which the first author (IKK) is transporting a dog, a large bag and a trolley by the Underground, and asks the second author (BW) a question, which helps to fix the IS of the reply.

(9) Q: How should I transport the DOG?

A: You should CARRY the DOG.
H* L+H*
Theme *Rheme* *Theme*

i. θ(9): $\lambda Q.\ Q(h, {}^{\star}dog_1)$ ρ(9): $\lambda x.\lambda y.\ {}^{\star}carry(x,y)$

ii. θ-AS(9): $\{\exists Q.\ Q(h,dog_1), \exists Q.\ Q(h,bag_3), \exists Q.\ Q(h,trolley_4)\}$
ρ-AS(9): $\{lead(h,dog_1), carry(h,dog_1), wheel(h,dog_1)\}$

(10) Q: Who should carry the DOG?

A: YOU should carry the DOG.
H* L LH%
Rheme *Theme*

i. θ(10): $\lambda x.\ carry(x, {}^{\star}dog_1)$ ρ(10): $\lambda Q.\ Q({}^{\star}h)$

ii. θ-AS(10): $\{\exists Q.\ Q(dog_1), \exists Q.\ Q(bag_3), \exists Q.\ Q(trolley_4)\}$
ρ-AS(10): $\{carry(h,dog_1), carry(s,dog_1), carry(officer_5,dog_1)\}$

[4] Alternative terms used for similar (but not identical) IS partitions in other works are, e.g., Topic-Focus (Sgall et al. 1986), Background(=Link+Tail)-Focus (Vallduví 1992). We adopt Steedman's terms, but add the subscripts in Theme$_{is}$, Rheme$_{is}$ and Background$_{is}$, Focus$_{is}$ in order to avoid confusion with some other uses of the same terms.

[5] For the time being, we ignore the modality introduced by "should" and any aspects of the speech-act beyond simple assertion.

(11) Q: What should I CARRY?

A: You should CARRY the DOG.
L+H* LH% H*LL%
Theme *Rheme*

i. θ(11): $\lambda x.\ ^{\star}carry(h,x)$ ρ(11): $\lambda Q.\ Q(^{\star}dog_1)$

ii. θ-AS(11): $\{\exists x.\ wheel(h,x),\ \exists x.\ push(h,x),\ \exists x.\ carry(h,x)\}$
ρ-AS(11): $\{carry(h,bag_3), carry(h,dog_1)\}$

For each sentence, (i) provides a simplified *IS-partitioned logical form*, where θ and ρ are operators which 'wrap' Theme$_{is}$ and Rheme$_{is}$, respectively. Within Theme$_{is}$ and Rheme$_{is}$, asterisks on terms (e.g., $\star carry$) indicate elements that belong to the respective Focus$_{is}$. These IS-partitioned logical forms represent the linguistic meaning of the sentences, and serve as input for a discourse (context) *update function* described below. (ii) indicates the Theme$_{is}$ alternative set (θ-AS) and Rheme$_{is}$ alternative set (ρ-AS), which are explained below. Because each example contains Focus$_{is}$ within Theme$_{is}$ (indicated by a *-term), which entails contrast with a previous Theme$_{is}$ (and hence alternatives to contrast with), each θ-AS contains more than one element. (Without pitch accents in Theme$_{is}$, and thus without contrast, the θ-AS would be a singleton set.)

2.1 Alternative Set Semantics for IS

Elaborating on Rooth's alternative semantics (Rooth 1992), Steedman assigns the following semantics to IS (cf. Steedman 2000a):

- Theme$_{is}$ presupposes a *Rheme$_{is}$-alternative set* (ρ-*AS*).
- Focus$_{is}$ within Rheme$_{is}$ restricts the ρ-AS to the singleton set corresponding to the asserted proposition.
- Theme$_{is}$ also presupposes a *Theme-alternative set* (θ-*AS*).
- Focus$_{is}$ within Theme$_{is}$ restricts the θ-AS to the singleton set corresponding to Theme$_{is}$.

ρ-AS corresponds to what Rooth calls the *contextual alternative set* (Rooth 1985; Rooth 1992). θ-AS is a set of alternative themes with respect to the context, corresponding to what Rooth calls the *question alternative set.* The notion of alternative set is also closely related to the notion of *secondary denotation* (Karttunen and Peters 1979).

Following (Steedman 2000a), we take ρ-AS to be a subset of the propositions supported by the context, whose characteristic function is obtained systematically from the IS-partitioned logical form. As noted in (Steedman 2000a, p.10), alternative sets may not be exhaustively known to hearers, and in practice one would want to compute with a more abstract form.

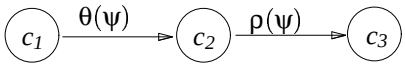


Figure 10.1: IS-sensitive update of context c_1 with ψ: $c_1[\theta(\psi)]c_2[\rho(\psi)]c_3$

2.2 IS-sensitive Context Updating

We follow (Krifka 1993; Kruijff-Korbayová 1998; Steedman 2000a) in defining the updating of an input context c_1 with an IS-partitioned logical form p as comprising two phases, a *Theme$_{is}$ update phase* ($c_1[\theta(\psi)]c_2$) and a *Rheme$_{is}$ update phase* ($c_2[\rho(\psi)]c_3$), where c_2 and c_3 are resulting contexts. (See Figure 2.2).

In the *Theme$_{is}$ update phase*, the input context c_1 is checked as to whether it supports or can accommodate the presuppositions of the theme $\theta(\psi)$ – namely, the Theme$_{is}$-alternative set θ-AS and the Rheme$_{is}$-alternative set ρ-AS. This yields a restricted context c_2 where $\theta(\psi)$ holds. In the *Rheme$_{is}$ update phase*, one alternative according to the ρ-AS is selected, which yields the final context c_3. Updating fails if either update phase does.

3 IS and "otherwise": single-clause antecedents

As noted earlier, Webber et al. (1999) have argued that "otherwise" has one argument established anaphorically, and one provided structurally. It is the anaphoric argument that provides the condition that "otherwise" appeals to and whose IS, we are arguing, the interpretation of "otherwise" must be sensitive to. This does not mean, however, that the antecedent of "otherwise" is limited to IS-partitioned utterances: just that IS-partitioning provides relevant possibilities.

Because example (5), given earlier, contains two different clauses (main and subordinate) that can serve as antecedents for "otherwise", there are at least the two possible conditions — shown in (6a) and (6b) — that *otherwise* β can derive and apply to the interpretation of β. Examples such as this are discussed in Section 4. Here we focus on cases where the condition that "otherwise" appeals to derives from a single clause antecedent.

Even here, the analysis in (Webber et al. 1999) must be refined in two ways to take account of IS:

1. The antecedent (A) of "otherwise" should not be treated as an atomic unit: rather, "otherwise" can appeal to a condition "C" deriving from either A's Theme$_{is}$ or its Rheme$_{is}$.

2. The context that β is asserted with respect to is not strictly worlds consistent with the real world or the current discourse context other than those "C" worlds: rather, it may or may not be consistent with the Theme$_{is}$ of its antecedent A as well.

(In the following examples, "Otherwise β" itself has an IS-partitioning. However, we do not explicitly indicate it, because it is not relevant to the points we are advancing. We will make a point about the IS-status of "otherwise" itself at the end of this section.)

The examples below address the first point, showing that the condition that "otherwise" appeals to may derive either from the Theme$_{is}$ of its antecedent — we call this the *full Theme$_{is}$-complement condition*– as in (12i), or from its Rheme$_{is}$ — we call this the *full Rheme$_{is}$-complement condition*– as in (12ii). (The corresponding paraphrases of the "otherwise β" are shown in italics.)

(12) Q. What should I do at a RED LIGHT?

i. At a red light, STOP. Otherwise you can go straight on.
H*LL%
Theme *Rheme*
Na červenou zastavte. Jinak můžete jet rovně.
At red$_{acc}$ stop$_{imp2pl}$ Otherwise can$_{2pl}$ go$_{inf}$ straight.
If the light is not red, you can go straight on.

ii. At a red light, STOP. Otherwise you will get a ticket.
H*LL%
Theme *Rheme*
Na červenou zastavte. Jinak dostanete pokutu.
At red$_{acc}$ stop$_{imp2pl}$ Otherwise get$_{2pl}$ fine$_{acc}$
If (the light is red and) you do not stop, you will get a ticket.

The full Theme$_{is}$-complement condition in (12i) corresponds to "not being at a red light" (or, "the light not being red"). The full Rheme$_{is}$-complement condition in (12ii) corresponds to "not stopping".

As to our second point, the context in which β is asserted, at issue is the Theme$_{is}$ of the antecedent of "otherwise". When "otherwise" appeals to the full Theme$_{is}$-complement condition, there is only one context with respect to which β can be asserted, namely the initial context before asserting α. The context updated with α's Theme$_{is}$ is irrelevant, because it is incompatible with the full Theme$_{is}$-complement condition: example (12i) cannot be interpreted as *If the light is red and if the light is not red, you can go straight on*.

On the other hand, when "otherwise" appeals to the full Rheme$_{is}$-complement condition, it appears that the IS-partitioning makes two contexts available with respect to which β can be asserted: the initial context before asserting α's Theme$_{is}$ (as in (13i)), and the context updated with α's Theme$_{is}$ (as in (13ii)).

(13) Q. When should I STOP?

i. Stop at a red LIGHT. Otherwise you can go straight on.
H* LL%
Theme *Rheme*

Zastavte na červenou. Jinak můžete jet rovně.
Stop$_{imp2pl}$ at red$_{acc}$. Otherwise can$_{2pl}$ go$_{inf}$ straight.
If the light is not red (i.e., in other conditions than being at a red light), you can go straight on.

ii. Stop at a red LIGHT. Otherwise you might get rear-ended.
H* LL%
Theme *Rheme*
Zastavte na červenou. Jinak by do vás
Stop$_{imp2pl}$ at red$_{acc}$. Otherwise be-aux$_{3sg}$ into you$_{gen}$
někdo mohl narazit.
somebody$_{nom}$ might$_{sg}$ bump$_{inf}$
If you stop and the light is not red, you might get rear-ended.

The sense that (13ii) conveys both in English and in Czech that one should *only* stop at a red light, comes from this interpretation of "otherwise" in terms of stopping under all conditions other than the light being red.

There is one further point to make before presenting our analysis of this phenomenon and review of these examples in more detail. That is that "otherwise" itself is a contrastive (part of the) Theme$_{is}$, and what we have seen here are different ways in which it relates to the input context: in example (12i), "otherwise" contrasts with the preceding Theme$_{is}$ (and therefore picks up the full Theme$_{is}$-complement condition), while in examples (12ii), (13i) and (13ii), it contrasts with the preceding Rheme$_{is}$ (and therefore picks up the full Rheme$_{is}$-complement condition). Example (12i$'$) below illustrates this Theme$_{is}$ contrast even more vividly, in that the pitch accents on "red" as Focus$_{is}$ within the Theme$_{is}$ of the first sentence indicates the speaker's awareness of alternatives that the "otherwise" sentence then explicates.

(12i$'$) At a RED light, STOP. Otherwise you can continue.
L+H* LH% H*LL%
Theme *Rheme*

3.1 Analysis

We propose the following IS-sensitive refinement of the analysis of "otherwise" in (Webber et al. 1999): Let us assume that α is the antecedent of *otherwise* β, and c_0 is the context prior to updating with α (rather than the real world). The IS-sensitive update enables us to distinguish between the following subsets of c_0:

- the subset where α's Theme$_{is}$ and alternatives to α's Rheme$_{is}$ hold (i.e., excluding α itself);
- the subset where alternatives to α's Theme$_{is}$ hold;
- the subset where alternatives to α's Rheme$_{is}$ hold (irrespective of Theme$_{is}$).

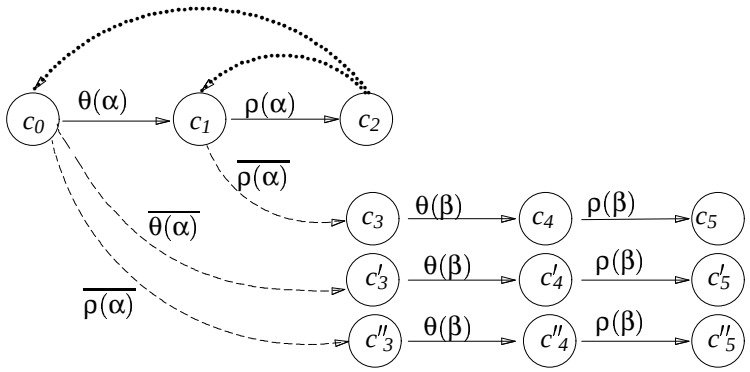


Figure 10.2: IS-sensitive updating with "α. Otherwise β".
The dotted arcs indicate the two possible ways of resolving "otherwise" with respect to the simplest IS-partitioning of the antecedent, and the dashed arcs indicate the transitions to the corresponding contexts.

In (Webber et al. 1999), β is asserted solely with respect to $c_0 - c_\alpha$, the subset of c_0 where alternatives to α hold. Here we refine this with respect to the above three subsets, defining how an input context is updated with the sequence "α. Otherwise β" for a single clause α:

1. c_0 is updated with α as described in Section 2.2: $c_0[\theta(\alpha)]c_1[\rho(\alpha)]c_2$.
2. Updating with "otherwise β" involves either:
   - c_1 being updated with "otherwise β", which involves constructing context c_3 as the Rheme$_{is}$-complement of c_2 with respect to c_1 and then updating this context with β: $c_1[\overline{\rho(\alpha)}]c_3[\theta(\beta)]c_4[\rho(\beta)]c_5$
   - c_0 being updated with "otherwise β" in one of two ways:
     - Context c_3' is constructed as the Theme$_{is}$-complement of c_1 with respect to c_0 and then c_3' is updated with β: $c_0[\overline{\theta(\alpha)}]c'_3[\theta(\beta)]c'_4[\rho(\beta)]c'_5$
     - Context c_3'' is constructed as the Rheme$_{is}$-complement of c_1 with respect to c_0 and then c_3'' is updated with β: $c_0[\overline{\rho(\alpha)}]c''_3[\theta(\beta)]c''_4[\rho(\beta)]c''_5$

These context-updating possibilities are shown schematically in Figure 10.2.

3.2 Examples

We now demonstrate this detailed IS-sensitive updating analysis for examples from the introduction to this section. Example (12i) repreated in (14) shows how the analysis applies to the case where a full Theme$_{is}$-complement condition is derived from the Theme$_{is}$ of the antecedent of "otherwise" and β is asserted with respect to the initial context, c_0. (Recall that this is the *only* context-updating possibility.)

(14) At a red light, STOP. Otherwise you can go straight on.

H*LL%

Theme *Rheme*

$c_0[\lambda P.at(h, red\_light) \wedge P]c_1[stop(h)]c_2$

$c_0[\overline{at(h, red\_light)}]c'_3[\lambda Q.Q(h)]c'_4[go\_straight(h)]c'_5$

If you are not at a red light, you can go straight on.

Example (13i) repeated in (15) shows how the analysis applies to the case where a full Rheme$_{is}$-complement condition is derived from the Rheme$_{is}$ of the antecedent. As shown above, there are two possible contexts against which β can be asserted. In (15), β is asserted with respect to the initial context, i.e. c_0.

(15) Stop at a red LIGHT. Otherwise you can go straight on.

H* LL%

Theme *Rheme*

$c_0[\lambda P.P \wedge stop(h)]c_1[at(h, red\_light)]c_2$

$c_0[\overline{at(h, red\_light)}]c_3[\lambda Q.Q(h)]c_4[go\_straight(h)]c_5$

If the light is not red (in other conditions than being at a red light), you can go straight on.

In contrast with (13i) is example (13ii), repeated in (16). While it appeals to the full Rheme$_{is}$-complement condition, β is asserted with respect to the context updated with α's Theme$_{is}$, i.e. context c_1.

(16) Stop at a red LIGHT. Otherwise you might get rear-ended.

H* LL%

Theme *Rheme*

$c_0[\lambda P.P \wedge stop(h)]c_1[at(h, red\_light)]c_2$

$c_1[\overline{at(h, red\_light)}]c_3[\lambda Q.Q(h)]c_4[get\_rear\_ended(h)]c_5$

If you stop and the light is not red, you might get rear-ended.

The examples in this section demonstrate a range of possible antecedents for "otherwise" that are not available without taking IS into account.

4 IS and "otherwise": complex-clause antecedents

We now turn to examples of the form considered in (Webber et al. 1999), where the condition used for interpreting "otherwise" comes from a complex sentence of the form *If* ϕ*, then* ψ. Here we show that the same analysis holds as before, with one addition:

- When both the subordinate clause and some element(s) from the main clause are included in the Theme$_{is}$, a third possibility for deriving the condition to which "otherwise" appeals is made available: the condition can be derived from a part of the Theme$_{is}$ of the antecedent.

There are somewhat more examples to review, because in both English and Czech, the main clause can belong entirely to the Theme$_{is}$ or to the Rheme$_{is}$ (i.e., the boundary between Theme$_{is}$ and Rheme$_{is}$ can coincide with the clause boundary between ϕ and ψ), or the main clause can be divided over the Theme$_{is}$ and the Rheme$_{is}$ (i.e., the boundary between Theme$_{is}$ and Rheme$_{is}$ splits ψ). The first of these possibilities is discussed in Section 4.1, the second in Section 4.2.

4.1 IS-boundary coinciding with clause boundary

When the IS-boundary between Theme$_{is}$ and Rheme$_{is}$ coincides with the clause boundary between ϕ and ψ, the Theme$_{is}$ (Rheme$_{is}$) consists of ϕ, and the Rheme$_{is}$ (Theme$_{is}$) of ψ. The examples below show that, as with simple clause antecedents, the condition that "otherwise" appeals to may derive either from the Theme$_{is}$ of its antecedent (the *full Theme$_{is}$-complement condition,* as in (17i)), or from its Rheme$_{is}$ (the *full Rheme$_{is}$-complement condition,* as in (17ii)).

(17) Q. What should I do if the light is RED?

A. If the light is RED, stop at the CORNER.
L+H*LH% H*LL%
Theme *Rheme*

A′. Stop at the CORNER if the light is RED.
H*LL% L+H*LH%
Rheme *Theme*

i. Otherwise you can go straight on.
If the light is not red, go straight on.

ii. Otherwise you will get a ticket.
If the light is red and you do not stop at the corner, you will get a ticket.

The full Theme$_{is}$-complement condition in (17i) corresponds to "the light not being red", and the full Rheme$_{is}$-complement condition in (17ii) corresponds to "not stopping" (as with the simple antecedent examples in (12i) and (12ii)).

When "otherwise" appeals to the full Theme$_{is}$-complement condition, there is only one context with respect to which β can be asserted, namely the initial context before asserting α. The context updated with α's Theme$_{is}$ is irrelevant, because it is incompatible with the full Theme$_{is}$-complement condition: example (17i) cannot be interpreted as *If the light is red and if the light is not red, you can go straight on.*

On the other hand, when "otherwise" appeals to the full Rheme$_{is}$-complement condition, it appears that the IS-partitioning makes two contexts available with

respect to which β can be asserted: the initial context before asserting α's Theme$_{is}$ (as in (18i)), and the context updated with α's Theme$_{is}$ (as in (18ii)).

(18) Q. When (i.e., under what conditions) should I stop at the CORNER?

A. Stop at the CORNER (L+H*LH%) — *Theme* if the light is RED. (H*LL%) — *Rheme*

A′. If the light is RED (H*LL%) — *Rheme* stop at the CORNER. (L+H*LH%) — *Theme*

i. Otherwise you can go straight on.
If the light is not red, (you needn't stop) and you can go straight

ii. Otherwise you might get rear-ended.
If you stop at the corner and the light is not red, you might get rear-ended.

Again, the sense that (18ii) conveys that one should *only* stop at a red light, comes from this interpretation of "otherwise" in terms of stopping under all conditions other than the light being red. It appears very difficult to get the variant of (18) with the preposed rhematic "if"-clause. We think that this is because this IS-partitioning requires a marked intonation pattern that may be difficult in English.

4.2 IS-boundary splitting the main clause

When the IS-boundary between Theme$_{is}$ and Rheme$_{is}$ "splits" ψ, the Theme$_{is}$ (Rheme$_{is}$) consists of ϕ and a part of ψ, while the rest of ψ belongs to the Rheme$_{is}$ (Theme$_{is}$). As before, "otherwise" can appeal to the *full Theme$_{is}$-complement condition* and the *full Rheme$_{is}$-complement condition,* but another possibility is that the condition derives from just that part of the Theme$_{is}$ in the matrix clause, as illustrated below. This we call the *partial Theme$_{is}$-complement condition.*

(19) Q. Where do you buy wine if it's SUNDAY?

A. If it's SUNDAY , we buy wine (L+H*LH%) — *Theme* over the STATE LINE. (H*LL%) — *Rheme*

Otherwise we just buy beer.

a. *If we don't buy wine, we buy beer.*

b. *If it is Sunday and we don't buy wine, we buy beer.*

The partial Theme$_{is}$-complement condition in (19) corresponds to "we do not buy wine". The reason we give two possible paraphrases of "otherwise we just buy beer" (*otherwise* β) is that the initial contex c_0 can be updated with this partial Theme$_{is}$-complement (γ) in either of two ways (shown schematically in Figure 10.3):

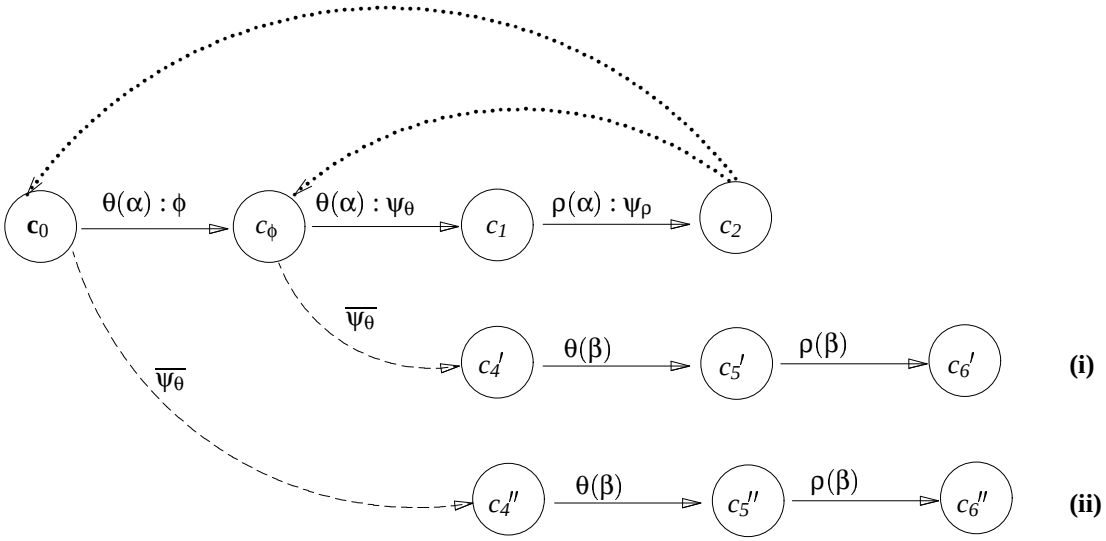


Figure 10.3: IS-sensitive updating with "α. Otherwise β" involving a partial Theme$_{is}$-complement condition.

- It can be updated just with γ, asserting β with respect to the result, as in (19a);
- It can be just updated with that part of α's Theme$_{is}$ in the "if-clause" and then γ, asserting β with respect to the result, as in (19b).

c_0 cannot be updated with respect to the entire Theme$_{is}$ and then γ because the two are incompatible.

Finally, we consider the case where the "if-clause" belongs to the Rheme$_{is}$ of α, as in example (20a). Of concern is the possibility that "otherwise" appeals to a condition derived from that part of the Rheme$_{is}$ in the matrix clause, what we call the *partial Rheme$_{is}$-complement condition.*

(20) Q. What should I do AFTER 5PM?

A. After 5pm take a BREAK, if you are TIRED.
LH% H* H*LL%
Theme *Rheme*

a. Otherwise, you'll start making mistakes.
If it is after 5pm, and if you are tired, and you do not take a break, you'll start making mistakes.

b. Otherwise, carry on until the job is done.
If it is after 5pm, and if you are not tired (and you do not take a break), carry on until the job is done.

In example (20a), the partial Rheme$_{is}$-complement condition corresponds to "you do not take a break". Here, the only context in which it makes sense to assert

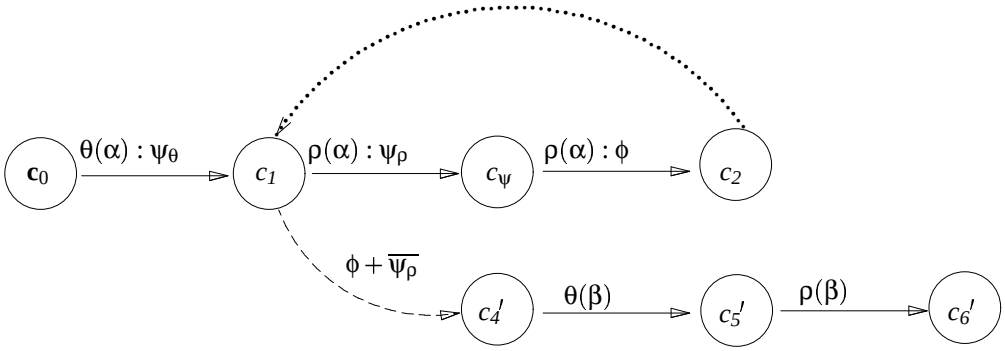


Figure 10.4: IS-sensitive updating with "α. Otherwise β" involving a partial Rheme$_{is}$-complement condition.

"Otherwise β" (with "otherwise" appealing to this condition) is the context resulting from updating the initial context c_0 with α's Theme$_{is}$ and with that part of α's Rheme$_{is}$ constituted by the "if-clause". This is shown schematically in Figure 10.4.

We note example (20b) because we are uncertain whether "otherwise" appeals to just the complement of the "if-clause" (i.e., part of α's Rheme$_{is}$) or the complement of the entire Rheme$_{is}$ of α. Here we feel that more research is needed concerning the status of (postposed) subordinated clauses with respect to the IS-partitioning, in particular, whether they should be treated within the matrix clause, or as separate utterances (with their own IS-partitioning) (cf. (Günthner 1996) for a discussion based on spoken data; cf. also Komagata's paper at this workshop).

5 Conclusions and Further Research

While we must still complete our discussion of "otherwise" with complex antecedents, we hope that we have convinced the reader that that IS is crucial to any account of the semantics of "otherwise". We recognize that several problems remain unaddressed:

- As already noted, we have not identified the range of things that can serve as antecedents (i.e., provide conditions) for "otherwise" nor identified from where in the discourse they can come, other then the previous clause or an embedded relative clause. More importantly, we have not said *why* they provide conditions for "otherwise": That is, we haven't addressed the basic problem of what (alternative) conditions a speaker may have in mind and what features of language give evidence for them.

  Here we have claimed that the *alternative sets* of Information Structure give such evidence. But they are clearly not the only evidence (e.g., multi-clause

antecedents seem possible for "otherwise", as do accommodated antecedents, both of which would be outside the realm of IS). And a more parsimonious analysis of the data we have presented may not involve IS at all: For example, Matthew Stone (personal communication) has pointed out that all our examples involve generics, which can be analysed as involving a set of *cases under discussion.* We must understand whether and if so, how, these two concepts are related.

- There are cases of postposed "if"-clauses that are best analysed as having their own IS – as in:

(21) Q. When should I stop?

A1. Stop at a red light — that is, if it's a weekend.
H*LL% H*LL%
Theme *Rheme* *Theme* *Rheme*

A2. Stop at a red light
H*LL%
Theme *Rheme*

— that is, if you don't want to get a ticket.
H*LL%
Theme *Theme* *Rheme*

In the case of (21:A2), the "if" clause is playing a role similar to an "otherwise" clause, so that adding an "otherwise" clause appears redundant. In the case of (21:A1), it may be that "otherwise" can either combine the rhemes into a single condition or consider the later one as a condition of its own.

- The role that the "otherwise" clause plays with respect to the preceding discourse is clearly tied, at least partially, with the condition it is taken to be otherwise to: In the complex "if"-clause antecedents we have discussed, being otherwise to the Rheme_{is} (in main or subordinate clause) provides an *explanation,* while being otherwise to the Theme_{is} provides a *elaboration* of what holds in other circumstances. While this may call into question the notion in Rhetorical Structure Theory (Mann and Thompson 1987) that there is an "otherwise" rhetorical relation signalled by the use of "otherwise", it still goes only a small way towards characterizing what is happening.

- Finally, we alluded earlier to ways in which the themes of subsequent utterances may be related and how "otherwise" was a prime example of a contrastive relation between themes or between theme and previous rheme. Discovering and enumerating these possibilities would do much to clarify the relationships between discourse structure and Information Structure.

Acknowledgements

I. Kruijff-Korbayová's work was funded by a Visiting Fellowship from the British Academy Postdoctoral Fellowship, November 1999 — April 2000, and a Postdoctoral Fellowship from the Royal Society / NATO, April 2000 — March 2001. We would like to thank Mark Steedman for advice and comments on earlier drafts.

Bibliography

Firbas, J. (1992). *Functional Sentence Perspective in Written and Spoken Communication.* Studies in English Language. Cambridge: Cambridge University Press.

Günthner, S. (1996). From subordination to coordination? verb-second position in German causal and concessive constructions. *Pragmatics 6*(3), 323–356.

Hajičová, E. and P. Sgall (1987). The ordering principle. *Journal of Pragmatics 11*(4), 435–454.

Halliday, M. A. (1970). *A Course in Spoken English: Intonation.* Oxford: Oxford Uniersity Press.

Halliday, M. A. (1985). *Introduction to Functional Grammar*. London, U.K.: Edward Arnold.

Karttunen, L. and S. Peters (1979). Conventional implicature. In C.-K. Oh and D. A. Dinneen (Eds.), *Syntax and Semantics: Presupposition,* Volume 11, pp. 1–56. Academic Press.

Krifka, M. (1993). Focus and presupposition in dynamic semantics. *Theoretical Linguistics 19*, 269–300.

Kruijff-Korbayová, I. (1998). *The Dynamic Potential of Topic and Focus: A Praguian Discourse Representation Theory*. unpublished Ph.D. thesis, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic.

Kruijff-Korbayová, I. and B. Webber (2000a, January 10-12). Concession, implicature, and alternative sets. In H. Bunt (Ed.), *Proceedings of the International Workshop on Computational Semantics IWCS-4*, Tilburg, the Netherlands, pp. 227–248.

Kruijff-Korbayová, I. and B. Webber (2000b, July). Discourse connectives, inference and information structure. In J. Bos and M. Kohlhase (Eds.), *Proceedings of the ICoS-2 Workshop on Inference in Computational Semantics*, Dagstuhl, Germany.

Mann, W. C. and S. A. Thompson (1987). Rhetorical structure theory: A theory of text organization. ISI Reprint Series ISI/RS-87-190, USC Information Sciences Institute, University of Southern Carolina.

Mathesius, V. (1975). On information bearing structure of the sentence. In S. Kuno (Ed.), *Harvard studies in syntax and semantics.* Harvard University Press.

Rooth, M. (1985). *A Theory of Focus Interpretation.* Ph.D. thesis, Graduate School of the University of Massachusetts, Amherst, Massachusetts.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics 1*, 75–116.

Sgall, P., E. Hajičová, and J. Panevová (1986). *The meaning of the sentence in its semantic and pragmatic aspects.* Dordrecht, The Netherlands: Reidel.

Steedman, M. (1996). *Surface Structure and Interpretation.* Cambridge, MA: M.I.T. Press.

Steedman, M. (2000a). Information structure and the syntax-phonology interface. *Linguistic Inquiry 31*(4), 649–689.

Steedman, M. (2000b). *The Syntactic Process*. Cambridge, MA: M.I.T. Press.

Vallduví, E. (1992). *The Informational Component.* New York: Garland.

Webber, B., A. Joshi, A. Knott, and M. Stone (2001). Anaphora and discourse semantics. Submitted to the Journal of Computational Linguistics.

Webber, B., A. Knott, M. Stone, and A. Joshi (1999). Discourse relations: A structural and presuppositional account using lexicalised TAG. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics*, College Park MD, pp. 41–48.

Connecting Information and Discourse Structure Levels through "Kontrast:" Evidence from Colloquial Russian Particles *–TO*, *ŽE*, and *VED'*

SVETLANA MCCOY
Boston University
smccoy@acs.bu.edu

ABSTRACT. The notion of kontrast, or the ability of certain linguistic expressions to generate a set of alternatives, originally proposed by Vallduví and Vilkuna (1998) as a clause-level concept, is re-analyzed here as connecting the level of information packaging in the clause and the level of discourse structure in the following way: kontrast is encoded at the clausal level but has repercussions for discourse structure. This claim is supported by evidence from the distribution properties of three colloquial Russian particles *–TO, ŽE,* and *VED'* which are analyzed as unambiguous markers of kontrast. Both the placement of these particles at the clausal level and their role in discourse are viewed as consequences of the type of the kontrast set and the cognitive status of information marked by each particle. [1]

1 Introduction

The notion of *kontrast,* introduced in Vallduví and Vilkuna (1998, V&V hereafter), establishes a conceptual distinction between two notions that have been conflated in the literature by the term of *focus*: one is *rheme,* a concept which, in opposition with *theme,* belongs to the domain of information packaging (Chafe 1976 inter alia) and the other is what the authors label as *kontrast,* the notion covering quantificational phenomena of a more formal semantic nature (Rooth 1985, 1992, Krifka 1991-92, etc.). V&V use the term *kontrast*

> as a cover term for several operator-like interpretations of focus that one finds in the literature: identificational foci, exhaustiveness foci, contrastive foci, contrastive topics, and also interrogative wh-words,

[1]This paper is based on chapter 8 of my Ph. D. dissertation. I would like to express gratitude to my committee members for all the help I received from them: Mary Catherine O'Connor, Bruce Fraser, Enric Vallduví, Catherine V. Chvany, and Shanley Allen. Special thanks go to my first reader Paul Hagstrom.

Paper presented at the ESSLLI 2001 Workshop on
Information Structure, Discourse Structure and Discourse Semantics
Edited by Ivana Kruijff-Korbayová and Mark Steedman.

> which have been seen as paradigmatic "foci" by many semanticists and syntacticians...
>
> The basic idea behind the notion of kontrast is the following: if an expression **a** is kontrastive, a **membership set** M={...,**a**,...} is generated and becomes available to semantic computation as some sort of quantificational domain. (V&V: 83)

V&V provide a convincing set of data from Finnish, Hungarian, and Catalan (with references to English and other languages) to demonstrate that kontrastiveness and rhematicity must be treated separately. They assign formal features [K:+] for kontrastive elements and [K:-] for non-kontrastive ones; similarly, [Rh:+] for rhemes and [Rh:-] for themes. Thus, there are four possible combinations of these features for any given element:

1. [K:+; Rh:+] – this combination has been discussed in literature under the label of *contrastive focus* (also, *identificational focus* and *exhaustiveness focus*);[2]
2. [K:+; Rh:-] – this combination has been known as *contrastive topics*;[3]
3. [K:-; Rh:+] – regular, non-kontrastive rhemes, or *foci*;
4. [K:-; Rh:-] – regular, non-kontrastive themes, or *topics.*

As pointed out by V&V, in different languages *kontrast* is expressed by different linguistic resources. For example, in English *kontrast* is signaled mainly by prosody while in Finnish it is encoded by syntactic position. Moreover, there is a conflict between a limited set of structural resources in a language and a set of interpretive categories that need to be expressed.

The goal of this paper is to demonstrate that the use of the notion of *kontrast* can be extended from a clause-level notion to one connecting the clause-level structure and the structure of discourse. Even though this set-generating power is encoded on the clausal level, it is one of the linguistic means that holds the discourse together and reflects the speaker's assumptions about the hearer's state of knowledge and attention.

The linguistic evidence for this claim comes from discourse particles of colloquial Russian *–to, že,* and *ved'*, which are analyzed as lexemes whose primary function is to signal, or unambiguously mark, kontrast and which are labeled "kontrastive markers," or "k-markers" (McCoy 2001).

[2] As pointed out by V&V, there is an additional variable present in this combination, i.e., the scope: [K:+] has a scope and [Rh:+] has a scope, and they need not but may coincide. I will indicate the scope of the kontrastive element within the *rheme* by $[_{+K/Rheme} \ldots]$.

[3] V&V (107, fn. 4) make an important point that not all thematic constituents can be kontrastive: only *links* (i.e., pointers to a specific file card for the entry of *rheme*) can while *tails* (i.e., the remaining part of *theme*) cannot (see Vallduví 1992 for terminology). The scope of the [K:+] element within the *link* will be marked here by $[_{+K/Link} \ldots]$.

Until recently, only descriptive analyses were available for this group of particles and particles *–to, že,* and *ved'* were labeled "emphatic," "contrastive," "intensifying," "expressive," "emotional," "strengthening," "enunciative," etc., with further classification of context-dependent multiple meanings or functions for each particle (see Vasilyeva 1972, collection of articles in three volumes of *Les Particules Enonciatives en Russe Contemporain*, etc.) However, descriptive approaches to these particles have proved inadequate and a search for an underlying meaning of (some of) these particle has been initiated in the following works: Bitextin (1994), Parrott (1997), Bonnot and Kodzasov (1998), Feldman (forthcoming). However, these early studies in the "unifying" direction have their limitations: they are either based on a single framework and/or choose to deal with unifying a single aspect of the particle(s), such as discourse role or cognitive status of information marked.

In the previous analyses, the placement rules of these particles at the clausal level made reference to either the position of the particle with respect to a prosodically prominent element or to some information structure construct as 'theme-rheme', 'topic-focus', etc. However, no comprehensive semantic explanation has been offered for explaining what the prosodic marking is used for and what the relationship between prosody and information packaging at the clausal level is. Similarly, the discourse role of these particles has been analyzed as giving rise to multiple implicatures and inference patterns and thus contributing to the cohesion and coherence of discourse.

In this paper it will be shown that such properties of these particles as their position in the clause and their role in discourse are consequences of their two essential properties as *k-markers*:

1. the type of set it marks and the type of membership within this set;

2. the cognitive status of the referents marked by it.

Below, these 3 particles will be given a unifying analysis which integrates the following current frameworks: the theory of "kontrast" (V&V); cognitive statuses of referents in discourse (Yokoyama 1986, Gundel, Hedberg, and Zacharski 1993); information packaging on the clausal level (Vallduví 1992); and hierarchical organization of discourse (Büring 2000[4]).

The data come from a variety of colloquial Russian texts, such as Protassova's corpus of the CHILDES database (MacWhinney 2000) and Zemskaja and Kapanadze (1978). The structure of the paper is the following: in the next three sections, the kontrastive properties and their consequences will be discussed for each particle individually, with particle *–to* given a more detailed analysis than particles *že* and *ved'*. The final section summarizes the findings with respect to what the analysis of these three particles as k-markers reveals about the role of kontrast at the clausal and discourse levels.

[4]For similar proposals see Kanerva and Gabriele (1995), van Kuppevelt (1995, 1996a,b), and Roberts (1996).

2 K-Marker *–TO*

I will first discuss two essential properties of *–to*: the type of set and the cognitive status of information marked by it. Then I will demonstrate how these essential properties determine the position of *–to* in the clause and its role in discourse.

2.1 Type of Set Marked by *-TO*

Particle *–to* marks a set of sets of related propositions (equivalently, a set of questions) which is generated by introducing alternatives to a kontrastive element within the *link* and a kontrastive element within the *rheme*. Consider an example from CHILDES:

(1) [Varja and her mother are looking at a picture of a dog who put her paw on a bear. They first discuss the dog, then start talking about the bear. Varja gets distracted by taking a scoop into her hand. Mother says:]

U tebja-***TO*** sovok, a chto u medvedja v lape? *(CHILDES, séance 2)*
At you-***TO*** scoop but what at bear in paw

YOU(-***TO***) have a SCOOP, but what does the BEAR have in his paw?

The information structure of the (English version of the) utterance containing *–to* is shown in (2). Below it, in (3), a generalized structure of the utterance with *–to* is given:

(2) [$_{+K/Link}$ You] (-***TO***) have [$_{+K/Rheme}$ a scoop] . . .

(3) [$_{+K/Link}$ *A*]-***TO*** *has property* [$_{+K/Rheme}$ *x*]

The default case for (3) is when the kontrastive (element within the) link *A* belongs to a set of entities: M_1={A, B, C, . . . } and the kontrastive (element within the) rheme *x* belongs to a set of properties: M_2={x, y, z, . . . }.[5] Thus, the proposition **x(A)** marked with ***–TO*** makes the hearer generate a set M_3, shown in three alternative ways in (4/5/6). In (4), the set M_3 is represented as a set of sets of propositions, in (5) the same set is shown as a set of questions (following Hamblin 1958/1973 and Karttunen 1977), and in (6) the set M_3 marked by *–to* is shown graphically:

[5]However, a more marked (i.e., less frequent) scenario is also possible: i.e., when the kontrast set for links is a set of properties, while the set of alternatives to the rheme consists of entities, as in (i):

(i) [Varja can't stop running. Mother is asking: who is the one running?]

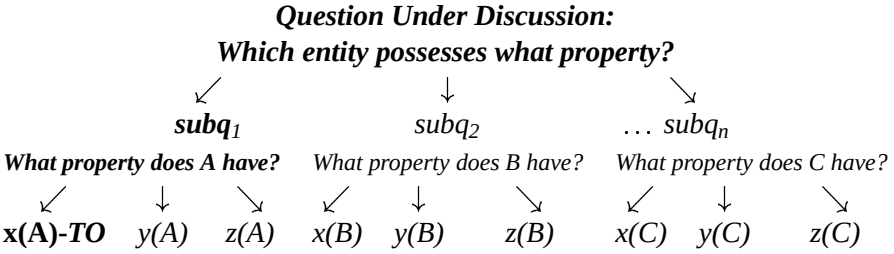
Nu [$_{+K/Rheme}$ kto] [$_{+K/Link}$ begaet]-***TO?*** *(CHILDES, séance 2)*
Well who is-running-***TO***

As for somebody running(-***TO***), who is doing this?

(4) M_3={ { **x(A)**, *y(A)*, *z(A)*, … };
{ *x(B)*, *y(B)*, *z(B)*, … };
{ *x(C)*, *y(C)*, *z(C)*, … }; … }

(5) M_3={**What is true of *A*?**; What is true of *B*?; What is true of *C*?; … }

(6) The set of sets of propositions (or, a set of questions) M_3 marked by ***–TO:***



So, the primary function of *–to* is a marker of a set of sets of propositions which is generated by introducing alternatives to the kontrastive link and the kontrastive rheme. While for the proposition containing *–to* the truth value is asserted, it is not the case with the alternative propositions: they are only made salient with the help of *–to*.

2.2 Cognitive Status of Referents Marked by *-TO*

Particle *-to* marks information (estimated by the speaker to be) known to the hearer but not activated in the hearer's mind at the time of the utterance.[6]

How can the speaker assume that the information is also known to the hearer? The speaker can safely assume that the hearer also knows the information if this information is shared through: code (language, culture), encyclopedic knowledge, deixis, or common experience (for more detail, see Yokoyama 1986). The example in (7) illustrates a speech situation where the source of the speaker's assumption about the proposition marked by *–to* being located in the hearer's knowledge set is deixis, while in (8) it comes from common experience:

(7) [A to B, after a long silence, on the top of a mountain before dawn]

Tišina-***TO*** kakaja! *(Vasilyeva 1972:68)*
Quietness-***TO*** *what*

'How quiet it is!' or:
'As for the state of quietness(-***TO***), how quiet it is!'
(Vasilyeva's translation: How wonderfully quiet!)

[6]Since all three particles are capable of marking information which is, in speaker's estimation, currently activated in the hearer's mind or inferable from information located there, I am primarily concerned with cognitive statuses that are located further away from the information which is at the center of the speaker's and the hearer's attention at the moment of the utterance.

(8) [Varja practices pulling laces through holes. After several other activities, Mother says:]

Oj, slushaj, a my s toboj [$_{+K/Link}$ chernen'kij**]-*TO*** shnurochek
Oh listen but we with you black-***TO*** *lace-DIM*
[$_{+K/Rheme}$ tak ni razu eshche i ne prodevali **]**, da?
so not once yet even NEG pulled-through, yes
(CHILDES, séance 5)

'Oh, listen, as for the BLACK(-***TO***) lace, you and I haven't pulled it through yet even once, right?'

To summarize the essential properties of *–to* as a k-marker: with respect to the type of set marked, *-to* marks a set of sets of propositions which differ from each other in the values of the kontrastive link and the kontrastive rheme. With respect to cognitive status of information marked by *–to*, this information is assumed by the speaker to be known to the hearer but not activated in the hearer's mind at the time of the utterance. In the next subsection I will demonstrate that these two essential properties of *-to* determine its position in the clause and its role in discourse.

2.3 Consequences: Position in the Clause and Role in Discourse

The position of *–to* in the clause is best analyzed with respect to the element marked [+kontrast]. Thus, *-to* is an enclitic to the kontrastive element within the *link.* This is a direct consequence of its kontrastive properties: first, the kontrast on the *link* value results in a set of sets of propositions (cf. similar treatment of "contrastive topics" in Büring 2000, etc.); second, the cognitive status of information marked by *–to* — known to the hearer but not activated in the hearer's mind — is more compatible with the function of the *link* than the *rheme*.

The role of *–to* at the discourse level is also a consequence of being a marker of a set of questions (equivalently, a set of sets of propositions).[7] In a discourse tree, as in (9), any question under discussion (QUD) can be constrained by k-marker *-to* in the way as shown in (6):

(9) **Discourse Situation**

↙ ↓ ↘
QUD$_1$ ***QUD***$_2$ ***QUD***$_n$

[7] Another consequence of the kontrastive nature of *–to* is its multiple implicatures proposed in the literature. Since *–to* is a marker of a set of sets of propositions, it also marks contrast and emphasis, functions as a theme/topic marker or a marker of contrastive topics, implies plurality, etc. The following implicatures are consequences of its second essential property—a marker of information (assumed by speaker to be) known to hearer but not activated in discourse yet: a marker of unexpectedness in addressing a topic; the fact that it adds a tone of familiarity, conversational spontaneity, intimacy, etc.

To summarize, the notion of kontrast is the core semantic meaning of particle *–to*: by analyzing it as a marker of a set of sets of propositions (equivalently, a set of questions) and as marking referents known to the hearer but not currently activated, its position in the clause and its role in discourse are accounted for.

3 K-Marker *ŽE*

K-marker *že* differs from *–to* in both the type of set and cognitive status of information marked and, accordingly, its other important properties (position in the clause, role in discourse, etc.) are also different from those exhibited by *–to*. However, as with *–to*, kontrastiveness is the core semantic meaning of this particle and determines its distribution.

3.1 Type of Set Marked by *ŽE*

Particle *že* marks a set of propositions which differ from each other in the value of (at least) one term. The kontrast set for *že* contains members which are mutually exclusive: if one proposition is true, the other one(s) is/are false. Thus, the relationship among the members of the *že* set often involves (binary) opposition, contradiction, or negation. The kontrast set marked by *že* is generalized in (10) and illustrated with a sample of naturally occurring data in (11):

(10) **The kontrast set marked by *ŽE*:**
M={ *X*, *X'* }, where *X*=¬*X'*
(*X* is true if and only if *X'* is false)

(11) [Varja notices a fly on the windowsill and asks her mother to kill it]
VAR: Ona muxa, muxa.
MOT: Muxa, muxa, da.
VAR: Ubit', ubit' ee!
MOT: Ona zhe uzhe ubita. *(CHILDES, séance 4)*

VAR: It's a fly, a fly.
MOT: A fly, a fly, yes.
VAR: Kill, kill it!
MOT: It (*že*) is already killed.

Gloss:

Ona zhe **[**$_{+\mathbf{K}}$ uzhe ubita**]**.
she ***že*** already killed (participle).

'(But) it (***že***) is already killed.'

In (11), mother's utterance with *že* corrects the presupposition of Varja's previous utterance *Kill it!*, i.e., *The fly is alive.* The members of the kontrast set marked by *že* are mutually exclusive: it is impossible for the fly to be alive (presupposition

of Varja's utterance) and to be already killed (mother's utterance containing *že*) at the same time.[8]

3.2 Cognitive Status of Referents Marked by *ŽE*

Particle *že* marks the membership set, one member of which is activated in the hearer's mind at the time of the utterance, while the other (incompatible) member is viewed by the speaker as though it is (or should have been) known to the hearer and should have been activated at this time.

For example, in (11), the activated member of the set is *The fly is alive* (presupposition of *Kill, kill it!*). The other member of the set (is treated by the speaker as though it) should be known to the hearer and, therefore, should be activated — i.e., in the speaker's estimation, there are enough visual cues for the hearer to make the correct conclusion (*The fly is already dead*).

Now let us consider some consequences of the essential properties of *že* as a k-marker.

3.3 Consequences: Position in the Clause and Role in Discourse

The placement of particle *že* can be defined with respect to the kontrastive element, which is a propositional term (usually within the *rheme*). The condition of mutual exclusiveness (or binary opposition) on the set marked by *že* makes this particle a good candidate of marking rhematic kontrasts, especially cases of kontrastive

[8] An interesting case is the use of *že* with wh-words. This case is somewhat cumbersome for the present analysis: how in the world can the kontrast set be mutually exclusive with wh-words? It seems that there is a way out of this paradox: *že* with wh-words is used to indicate that from the speaker's point of view, the set of possible answers is empty of reasonable alternatives (and is now open to unreasonable alternatives). While the set of possible answers is empty of reasonable alternatives for the speaker, it might not be so for the hearer: the speaker expects the hearer to provide an answer (which is probably not so unreasonable from the hearer's point of view). An example, recorded in a home setting, is provided in (i):

(i) [Father puts away his son's library books and tells him that mother took them back to the library (which she did not do). The boy asks mother if she indeed took them back. She says 'no'. He then says (in the presence of both parents):]

U kogo ***že*** mne togda sprosit'?
'Who (***že***) (in the world/the hell . . .) should I ask then?'

In (i), from the speaker's point of view, the set of reasonable answers (individuals who can provide an answer) is exhausted or empty; the child now throws the ball to the parents who might be able to provide an alternative explanation, which at this point in discourse is viewed by the child as an unexpected/non-predictable/unreasonable alternative. However, the answer to the child's question is indeed known to the parent(s) who opted not to reveal it to him. It seems that an explanation along these line would bring us as close to the mutually exclusive set as we could possibly get with wh-words.

verum foci[9] — i.e., kontrast not on the lexical meaning of the verb but on some verbal inflectional category, such as tense, aspect, etc. It is possible to translate utterances containing *že* as cases of rhematic polarity of verum focus; for example, alternative translations for the utterance with *že* in (11) are *It [=the fly] IS already killed* or *It is ALREADY killed.*

The role of *že* in discourse is also a consequence of its kontrastive properties: *že* refers back to a salient element or some unresolved (from the speaker's viewpoint) question in the discourse or discourse situation.[10]

To summarize, by analyzing particle *že* as a marker of a set of mutually exclusive propositions and as a marker of information which is related through the set relationship to information that is activated in discourse, its distributional properties at the clausal and discourse levels are largely accounted for.

4 K-Marker *VED'*

Etymologically, particle *ved'* is a form of the verb *vedat'* 'to know', from which it has inherited factivity.

4.1 Type of Set Marked by VED'

Similarly to *že*, particle *ved'* marks a set of propositions which differ from each other in the value of at least one term. However, there are important differences between them: for *ved'* the set membership is restricted to propositions which have illocutionary force of assertions (while *že* does not have this restriction). Also, the members of the set are not mutually exclusive (as in the case of *že*) but rather supplementary; when *ved'* does express contradiction, it is more of the '*yes, but* ... ' type.

The examples in (12-13) illustrate that the difference between *že* and *ved'* comes (partially) from the difference between a mutually exclusive set (*že*) and a supplementary set (*ved'*):

(12) [Varja and Grandmother are looking at a picture of birds standing on the ground. Varja believes that the birds have fallen down. Grandmother corrects Varja: since they are standing and not lying down, they have not fallen down:]
*VAR: Èta pit'ki.

[9]Thanks to Enric Vallduví for pointing this out.

[10]Multiple implicatures associated with *že* at the discourse level can also be viewed as following from its two essential properties. Since *že* is a marker of a restricted set, containing mutually exclusive propositions, it is subsequently a marker of contrast, a marker of emphasis, a marker of contrastive focus, a marker of contrastive topic, it adds a tone of an indisputable argument and is perceived as a verbal attack on hearer, etc. From being a marker of activated information which is to a large extent known to the hearer, it follows that *že* can also be analyzed as a marker of a reference point in the activated domain of reference, as being perceived to be a verbal attack on hearer, as adding a tone of an indisputable and irrefutable argument, and conveying emotions of irritation, annoyance, and impatience, etc.

%eng: These are birds.
*VAR: Upai.
%eng: Fell down.
*GPP: Oni ne upali.
%eng: They did not fall down.
*GPP: Pochemu ***zhe*** ty dumaesh', chto oni upali.
%eng: Why do you think, that they fell down.
*GPP: Odna stoit na penechke, drugaja stoit na zemle, a tret'ja stoit na zemle szadi penechka.
%eng: One is standing on a little stump, the second is standing on the ground, and the third is standing on the ground behind the stump.
*GPP: Vot tak, szadi penechka.
%eng: That way, behind the stump.
*GPP: A pochemu ty dumaesh', chto oni upali, oni ***zh*** ne lezhat, oni stojat.
%eng: And why do you think, that they fell down, they are not lying, but standing.
*GPP: Kogda kto-nibud' upadet, on lezhit.
%eng: When somebody falls, he lies down. *(CHILDES, séance 3)*

GPP: Pochemu ***zhe*** ty dumaesh', chto oni upali.
'Why (***že***) in the world do you think that they fell down.'

GPP: A pochemu ty dumaesh', chto oni upali, oni ***zh*** ne lezhat, oni stojat.
'And why do you think that they fell down, they (***že***) are not lying down, they are standing.'

In (12), *že* marks a set containing mutually exclusive members (thus it conveys emotions of contradiction, correction, and is perceived as a 'verbal attack on the hearer'): Varja believes that *The birds fell down* (since they are standing on the ground and not flying), while the Grandmother refutes that by saying that *The birds did not fall down* (since they are standing and not lying down). Later, this argument is brought to a mediator, Varja's mother (Katja). In presenting the gist of the argument to Katja, the Grandmother uses *ved'* and not *že* (since no opposition is expected from an adult who knows the lexical meaning of *upali* 'fell down'):

(13) [Grandmother re-tells the argument about the "fallen" birds to the mother]

Vot ona govorit, chto èti ptichki upali, a ja schitaju, chto
PART she says, that these birds-DIM fell-down, and I think that
oni **[**$_{+\mathbf{K}}$ ne upali **]**, oni **[**$_{+\mathbf{K}}$ stojat ***ved'*** na nozhkax **].**
they NEG fell-down, they are-standing ***ved'*** on feet-DIM
(CHILDES, séance 3)

'So she says that these little birds fell down and I think that they did not fall down, they are standing, (*ved'*) you know, on their feet.'

So, in (13) *ved'* marks a supplementary set, the members of which are synonyms rather than antonyms (thus, it constitutes an 'extension in logic' and not 'a verbal attack'): {*The birds did not fall down; The birds are standing on their feet*}.

4.2 Cognitive Status of Referents Marked by *VED'*

Similarly to *–to*, particle *ved'* marks information known to the hearer but not activated in the hearer's mind at the time of the utterance. The difference between *že* and *ved'* comes (partially) from the difference in the cognitive status of information marked. Consider example (14) from Zemskaja and Kapanadze (1978: 52-53; glosses and translation added):

(14) [Two retired ballet dancers had previously talked about what great actors they used to work with in their youth. Later, one of them realizes that they forgot to mention two of the most famous actors (Čirkov and Čerkasov) and reminds her friend about that]

A: Kogda my s toboj govorili o mjuzik-holle, my govorili: ax kakie byli aktery!
'When you and I talked about the music hall, we were saying "Oh what actors were there!"'

My s toboj ***ved'*** ne skazali, čto tam byl Čirkov, tam byl Čerkasov!...
We with you ***ved'*** NEG said that there was Čirkov, there was Čerkasov
'You know, we did not say that there was Čirkov there, there was Čerkasov there!..'

B: [overlapping with the end of the preceding utterance]
Da gospodi! Čerka-a-sov, Čirko-o-v...
'Oh, my god! Čerka-a-sov, Čirko-o-v...'

In (14), *ved'* marks information known to both the speaker and the hearer which is not activated in the hearer's mind at the moment: talking about what great actors were in the theater at the time when both of the interlocutors worked there, they forgot to mention two most famous actors.

4.3 Consequences: Position in the Clause and Role in Discourse

Similarly to the other two particles, the position of *ved'* in the clause is best defined with respect to the kontastive element: it is either a proclitic or an enclitic to such element.

The discourse role of *ved'* is determined by the type of set it marks: being a marker of a supplementary set, it relates members of the set to each other by building a super-question above them.[11]

[11] Also, multiple implicatures of *ved'* can be treated as consequences of its kontrastiveness: since *ved'* is a marker of a (supplementary) set of propositions (assertions), it is a marker of contrast, a marker of emphasis, an assertion marker, etc. Since *ved'* (like *–to*) is a marker of information that is

	-TO	***ŽE***	***VED'***
ESSENTIAL PROPERTIES:			
Type of set marked	A set of sets of propositions (= a set of questions); requires two [+K] elements: in the *link* and the *rheme*	A set of mutually exclusive propositions which differ in the value of at least one [+K] element	A supplementary set of propositions which differ in the value of at least one [+K] element
Cognitive status of referents	*Familiar* to the hearer but *not activated* in his/her mind	One member: *in focus* or *activated*; the other member: (treated as) *familiar* and *activated*	Same as for ***-TO***
CONSEQUENCES:			
Position in a clause	Enclitic to [+K] element, usually *link*	Enclitic to [+K] element	Proclitic or enclitic to [+K] element
Role in discourse	Generates a set of sub-questions (sisters dominated by the same QUD); partially answers QUD and makes salient the other sub-questions.	Refers back to salient element or some unresolved question in the previous discourse	Generates a set by building a super-question above the members of the set

Table 11.1: Comparing Kontrastive Markers *–TO, ŽE*, and *VED'*

To summarize: *ved'* has been observed to share certain properties with *že* and other properties with *–to*. By addressing the questions of what type of set and what cognitive status of information each of these particles marks, their differences and similarities are explained.

5 Conclusion

The notion of kontrast has been found to be a core semantic meaning of such lexical items as particles *–to, že*, and *ved'*: these lexemes are analyzed as unambiguous markers of kontrast. The core semantic meaning of each of these particles has been found to encompass two issues: the type of set and the cognitive status of referents marked by each of these particles. Among important properties of these particles, which are nevertheless treated as consequences of their essential kontrastive properties, are the placement of the particles at the clausal level and their role in the organization of discourse. Their distribution properties and the choice of particles in a particular context is also motivated by the differences in their kontrastive

assumed by speaker to be known to hearer but not activated yet, it is a marker of unexpectedness in addressing a topic, a marker of encyclopedic knowledge, and it is perceived as a (polite) reminder, etc.

nature. The findings are summarized in Table 4.3.

The example of these three colloquial Russian particles suggests that the notion of kontrast has applications beyond the level of the clause to serve as a construct which connects the levels of clausal and discourse structure.

Data Sources

CHILDES: Other Languages - Russian – Protassova. In: MacWhinney, Brian. 2000. *The CHILDES Database: Tools for Analyzing Talk.* 3d Edition. Vol. 2: *The Database.* Mahwah, NJ: Lawrence Erlbaum Associates.

Zemskaja, E. A. and L. A. Kapanadze, eds. 1978. *Russkaja Razgovornaja Reč': Teksty.* Moskva: Nauka.

Bibliography

Bitextin, Aleksandr Borisovič. 1994. Časticy *-to, že, ved'* i Vvodnye Konstrukcii Tipa *kak izvestno* kak Sredstva Ukazanija na Izvestnost' Propozicional'nogo Soderžanija Predloženija Slušajuščemu. AKD, Moskovskij Gosudarstvennyj Universitet.

Bonnot, Christine and Kodzasov, C. B. 1998. Semantičeskoe Var'irovanie Diskursivnyx Slov i Ego Vlijanie na Linearizaciju i Intonirovanie (Na Primere Častic *ŽE* i *VED'*). In: Kiseleva, K. and Paillard, D., eds. *Diskursivnye Slova Russkogo Jazyka: Opyt Kontekstno-Semantičeskogo Opisanija.* Moskva: Metatekst.

Büring, Daniel. 2000. On D-Trees, Beans, and B-Accents. Ms., UCLA.

Chafe, Wallace L. 1976. Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View. In: Li, Charles N., ed. *Subject and Topic.* New York: Academic Press, 25-55.

Feldman, Anna. Forthcoming. Discourse Markers - Accessing 'Hearer-Old' Information: The Case of the Russian *Že*. In: Melby, Alan et al., eds. *Proceedings of the 27th LACUS Forum.*

Gundel, Jeanette K., Nancy Hedberg, and Ron Zacharski. 1993. Cognitive Status and the Form of Referring Expressions in Discourse. *Language* 69 (2): 274-307.

Hamblin, C. L. 1958. Questions. *Australasian Journal of Philosophy* 36(3): 159-168.

Hamblin, C. L. 1973. Questions in Montague English. *Foundations of Language* 10: 41-53.

Kanerva, Jonni M. and Leslie A. Gabriele. 1995. Intonation and Focus Layers. In: Beckman, Jill N., ed. *Proceedings of the Northeast Linguistic Society 25 (NELS 25).* Vol. 1: Papers from the Main Session. University of Pennsylvania, 335-346.

Karttunen, Lauri. 1977. Syntax and Semantics of Questions. *Linguistics and Philosophy* 1: 3-44.

Krifka, Manfred. 1991-2. A Compositional Semantics for Multiple Focus Constructions. *Linguistische Berichte,* Suppl. 4: 17-53.

McCoy, Svetlana. 2001. Colloquial Russian Particles *–To, Že*, And *Ved'* As Set-Generating ("Kontrastive") Markers: A Unifying Analysis. Ph.D. dissertation, Boston University.

Les Particules Enonciatives en Russe Contemporain 1 (1986), 2 (1987), 3 (1988). Paris: Institut d'Études Slaves.

Parrott, Lillian. 1997. Discourse Organization and Inference: The Usage of the Russian Particles *že* and *ved'*. Ph. D. dissertation, Harvard University.

Roberts, Craige. 1996. Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics. In: Yoon, J. H. and Andreas Kathol, eds. *OSU Working Papers in Linguistics 49: Papers in Semantics,* 91-136.

Rooth, Mats. 1985. Association with Focus. Ph.D. dissertation, University of Massachusetts, Amherst, MA.

Rooth, Mats. 1992. A Theory of Focus Interpretation. *Natural Language Semantics* 1: 75-116.

Vallduví, Enric. 1992. *The Informational Component.* New York: Garland.

Vallduví, Enric, and Maria Vilkuna. 1998. On Rheme and Kontrast. *Syntax and Semantics* 29 (The Limits of Syntax), 79-108.

van Kuppevelt, Jan. 1995. Discourse Structure, Topicality and Questioning. *Linguistics* 31: 109-147.

van Kuppevelt, Jan. 1996a. Directionality in Discourse: Prominence Differences in Subordination Relations. *Journal of Semantics* 13: 363-395.

van Kuppevelt, Jan. 1996b. Inferring from Topics. *Linguistics & Philosophy* 19: 393-443.

Vasilyeva, A. N. 1972. *Particles in Colloquial Russian: Manual for English-Speaking Students of Russian.* Moscow: Progress Publishers.

Yokoyama, Olga Tsuneko. 1986. *Discourse and Word Order.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Contrast from a Contrastive Perspective

VALÉRIA MOLNÁR
Lund University, Sweden
valeria.molnar@tyska.lu.se

ABSTRACT. The main concern of the paper is to address some problems that arise when defining the concept of *contrast* and to clarify the relation between contrast and other concepts of information structure. As to the definition of the notion of contrast, we need to separate the semantic and pragmatic aspects of contrastiveness as well as identify different types within these two main categories. There is abundant cross-linguistic evidence for the hypothesis that certain types of contrast motivate a more fine-grained analysis of topicality and focusing. It will be argued, however, that contrastiveness not only is a relevant feature for the parametrization of topicality and focusing, but that contrast indeed represents an autonomous concept of information structuring.

1 Introduction

The main concern of the paper is to clarify the relation between *contrast* and the two main concepts of information structure *topic* and *focus*. This requires a *definition* of contrast and the specification of the *compatibility* of different contrast types with focusing and topicality.

I will argue for the following assumptions:

(i) the notion of contrast is a **linguistically relevant phenomenon** and does not only arise "from particular inferences which we draw on the basis of given conversational contexts" (Lambrecht 1994: 290): the feature contrastiveness has *syntactic* and *phonological* consequences for the realization of topic and focus;

(ii) the **distinction of different types of contrast** is necessary in order to explain its impact on linguistic forms, and above all the separation of the *pragmatic level* of contrastiveness from the *semantic*, quantificationally based, level is required;

(iii) the **linguistic status of contrast** must be differentiated: contrast is not simply a feature of topicality and focusing, but it seems plausible to regard contrast –in certain languages and in certain constructions– as a further *autonomous "packaging phenomenon"*. On the basis of cross-linguistic evidence contrast must be included into the inventory of pragmatic categories.

Paper presented at the ESSLLI 2001 Workshop on
Information Structure, Discourse Structure and Discourse Semantics
Edited by Ivana Kruijff-Korbayová and Mark Steedman.

I would like to start my discussion with some introductory remarks on the basic notions of contrast, topic and focus and the relations between them.

According to a theory neutral definition (cf. Bußmann 1990: 419), the notion of **contrast** has two main dimensions: firstly, it is used as a synonym for "opposition" – either on the paradigmatic or syntagmatic level – and secondly it also includes also another aspect, namely highlighting by accent.

The notion of **focus** is related to highlighting in some sense, too. The view that focus is the "information centre of the sentence" and contains the new, non-presupposed part of the utterance is widely accepted in the literature. It is foregrounded most often by stress while the rest of the utterance remains in the background. Besides this syntagmatic (horizontal) type of highlighting, it is claimed that highlighting in the paradigmatic (vertical) dimension may also be involved. The basic idea is that a set of alternatives exists for the focused constituent which stands in opposition to all of them. Obviously, the two main properties of contrast - opposition in two directions and highlighting - are also typical of the notion of focus. **Contrast** and **focus** are thus often regarded as very closely related concepts in linguistic research.

According to an extreme view, focusing is always contrastive - and as all utterances contain a focus, all utterances must necessarily be contrastive - or as Dretske (1952) puts it, "all contingent statements contrast […] one state of affairs with another". While Bolinger (1961: 87) defends the view that "in a broad sense every semantic peak is contrastive", Dretske (1972: 412) argues that "contrastive statements" must be regarded as a special class, because they "embody a dominant contrast, a contrastive focus, a featured exclusion of certain possibilities". In examples (1) and (2) from Dretske, different contrasts are featured: in (1) the verb *sold*, but in (2) the nominal phrase *my typewriter* stands out as the focal point and embodies the dominant contrast:

(1) I *sold* my typewriter to Clyde.

(2) I sold *my typewriter* to Clyde.

The decisison as to whether or not focusing is inherently contrastive depends on how important we judge dominant contrast to be. If not only highlighting, but further conditions such as "a featured exclusion of certain possibilities" need to be met for contrast, then contrast can only be an optional and not an obligatory property of focusing.

As to **topic**, three definitions dominate the linguistic landscape. Firstly those which define topic as the notion of *aboutness* (Reinhart 1982) or as an "address pointer", also called "link" by Vallduví (1992). According to another influential view, the topic should be regarded as the notion of *frame*: "the topic sets a spatial, temporal, or individual framework within which the main predication holds" (Chafe 1976: 51). A third definition of topic is given with recourse to *old information*: the topic is either identified as given information or in weaker versions, the "givenness condition" is an important part of topicality.

In none of these three definitions do the properties of topicality have a direct relation to the notion of contrast. The separation of the topic from the rest of the sentence can be made in an explicit, emphasized way, but this is not necessarily so. In the analysis of the **relation of contrast to topic**, their compatibility is a controversial issue, especially in theories where emphasis only can be assigned to focus.

2 The definition of the notion contrast

The most important criteria that are discussed in the literature in connection with the definition of contrast are listed below with regard to their hierarchy (in order of importance):

- highlighting
- dominant contrast
- membership in a set
- limited set of candidates
- explicit mentioning of alternatives

The relevance of the criteria is judged in different ways - leading to different types of distinctions between the contrastive and non-contrastive, regular foci. It is important to note that the judgements are partly dependent on the theoretical approach, partly on the type or number of languages taken into consideration and the presence of obligatory formal marking of the distinction in these languages.

The basic requirement for contrastiveness is the existence of highlighting. Not only Bolinger links every semantic peak to contrastiveness, but also Lambrecht (1994) regards highlighting as a sufficient condition of contrast. According to the overwhelming majority of linguists, though, the existence of a "dominant contrast" dividing the utterance into two parts - background (also called "presupposition") and focus - is a necessary requirement of contrast. Rochemont (1986: 52) distinguishes contrastive focus from presentational focus on the basis of this criterion.

A further condition of contrast is membership in a set, namely that we can generate a set of alternatives for the focused constituent - or as Jackendoff (1972: 243) puts it "a coherent class of possible contrasts with the focus." This condition is regarded as a sufficient prerequisite of contrast also by linguists working in the framework of "Alternative semantics" (e.g. Rooth 1985). In many approaches, however, "membership in a set" is narrowed down to the requirement of a *closed* set. The decisive criterion for contrastiveness is thus, according to Halliday (1967), Chafe (1976) and Rooth (1992), the availability of a limited number of candidates. Halliday (1967: 206) defines "contrastive" "as contrary to some predicted or stated alternative" and Chafe (1976: 34) also favours the view – opposed

to Bolinger's suggestion – that "contrastive sentences are qualitatively different from those which simply supply new information from an unlimited set of possibilities". When the set of possibilities is unlimited, the sentence supplies only "new information" and fails to be contrastive.

While Chafe regards "the limited number of candidates" as the essential distinction between contrastive and non-contrastive readings, Jacobs (1988: 113) claims that the candidates *excluded* must be explicitly mentioned in the context. According to Jacobs, the focus of negation is thus inherently contrastive (3), while the contrastiveness of foci of the illocutionary operators is context-dependent, only present in case of explicit mentioning of alternatives (4):

(3) Dieses Buch hat mir nicht$_2$ Kláus empfohlen, sondern Gérda.
F_2

(4) A: Ich fürchte, daß wieder die Schwéden gewonnen haben.
B: Keine Angst! Diesmal haben wír gewonnen.
F_1

(5) Wießt du, wer gewonnen hat? Wír (haben gewonnen)!
F_1

There are further controversial questions in connection with the definition of contrast, namely whether it should be regarded as a gradient or distinct notion and whether or not contrast is an independent phenomenon of information structure.

If one shares Bolinger's and Lambrecht's opinion, according to which focusing is always contrastive, then utterances can only be *used* contrastively and it is only possible to speak of "clear or less clear instances of contrastiveness". Even for this gradient approach the "clearest instances of contrastiveness" are those "in which a focus designatum explicitly contradicts a stated or predicted alternative" (Lambrecht 1994: 290). Most approaches argue, however, for the non-gradient character of contrastiveness and for a distinction between contrastive and non-contrastive cases on the basis of one of the above-mentioned criteria. A relevant task for linguistic theory is to take a stand in this controversy, by expanding the data to a large number of languages so that cross-linguistically valid arguments can be made.

The second important theoretical question is the linguistic status of contrast - based on the impact of contrast on linguistic realization. The standard view represented by linguists working within different theoretical frameworks (see above Dretske, Bolinger, Halliday, Rochemont, Rooth, Jacobs etc) is that contrast cannot be regarded as an independent phenomenon of information structure, but only as a feature of focusing and topicality, which serves to further parametrize these notions. Analysing contrast as a "packaging phenomenon" was, however, suggested by Chafe in 1976 and also by recent linguistic analyses of Finnish data.

3 The linguistic relevance of contrast - *within* focusing and topicality

In opposition to the view according to which contrastive cases cannot clearly be separated from non-contrastive ones, there is abundant evidence (i) that contrast in many languages can be optionally marked or even *must* be marked by grammatical means, and (ii) that certain formal phenomena (syntactic, morphological and phonological) can only be explained with recourse to the notion of contrast. Consequently, contrast should be a linguistically relevant phenomenon, and not only a cognitively motivated category.

3.1 Contrast vs. Focus

Contrast – or the presence of certain kinds of contrastiveness – motivates the internal differentiation of focusing. It is plausible to assume that the notion of contrast applies only to those foci which operate on predicted or stated alternatives and do not simply express new information. A further differentiation of contrastive foci seems to be motivated by the status of alternatives: explicitly mentioned alternatives trigger other syntactic patterns than simply predicted alternatives.

A distinction between two main types of focusing is suggested on the basis of empirical data from Hungarian by É. Kiss (1998). The **dichotomy** "identificational focus" vs. "information focus" is based on differences both in syntactic realization and semantically relevant content. The most important *syntactic* difference is that the information focus remains in situ (6) while the operator ("identificational") type must be moved to a special operator position, in Hungarian e.g. into the preverbal position (7) (cf. É. Kiss 1998: 249):

(6) Mari ki nézett magának EGY KALAPOT.
'Mary picked for herself A HAT.'

(7) Mari **egy kalapot** nézett ki magának.
Mary a hat.ACC picked out herself.DAT
'It was a **hat** that Mary picked for herself.'

The *semantic* difference can be specified as follows: while the focus operator "represents a subset of the set of contextually or situationally given elements for which the predicate phrase can potentially hold" (É. Kiss 1998: 245), the information focus only conveys new, nonpresupposed information. The sentence (7) containing a focus operator expresses that it was *only a hat and nothing else* Mary picked out for herself, whereas the information focus in (6) merely presents *a hat* as new information, without suggesting that the hat was the only one of a set of relevant things for Mary.

From a cross-linguistic perspective, the different types of contrastiveness demand, however, a further division of the focus field - not only a dichotomy, but even a **trichotomy** seems plausible. This is the result of a necessary distinction

of two types of contrastive foci, also represented syntactically: The focus operator operating either on a contextually *open* or *closed* set requires movement to a verb-related Focus-position. Movement into the verb-adjacent Focus position - which according to proposals made within generative frameworks is located in the left periphery of the sentence, i.e. in the Focus projection of the CP-domain (see below 5.2.) - is overt in so-called Focus languages (e.g. Hungarian, Basque), but can also be covert and delayed until LF (as in English). On the other hand, for the focus type by which alternatives are excluded from a *closed* set, movement is not necessarily verb-related, but in this case the left-peripheral position is decisive. Contrastive foci are moved in syntax into a projection appearing sentence-initially in surface structure in a great number of languages (Italian, Spanish, Catalan, Rumanian, Arabic, Greek, Russian, Finnish etc).

The difference between the two types of Focus movement can be illustrated by comparing Finnish and Hungarian data: in Finnish the leftmost position of the sentence is responsible for the expression of contrastiveness (8-B) and can be occupied only when the focused constituent refers to alternatives in a contextually limited set where the alternatives are known to the participants of the discourse. This explains why (9-A) is not an appropriate answer to a general informative question (9-F1) or to a constituent question (9-F2) if the alternatives are not explicitly mentioned or at least present in the minds of the discourse participants (cf. Järventausta & Molnár, 2001):

(8) A: Pekka lensi Tukholmaan.
'Pekka flew to Stockholm.'

B: [KONTRAST **Reykjavikiin**] Pekka lensi .
'To Reykjavik, Pekka flew.'

(9) F1: What did Pekka do?

F2: Where did Pekka fly?

A: * [KONTRAST **Reykjavikiin**] Pekka lensi.
Reykjavik-to Pekka flew

In Hungarian the opposite is true: the focus operator must be adjacent to the finite verb in surface structure (10-B1), but is not necessarily related to contextually specified alternatives:

(10) A: Hova repült Péter?
Where flew Peter
'Where did Peter fly?'

B1: Pter [FOCUS **Reykjavikba**] repült.
Peter to Reykjavik flew
'Peter flew to Reykjavik.'

B2: * [FOCUS **Reykjavikba**] Péter replt.

Basque yields also strong evidence for the assumption of two contrast-related focus positions. In Basque, both types of contrast-related focusing are possible: Etxepare (1997) distinguishes between "emphatic focus" (11-a), the verb-related focus (here incidentally in sentence-initial position, but as (11-b) shows its verb-adjacent position is essential for grammaticality), and "contrastive focus" on the left periphery of the sentence (12-b) excluding a contextually explicit alternative (12-a) - however without the restriction on verb-adjacency:

(11) a. PATATAK **maite ditu** Jonek.
potatoes love AUX Jon
'It is potatoes that Jon loves.'

b. * PATATAK Jonek **maite ditu**.

(12) a. MIKELI, ardoa **ekarri diote**.
for-Mikel wine bring AUX
'It is for Mikel that they brought wine.'

b. JONEK, ardoa **ekarri du**.
Jon wine bring AUX
'It is Jon that brought the wine.'

3.2 Contrast vs. Topic

Contrastiveness is not only marked in the field of focusing, but also in combination with topicality, in the latter case often by the use of special **syntactic** patterns such as topicalization and left-dislocation. Left-dislocation and topicalization seem highly restricted in discourse: they serve to change the current topic of conversation, and contrastiveness in some sense is important for the adequate use of these marked constructions. The relevance of contrastiveness in the case of topicalization in English is illustrated by Prince (1984: 218) by the difference in appropriateness of examples (13) and (14):

(13) A: You see every Woody Allen movie as soon as it comes out.
B: No - *Annie Hall* I saw (only) yesterday.

(14) A: Why are you laughing?
B: # *Annie Hall* I saw yesterday. I was just thinking about it.

The proper use of left-dislocation in (15) is also motivated by the fact that the left-dislocated constituent *gallstones* is contrasted with *nervous breakdown* (cf. Prince 1984: 221):

(15) "Everybody has their little bundle, believe me. I'll bet she had *a nervous breakdown.* That's not a good thing. *Gallstones,* you have them out and they're out. But *a nervous breakdown,* it's very bad … "

Contrastiveness in combination with topicality is not only expressed by marked syntactic constructions but also by **intonational** patterns. Here the "I-contour" (fall-rise accent) in German or the "B-accent" in English are the most thoroughly investigated phenomena. Büring (1997: 69) claims that recourse to alternatives is typical not only for focusing but also for topics realized by the fall-rise, in the latter case giving rise to the "residual topic":

(16) Q: What did the pop stars wear?
A: The [FEMALE]$_T$ pop stars wore [CAFTANS]$_F$.
Residual Topic: What did the male pop stars wear?

On the basis of the data presented we can conclude that the notion of contrast must be regarded as a linguistically relevant phenomenon, because contrast is a necessary condition for the use of certain syntactic and phonological means - both in the field of focusing and topicality. The data from different languages show, however, that this notion is important to a different degree - depending on which language and which specific structure is considered. In those languages where formal marking of contrast is only optional and not unambiguous (as in English or German), the strict distinction of the contrastive and non-contrastive cases is empirically not so well-founded - supporting the gradient view of contrast. From a cross-linguistic perspective it seems well-motivated, though, to assume that the notion of contrast has a crucial impact on linguistic forms, in many languages demanding an obligatory formal realization.

4 Contrast types

The data give empirical evidence also for the hypothesis that there is a need for differentiating the notion of contrast, because languages require a division between contrast and non-contrast with recourse to *different* criteria presented above. Firstly, the distinction between contrasting within a closed, limited set on the one hand, and recourse to alternatives within an open set, on the other, seems motivated: in the former case, contrastiveness has a **pragmatic** character and the realization of contrast is relevant, whereas in the latter case contrast is **semantically** anchored and operation on quantificational domains seems important. Secondly, we need to distinguish different semantic operations on quantificational domains and assume different types of exclusion.

4.1 Contrast: pragmatic vs. semantic level

The distinction between the pragmatic and semantic level of contrasting in the case of **focusing** is motivated by the syntactic difference in their realization. In both cases the focus type involved is "narrow focus", leaving one part of the sentence backgrounded.

Movement of a narrow focus into the left sentence-peripheral position is possible in many languages (in Finnish, Spanish, Italian, Rumanian, Russian), but only when the relevant set is limited. This type of contrast I will call **CONTRAST** - using a special notation with capital letters. The difference between the grammaticality of the Italian examples (17) and (18) (cf. Zubizzareta, 1998: 20) can only be explained by the fact that the explicit mentioning of the relevant alternative(s) (or at least their salience in the speaker's and hearer's minds) is a necessary condition for movement into the left periphery:

(17) Who ate an apple?

*Gianni ha mangiato una mela.
'Gianni has eaten an apple.'

(18) GIANNI ha mangiato una mela (non Piero).
'Gianni has eaten an apple (not Piero).'

For the semantically anchored "narrow focusing", where recourse to a *closed* set is not absolutely necessary, the target position of the focus movement is obligatorily verb-related - without the restriction on the location on the left-periphery. In Hungarian the preverbal position (Spec-F-position) is involved (19) and Focus movement takes place in surface structure:

(19) A: Ki evett almát tegnap?
who ate apple-Acc yesterday
'Who ate apple yesterday?'

B1: Tegnap [FOCUS **János**] evett almát.
yesterday John ate apple

B2: * [FOCUS **János**] tegnap evett almát.

B3: * Tegnap evett almát [FOCUS **János**].

A contextually anchored contrast is even possible without the semantically based exclusion. This is typical for contrastiveness in the field of **topicality** - expressed by means of topicalization and by using different types of left dislocated structures which are employed "to mark a shift in attention from one to another of two or more already active topic referents" and "often have a 'contrastive' function" (Lambrecht 1994: 183). For this type of contextually anchored contrast I would like to suggest the label **LD-CONTRAST**.

4.2 Semantic distinctions

In the linguistic literature two types of "exclusion" of alternatives are distinguished. Most attention has been paid to the type of exclusion typical of the Hungarian **Focus Operator** specified as "exhaustive identification" (É. Kiss 1998).

The identification of one member of a set performed by an operator focus in the Hungarian sentence (20-B) entails that all other members of the set are excluded:

(20) A: Ki járLundban egyetemre?
who is going Lund-in university-to
'Who is studying in Lund?'

B: [PÉTER]$_F$ járLundban egyetemre.
Peter is going Lund-in university-to
'It is PETER who is studying in Lund.'

A further type of exclusion is the operation type which is performed when elements are realized with the fall-rise accent. Even by highlighting with the prenuclear rising accent of an "I-contour", a set of alternatives is induced and excluded. Nevertheless, with the rising accent the exclusion is "weakened". This accent type signals thus not the exclusion of *all* other members of the set (21), but that there is *at least one* member for which the predication (or a part of it) does not hold:

(21) [PÉTER]$_{CT}$ [LUNDBAN jár egyetemre]$_F$.
'As for PETER, he is studying in LUND.'

Also Jacobs (1997) (commenting on Büring's 1997 proposal) emphasizes that the first highlighting of "I-Topikalisierung" contains recourse to alternatives, but this has not the same status as the one connected with focusing. The crucial difference between them is that only after utterances with "I-Topics" can questions be left open; this is not possible with utterances that only contain focus.

To express the similarity between the operator focus and the contrast expressed by the fall-rise, it was suggested in Molnár (1998) that both notions are associated with the feature [+ exclusive]. This feature appears, though, in the two cases in combination with different values of another feature [+/- exhaustive]: For the focus operator the features [+ exclusive] and [+ exhaustive] are characteristic expressing that all alternatives are excluded, whereas the contrastive topic contains the combination [+ exclusive] and [- exhaustive] , indicating the exclusion of only one or some of the relevant alternatives. The difference between the two exclusion types can thus be described by the opposition of "**strong exclusion**" ("*all*-exclusion") and "**weak exclusion**" ("*some*-exclusion").

The information structural value of the weak exclusion is differently judged in literature: it is subsumed either only under focus, called "Contrafocus" by Kenesei (1989), or only under topic - called "S-Topic" by Büring (1997), "I-Topik" by Jacobs (1997). I would like to suggest the label "**I-CONTRAST**" for this exclusion type, emphasizing by this term not only the importance of intonation for this type of contrast, but also that it cannot be restricted simply to topicality or focusing.

(22) **CONTRAST TYPES**:

	focus?	topic?
closed set	**CONTRAST**	**LD-CONTRAST**
open set	**FOCUS OPERATOR**	**I-CONTRAST**

5 The linguistic status of contrast

In the last part of the paper I would like to argue that the linguistic status of contrast must be differentiated: contrastiveness is not only a feature of topicality and focusing, but it seems plausible to regard contrast - in certain languages and certain constructions - as an autonomous phenomenon of information structure.

This proposal is, however, not uncontroversial. Lambrecht (1994: 290) claims that "contrastiveness, unlike focus, is not a category of grammar but the result of the general cognitive processes referred to as 'conversational implicatures"'. Quite different is Chafe's approach, according to whom "the status focus of contrast is different from the status new" (1976: 38). Chafe separates the notion of contrastiveness from the notion of focus, on the one hand, claiming that the latter participates only in the given-new distinction. On the other hand, he argues for the distinction between "real topics" of the topic-prominent languages and contrastiveness in the so-called "double foci of contrast". The latter is found in the left dislocation (23) and topicalization structures of English (24), containing "possible pairings of theatrical events with certain times":

(23) As for the pláy, John saw it yésterday.

(24) Yésterday, John saw the pláy.

Chafe emphasizes that contrastiveness in both cases - in "single focus of contrast" as well as in "double focus of contrast" - should be regarded as a further "packaging phenomenon" *besides* topic and focus. Data from several languages demonstrate, however, that contrast can overlap with topicality and focusing in different ways. According to the main hypothesis of this paper, from a cross-linguistic perspective, the notion of contrast cannot be reduced to a feature serving to differentiate topicality and focusing, nor can it be regarded as a notion standing *besides* topic and focus in information structure. As contrast – in certain languages, in certain structures – can be superimposed on and combined with the notions of topic and focus, the notion of contrast should be added to the inventory of relevant information structural categories. The two main instances of contrast attested empirically which support this claim will be discussed below.

5.1 I-CONTRAST

As to the phonologically realized type of contrastiveness, the "**I-CONTRAST**", it is easy to prove that this contrast type does more than simply identify topicality: the analysis of the constituent carrying the (fall-)rise accent in (25) as topic seems

plausible, whereas the contrast marked by the fall-rise in (26) cannot be related to topicality. In these case one could only claim a secondary focusing, without topicality, though:

(25) [Auf der /NEUNundfünfzigsten Straße]$_T$ habe ich [die SCHUHE\]$_F$ gekauft.
"On fifty NINTH Street I bought the SHOES." (Büring 1997)

(26) Man \/MUSS das Buch \NICHT mögen (, aber man KANN).
one must the book-Acc not like but one can
(Jacobs 1997)

In English the assignment of the fall-rise accent is not restricted to intonation structures containing two pitch accents, but is possible with a single pitch accent. Hetland (2001) claims that in these cases (27) the constituent with the fall-rise accent must be regarded as the nuclear focus of the sentence:

(27) Q: Did you feed the animals?
A: I fed the ˇcat.
[Inf.Focus, fall-rise]

The function of "I-Contrast" seems identical in all analyzed cases - independently of its cooccurence with topic or focus: it evokes alternatives and is connected with open questions, motivating the analysis that I-Contrast is a superimposed notion on both topicality and focusing.

5.2 KONTRAST

The second type of contrast which can be claimed to be an autonomous notion of information structure is **KONTRAST** in Finnish. In Finnish, the *type* of pitch accent (i.e. the fall-rise contour) is, however, not an obligatory, decisive means for expressing contrastiveness. Here, it is primarily the syntactic position that is involved in the formal realization of contrastiveness. The Finnish data demonstrate clearly that "the semantic function of contrast may cut across, and supersede the functions of topic and focus" (É. Kiss 1995: 6). The KONTRAST-Position can apparently host not only a contrastive focus, but also a contrastive topic and appears in two different sentence patterns called by Vilkuna (1995: 249) "TOP pattern" and "FOCTOP pattern". Whereas in the so-called FOCTOP-sentences (29-b), the KONTRAST-position is occupied by the nuclear focus of the sentence, in the TOP-sentences (30-b), the highlighted constituent in the KONTRAST-position is topic, followed by a further highlighted constituent, namely the nuclear focus. Cf. Järventausta & Molnár (2001):

FOCTOP-sentence (28-b):

(28) A: Pekka lensi **Tukholmaan**.
Pekka flew to Stockholm
'Pekka flew to STOCKHOLM.'

B: [KONTRAST **Reykjavikiin**] Pekka lensi.
'Pekka flew to REYKJAVIK."

TOP-sentence:

(29) a. [KONTRAST **Tukholmaan**] Pekka lensi [FOCUS **Finnairilla**],
'To STOCKHOM, Pekka flew by FINNAIR,'

b. [KONTRAST **Reykjavikiin**] (Pekka lensi) [FOCUS **Icelandairilla**].
'To REYKJAVIK, Pekka flew by ICELANDAIR."

As to the *semantic* operations performed by KONTRAST in the case of a contrastive topic and a contrastive focus, É. Kiss (1998: 270) notes that they are different in the two cases: while "the contrastive focus exhaustively identifies the subset for which the predicate holds", "the identification performed by the contrastive topic is nonexhaustive." The Finnish contrast is, thus, as opposed to the Hungarian focus operator, semantically fairly underspecified (Vallduví & Vilkuna 1998: 83).

The *syntactic analysis* of KONTRAST in Finnish requires that several syntactic and pragmatic features be considered. The position is *left peripheral* and can only host a constituent if it refers to explicitly mentioned - or at least contextually salient - alternatives within *a closed set.* The KONTRAST-position is somewhat different from the sentence-peripheral CONTRAST-position of other languages (like Italian or Russian), in that it cannot only be occupied by the contrastive focus, but even by a contrastive topic. The KONTRAST-position in Finnish – as opposed to the Focus-position in Hungarian – is *not verb-related,* which is also typical of the Topic position in Hungarian: the difference is, though, that the Topic position can be iterated, whereas *KONTRAST is unique.*

The listed features require the modification of the structural representation of clauses proposed by Rizzi (1997) within the generative framework. Rizzi dissolves the left-peripheral complementizer layer - the category of CP - into a number of different functional projections, hosting besides a free functional morpheme various operator-like elements such as topics, focalized elements, interrogative and relative pronouns. Below the highest functional phrase ForceP which provides information about the clausal type (or the "specification of Force"), he assumes a topic-focus field, which in Italian involves a FocP surrounded by recursive TopPs. In Rizzis SPLIT-CP-model (1997), however, only the TOPIC- and FOCUS-projections host constituents with specific discourse interpretations, and its expansion thus seems necessary by introducing a KONTRAST-P in the left periphery of the sentence (30):

(30) ForceP **KontrP** TopP* FocP TopP* FinP

The KONTRAST-projection in the articulated CP-domain is not verb-related and cannot be licensed by the same type of "Affective Operator Criterion" (cf. Rizzi 1991) as the Focus. According to my proposal, a K-feature should be made responsible for the licensing of the constituent in the Specifier-position of the KONTRAST-projection. This K-feature must guarantee the so-called "discourse connection" (cf.

Haegeman & Guéron 1999: 536) and requires (i) the absolutely leftmost position, (ii) the explicit mentioning of relevant excluded alternatives within a closed set, (iii) - or alternatively, in case of Contrastive topic, that the discourse connection be warranted by the givenness or high degree of "identifiability" of this constituent.

6 Conclusion

The data surveyed in this paper have served to show that the notion of contrast is a linguistically relevant phenomenon and cannot merely be regarded as "conversational implicature" arising from focusing in certain conversational contexts (Lambrecht 1994: 291). Even though in certain languages formal marking of contrast is optional – thus not allowing a clear-cut distinction between contrastive and non-contrastive cases – empirical data from a great number of languages show that contrast has a crucial impact on linguistic forms, requiring the obligatory use of certain phonological and syntactic means.

It was also argued that further distinctions within the notion of contrast itself seem necessary, in order to make explicit the impact of contrastiveness on linguistic forms: above all the separation of the pragmatic level from the semantic - quantificationally based - level is motivated by linguistic data. The different types of contrast – realized in certain positions and with specific accent types – are, however, not directly bound to certain discourse pragmatic interpretations. Contrast is thus compatible both with focusing and topicality, with the restriction that the preferred relations are dependent on the type of contrast. The possibilities and restrictions of the co-ocurrence of the different contrast types with topic and focus are partly due to universal, and partly to language specific regularities.

The most important claim of the work presented is that the linguistic status of contrast must be analyzed in a more differentiated way than has been the case hitherto in linguistic research: from a cross-linguistic perspective, contrastiveness is not simply a feature of focusing and topicality. Contrast must be established - on the basis of evidence from certain languages and in certain constructions - as a further category of information structure, *superimposed* ontopic and focus.

Bibliography

[1] Bolinger, D.L. 1961. Contrastive accent and contrastive stress. *Language* 37 (1), 83-96.

[2] Bußmann, H. 1990. *Lexikon der Sprachwissenschaft.* Kröner.

[3] Büring, D. 1997. *The meaning of topic and focus: the 59th Street Bridge accent* (Routledge studies in German linguistics 3]. London: Routledge.

[4] Büring, D. 1999. Topic. In P. Bosch and R. van der Sand (eds) *FOCUS. Linguistic, cognitive, and computational perspectives. Part II Semantic interpretation of focus phenomena.* [Studies in Natural Language Processing], 142-165. New York and Cambridge: Cambridge University Press.

[5] Chafe, W. L. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C.N. Li (ed.) *Subject and Topic*, 25-55. New York, San Francisco and London: Academic Press.

[6] Dretske, F. J. 1972. Contrastive statements. *The Philosophical Review* 81: 411-437.

[7] É. Kiss, K. 1995. Discourse configurational languages. Introduction. In É. Kiss, K. (ed.) *Discourse Configurational Languages.* 3-27. New York and Oxford: Oxford University Press.

[8] É. Kiss, K. 1998. Identificational focus versus information focus. *Language* 74 (2): 245-273.

[9] Etxepare, R. 1997. Two types of focus in Basque. In B. Agbayani and S.-W. Tang (eds) *Proceedings of the 15th West Coast Conference on Formal Linguistics*, 113-127. Stanford. CSLI. Distributed by Cambridge: Cambridge University Press.

[10] Haegeman, L. and Guéron, J. 1999. *English Grammar. A generative perspective.* Malden, Mass.: Blackwell.

[11] Halliday, M.A.K. 1967. Notes on transitivity and theme in English II. *Journal of Linguistics* 3: 199-244.

[12] Hetland, J. 2001. Contrast, the fall-rise accent, and information focus. Ms.

[13] Jackendoff, R.S. 1972. *Semantic interpretation in generative grammar* [Studies in Linguistics Series 2]. Cambridge, Mass. and London: The MIT Press.

[14] Jacobs. J. 1988. Fokus-Hintergrund-Gliederung und Grammatik. In H. Altmann (ed.) *Intonationsforschungen*, 89-134. [Linguistische Arbeiten 200] Tübingen: Niemeyer.

[15] Jacobs, J. 1997. Bemerkungen zur I-Topikalisierung. *Linguistische Berichte* 168, 91-133.

[16] Järventausta, M. and Molnár, V. 2001. Diskurskonfigurationalität im Finnischen und im Ungarischen. Ms.

[17] Kenesei, I. 1989. Logikus-e a magyar szórend? [Is Hungarian word order logical?]. *Általános Nyelvészeti Tanulmányok* 17: 105-152.

[18] Lambrecht, K. 1994. *Information structure and sentence form.* Cambridge University Press: Cambridge.

[19] Molnár, V. 1998. Topic in focus. On the syntax, phonology, semantics and pragmatics of the so-called 'contrastive topic' in Hungarian and German. *Acta Linguistica Hungarica* vol. 45 (1-2): 89-166. Dordrecht: Kluwer.

[20] Prince, E. 1984. Topicalization and left-dislocation: A functional analysis. In S.J. White and V. Teller (eds) *Discourses in reading and linguistics.*

[Annuals of the New York Academy of Sciences 433], 213-225. New York: New York Academy of Sciences.

[21] Reinhart, T. 1982. Pragmatics and linguistics: An analysis of sentence topics. Reproduced by the Indiana University Ling. Club.

[22] Rizzi, L. 1991. Residual verb-second and the Wh-criterion. *Technical Reports in Formal and Computational Linguistics 2* (Faculté des Lettres, Université de Genève).

[23] Rizzi, L. 1997. The fine structure of the left periphery. In L. Haegeman (ed.) *Elements of grammar*, 281-337. Dordrecht: Kluwer.

[24] Rochemont, M.S. 1986. *Focus in generative grammar* [Studies in Generative Linguistic Analysis 4]. Amsterdam and Philadelphia: John Benjamins.

[25] Rooth, M.E. 1985. *Association with focus.* Ph.D.diss. University of Massachusetts at Amherst.

[26] Rooth, M.E. 1992. A theory of focus interpretation. *Natural Language Semantics* 1: 75-116.

[27] Vallduví, E. 1992. *The informational component.* New York and London.

[28] Vallduví, E. and M. Vilkuna. 1998. "On rheme and kontrast". In P.W. Culicover and L. McNally (eds) *The limits of syntax*, 79-108. [Syntax and Semantics, Volume 29].

[29] Vilkuna, M. 1995. Discourse configurationality in Finnish. In K. É. Kiss (ed.) *Discourse configurational languages*, 244-268. New York and Oxford: Oxford University Press.

[30] Zubizarreta, M.L. 1998. *Prosody, focus, and word order* [Linguistic Inquiry Monographs 33]. Cambridge, Mass. and London: The MIT Press.

Complement set reference

RICK NOUWEN
UiL-OTS Universiteit Utrecht
rick.nouwen@let.uu.nl

ABSTRACT. The paper considers the phenomenon of complement anaphora and offers an analysis in the framework of optimality theoretic semantics / pragmatics. I argue that from an interpretation perspective there is a preference for reference to the so-called REFSET over all other sets associated with quantificational structures. Reference to the complement set is thus only possible when two conditions are met: (i) the complement set satisfies a witness-constraint and (ii) the reference set is ruled out as an antecedent for independent reasons.

1 Introduction

Given a quantificational structure $D(A)(B)$, there are three sets that can be associated with it[1]: (i) the reference set (or REFSET), corresponding to the intersection of A and B; (ii) the maximal set (MAXSET) which equals the restrictor A; and (iii) the complement set (COMPSET) which is the set of entities in A which do not have property B. Anaphoric reference to all these three sets is possible.

(1) a. Most students went to the party.
They had a good time. REFSET

b. Most students went to the party.
They like to have a good time. REFSET/MAXSET

c. Few of the students went to the party.
They went to the beach instead. COMPSET

In (1a), the plural pronoun refers to the set of students that went to the party, the reference set. The pronoun in (1b) can also be interpreted as referring to this set, but could also refer to students in general, which corresponds to a MAXSET-interpretation. Finally, in (1c) we encounter reference to the complement set. The pronoun here refers to the students that did not go to the party. This is a case of complement anaphora.

[1]Here are some formal conventions:
In a structure $D(A)(B)$ we call D a determiner, $D(A)$ a quantifier, A the restrictor and B the nuclear scope. A determiner D is monotone increasing or upward entailing in its right argument if for any $C \supseteq B$ and any A, if $D(A)(B)$ holds, then $D(A)(C)$ holds as well. A determiner D is monotone decreasing or downward entailing in its right argument if for any $C \subseteq B$ and any A, $D(A)(B) \rightarrow D(A)(C)$.

Paper presented at the ESSLLI 2001 Workshop on
Information Structure, Discourse Structure and Discourse Semantics
Edited by Ivana Kruijff-Korbayová and Mark Steedman.

What makes complement anaphora interesting is that their distribution relates to formal properties of the determiner involved in the antecedent sentence. Reference to COMPSET is not possible following non-downward entailing quantifiers. This is in contrast with REFSET reference.

(2) a. Few of the students went to the party.
They went to the beach instead. COMPSET

b. Few of the students went to the party.
They had a good time. REFSET

(3) a. Most students went to the party.
#They went to the beach instead. COMPSET

b. Most students went to the party.
They had a good time. REFSET

Reference to COMPSET is bad following cardinal downward entailing quantifiers.

(4) a. Less than half of the students went to the party.
They went to the beach instead. COMPSET

b. Less than ten students went to the party.
??They went to the beach instead. COMPSET

Reference to COMPSET is better following partitive cardinal downward entailing quantifiers.

(5) a. Less than twenty-five of the fifty students went to the party.
?They went to the beach instead. COMPSET

These are the basic facts concerning complement anaphora. In section 2 we will discuss the source of the discussion surrounding complement anaphora: a range of psycholinguistic experiments conducted by Moxey and Sanford. Section 3 discusses two analyses that try to explain complement anaphora. In section 4 we will present some new data, focusing on the *interpretation* of plural pronouns. I will argue that from this perspective reference to a complement set is marked. Finally, in section 5 I will give an analysis.

2 Moxey and Sanford's experiments

As we have seen anaphoric links to COMPSET are subject to other constraints than reference to the other sets associated with quantificational structures. This phenomenon has been thoroughly studied in a series of psycholinguistic experiments (see (Sanford and Moxey 1993)). In these experiments subjects were confronted with a single quantified statement and were asked to make up a sensible continuation beginning with the plural pronoun *They*.

(6) Q of the MPs were at the meeting. They ...

Subjects were also asked to indicate what the plural pronoun referred to in their continuation. Here they could make a choice between five categories: *MPs in general, all MPs, MPs who went to the meeting, MPs who did not go to the meeting* and *none of the above*. Independent judges checked all the utterances and reference indications. In 98% of these cases the judges agreed with the judgments of the subjects.

An alternative experiment tested for intra-sentential complement anaphora, using a structure like (7).

(7) Q of the MPs attended the meeting, because they ...

The results showed a preference for COMPSET reference following the determiners: *hardly any, not many, very few and few*. This preference was not present with the determiners *a few* and *many*. A special case was *only a few* which only showed COMPSET reference in the task involving a structure like (7).

In a different study complement set reference was studied with proportional numerical expressions. The continuation method and the use of judges was as in the experiment above (cf. (Sanford and Moxey 1993):77). Here it was found that COMPSET continuations were favored following the determiners: *less than n%*. The other determiners (*n%, only n% and more than n%*) showed few to no continuations containing complement set reference.

It is important for the current paper to note that *preference* for a certain continuation should be interpreted rather weakly as roughly indicating that more than half of the subjects use a complement anaphor. That is, REFSET continuations *did* occur following downward monotone quantifiers.

In their analysis Moxey and Sanford's basic assumption is the functional usefulness of COMPSET. Earlier experiments have shown that expressions like *few*, *not many*, *a few* and *only a few* roughly specify the same quantity (see (Sanford and Moxey 1993):ch2). There is however a fundamental difference between these determiners in that some of them seem to report on a deviation from expectation, while others do not. Moxey and Sanford link this observation to the notion of *negativity* (cf. (Klima 1964)). Negative determiners, the argument continues, put *focus* upon their COMPSET. This means that in a discourse the COMPSET is more prominent than the REFSET. There will be a preference for reference to the set focused by the determiner. Moreover, this focus will have the thematic effect that, following a COMPSET-licensing quantificational structure a reason is given why REFSET is smaller than was to be expected. This is confirmed by a series of experiments showing that COMPSET-continuations generally are (in Moxey and Sanford's terminology) of a *reason-why-not* nature. In these experiments independent judges were asked to indicate which of four types a produced continuation belonged to. The COMPSET continuations were dominantly classified as indicating "the reason why the predicate is not true of the refset" (see (Sanford and Moxey 1993):66).

3 Explaining COMPSET reference

A common reaction within the (formal) semantic community on the results of Moxey and Sanford's experiments was that complement anaphora do not *really* involve complement set reference, but instead are a case of some sort of *pseudo-reference* established by a generalization over the maximal set (see especially (Corblin 1996)). This view still enjoys considerable popularity although it creates some fundamental problems. For one, the generalization seems to be allowed since the antecedent sentence expresses the inferiority of the REFSET. Problematic then is that this generalization should also be allowed following a sentence like *Not all A B*, where the COMPSET might consist of but a few exceptions. Other authors have proposed other variants of pseudo-reference (e.g. (Geurts 1997)) and the result is an interesting discussion on the *reality* of complement set reference. Due to space considerations I will not go into that discussion here and presuppose the reality of complement anaphora. (For a detailed elaboration on this issue as well as more arguments against pseudo reference cf.(Nouwen 2001)).

Once we accept that complement anaphora involve reference to the complement set, the question arises how this anaphoric link comes about.

3.1 Emptiness

In their presentation of optimality theoretic semantics (de Hoop and Hendriks 2001) suggest that the data can be explained by the interaction of pragmatic constraints.

The basic assumption behind optimality theoretic semantics is the free interpretation hypothesis. This is nothing but the driving force behind the generator of the optimality theoretic system of finding the most optimal interpretation for a certain linguistic form. Relevant to our story is that given this hypothesis a generalized quantifier can take any domain of quantification. Thus REFSET reference and COMPSET reference of a pronoun compete in the candidate set of the interpretation of a continuation following a quantificational structure.

The basis of de Hoop and Hendrik's explanation is the constraint EMPTINESS.

(8) EMPTINESS: As the antecedent of an anaphoric expression, do not choose a set that is or may be empty

By itself, this constraint can already explain the correlation between downward monotonicity and complement set reference. The reference set of a downward monotone quantifier can, of course, be empty. Reference to this set will thus violate EMPTINESS. The violation pattern is reversed for monotone increasing quantifiers.

The constraint is especially interesting with respect to cardinal downward monotone quantifiers. In those cases both REFSET and COMPSET can be empty: the first due to downward monotonicity; the second due to the fact that for all these quantifiers D(A)(A) holds. Both interpretations thus violate EMPTINESS. The optimality

decision is now left to a constraint which is ranked lower. De Hoop and Hendriks have a constraint called *forward directionality* (or FORWD), which expresses a preference for referring to the REFSET.[2]

There is however a complication. Emptiness is not the only constraint in competition with forward directionality. There is also the high ranked *avoid contradiction* or AVOIDC. If ranked higher then forward directionality, avoid contradiction can explain the facts in (9). In the second sentence in (9a), the determiner is interpreted as quantifying over the reference set of the preceding sentence. Only when a contradiction is encountered, as in (9b), does interpretation prefer returning to the larger domain, namely that of students (cf. (de Hoop and Hendriks 2001):(28)/(32)).

(9) a. Ten students attended the meeting. Three spoke.
b. Ten students attended the meeting. Twelve spoke.

This additional interaction causes problems for the analysis of cardinal decreasing quantifiers. As we have seen, both the reference set and the complement set of these structures violate emptiness. Forward directionality then prefers the reference set. But in the examples of interest, like (4b) repeated below, avoid contradiction has an unwanted decisive preference for the complement set.

(4b) Less than ten students went to the party.
??They were too busy.

There is another problematic side effect of their proposal. There seems to be no room for optionality. Remember that the results from Moxey and Sanford's experiments did not show that REFSET reference does not occur following downward monotone quantifiers. They merely showed there to be a preference. The optimality model proposed by de Hoop and Hendriks suggests that REFSET continuations will never follow non-cardinal downward entailing quantifiers.

We will return to de Hoop and Hendrik's analysis but next consider another explanation of the complement anaphora phenomenon. As we will see in this approach optionality does play an important role.

3.2 Dynamic Quantifiers

Rodger Kibble ((Kibble 1997a)) gives a technical explanation for Moxey and Sanford's experimental results. In dynamic semantics, quantifiers are made dynamic by giving them an existential structure, as in (10).

(10) $Q(A)(B) \;=\; \exists x \wedge \exists y \wedge \max_y(A[x/y]) \wedge \max_x(x \leq y \wedge B) \wedge Q'(y)(x)$

This says that a quantificational structure $Q(A)(B)$ is to be interpreted as the introduction of two (maximal) sets, one satisfying the restrictor A and the other, a

[2]This constraint is actually a much more general constraint expressing the preference for an ongoing reduction of topic range.

subset of the former, satisfying the scope. The quantification succeeds if the two introduced sets are in the quantificational relation Q' (the extension of the determiner Q). The effect is that every quantificational structure introduces MAXSET and REFSET into the context. A well-known problem, however, is that this mechanism only works for MON↑ determiners. Because of the existential nature of the construction, it merely desires there to be *some* pair of maximal sets satisfying restrictor and scope and the Q'-relation. It does not enforce *all* pairs of maximal sets satisfying A and B to be in this relation. Kibble takes this problem to be the source of complement anaphora. A natural way of solving the problem is by defining MON↓ determiners in terms of their duals. There are, however, two possible constructions of duals for each determiner D: one based on the so-called *complement* $-D$ and one based on the *contradual* $D-$. The first can be compared with ordinary wide scope negation of the quantificational structure. This way, the complement of $\mathit{few}(A)(B)$ will be $\mathit{not\ few}(A)(B)$ or $\mathit{many}(A)(B)$. To return to *few* we simply negate once more: $\mathit{not\ many}(A)(B)$. Now we have an alternative to *few* in terms of the (unproblematic) MON↑ determiner *many*. The other way of doing this is by using the contradual: $\mathit{few}(A)(\mathit{notB})$. It is easy to see that $D-$ (just like $-D$) is MON↑ whenever D is MON↓ . Negating once more gives us the second alternative to *few*. All this is illustrated in (11).

(11) Few of the students went to the party.

a. Not many of the students went to the party.

b. Many of the students did not go to the party.

Using these insights to come to a dynamic interpretation for MON↓ determiners results in an ambiguity. When the complement is used, the dynamic structure introduces the reference set $A \cap B$ into the context. But when the contradual is used the reference set will be $A - B$. This, according to Kibble, explains why (and why only) we encounter complement anaphora with downward entailing quantifiers.

At first sight, Kibble's solution to the complement anaphor puzzle might seem a side effect of a technical oddity. But there is an important insight hidden behind all this. The correlation between monotonicity properties of a determiner and its possibility to license complement set reference follow from logical inference. As Kibble puts it: "[A] plural pronoun can pick up an antecedent which is either explicitly introduced or logically inferable from antecedent information, provided that the referent functions as a *witness set* which serves to verify the antecedent sentence(s)." ((Kibble 1997b):126). Thus, while a small proportion of MPs being absent is not a witness for *Most MPs attended the meeting* (there could be more of them), a large set of absentees will verify that *few of the MPs attended the meeting* (Kibble refers to such witnesses as *negative witnesses*). With cardinal quantifiers there is no way of knowing whether a large set verifies the antecedent sentence or not, simply because the domain is not known. Thus, COMPSET can only be a witness set for a downward entailing proportional quantification.

There is however a gap between the formal explanation and Kibble's conceptual story involving witnesshood. Formally, nothing prevents us from forming the

duals of a monotone increasing quantifier and such a dual would also output a complement set. The only reason why it is not there is because MON↑ quantifiers do not have the same problems with existential structure as the downward ones have. On the conceptual side, then, it is not clear what it means for a set to *verify* its antecedent sentence and why we wouldn't assume such a set to be maximal. One also can start wondering whether REFSETs are always witnesses. Following Kibble's analysis a small set of party going students, for instance, would not verify a sentence expressing that few of the students attended the party, but still pronominal reference to this set *is* possible. So apparently reference sets have unconditional witnesshood. Moreover, a set of students not going to a certain party can only count as a negative witness for few of the students attending that party, once we have established that this set is a large proportion of the whole set of students. But in order for us to know that, we have to know the cardinality of the (contextual) domain of students. Genuine complement sets are thus also in need of cardinality information. But exactely this was why Kibble excluded non-proportional determiners from producing negative witnesses.

Although Kibble makes clear that some sense of witnesshood is needed to explain complement set reference, his formulation of this key notion seems to raise a lot of questions. In the next section, however, we will see that accounting for the class of COMPSET licencing determiners is not the only problem for the two analyses we have considered.

4 COMPSET from the hearer's perspective

Notice a detail in both Hendriks and de Hoop's analysis and that of Kibble's. Both implement the insights in a framework meant to describe natural language interpretation, while Moxey and Sanford's data shows us primarily facts of production. Although this data clearly has repercussions for comprehension (e.g. the expectation of a certain kind of reference following a certain type of determiner), I will try to show here that from an interpretation perspective the data is more subtle.

We have already seen that Moxey and Sanford found a thematic effect in COMPSET continuations: they usually specify the reason why a relatively large proportion of the domain did *not* satisfy the predication. From an interpretation perspective it is interesting to see what happens if we use complement anaphora in continuations other than the thematic preference found by Moxey and Sanford. Consider the examples in (12).

(12) a. Few of the students went to the party.
I know who they are.

b. Few of the American presidents in the 20th century were in power for two consecutive terms. My history teacher made me learn their names by heart.

c. Few of these balls are blue.
Can you point them out for me?

In all these discourses the predications in the first and second sentence are neutral with respect to one another. Resolving the plural pronoun shows a clear preference for REFSET reference. It appears then that the default interpretation for plural pronouns is the intersection of restrictor and scope no matter the formal properties of the determiner. As far as I know, this asymmetry between production and interpretation has not been noticed.

The point made here is that REFSET is (if possible) the preferred resolution of a plural pronoun. The complement set has no such status. During resolution it is overruled by the reference set in neutral situations. It seems to me that COMPSET interpretation is the result of a last resort strategy. Notice that a side effect of *reason-why-not* continuations is that resolving the plural pronoun to refer to the REFSET results in a contradiction. The only non-contradictory resolution is, of course, the COMPSET.

In (Moxey and Sanford 1987), we find an example showing that *reason-why-not* continuations are not obligatory. In (13), there is a clear case of complement set reference, but the fact that the members of COMPSET send their apologies does not really indicate why so few MPs were at the meeting.

(13) Few MPs were at the meeting. They sent apologies for being absent.

Still, once again we see that resolving the plural pronoun to REFSET reference would result in a contradiction.

Notice how both accounts given in the previous section are not able to cope with the general preference for REFSET reference. De Hoop and Hendriks' analysis will always prefer the non-empty set over the (possibly) empty one, no matter the predications involved (i.e. ignoring the problematic interaction with avoid contradiction, but more on that later). Following Kibble's story, COMPSET *is* a reference set, so the desired difference between the two sets does not exist.

More support for the view that the interpretation of anaphora shows a preference for REFSET comes from explicit reference to the complement set. Notice the following contrast.

(14) a. Few of the students went to the party.
The others stayed at home instead.

b. Few of the students went to the party.
The others had a good time. non-party-goers$>>^{??}$party-goers

In (14a) we see that we can replace the complement anaphor with an explicit reference to the complement set: *the others*. This definite description takes the complement of the REFSET relative to some domain of quantification (the MAXSET). If we take Kibble's analysis serious, there is really no difference between COMPSET and REFSET other than that they are the reference set of dynamically different but logically equivalent representations of a quantificational structure. In other words, there is no apparent reason why *the others* would not accept the COMPSET as an antecedent reference set. This however gives us the odd (14b).

Summarizing, for the interpretation of plural anaphora, we find the following paradigm in terms of emptiness. (A) corresponds to monotone increasing contexts; (B) to cardinal downward monotone ones and (C) represents the cases with true COMPSET licensers (proportional monotone decreasing determiners).

(A)	COMPSET possibly empty, REFSET non-empty:	*COMPSET/REFSET
(B)	COMPSET and REFSET both possibly empty:	*COMPSET/REFSET
(C)	REFSET possibly empty, COMPSET non-empty:	REFSET>COMPSET

5 Analysis

The first question arising from all this is how to derive the potential cancelling of the preference for REFSET. One way to account for this is to return to an optimality approach and add the high-ranked constraint *avoid contradiction* (henceforth: AVOIDC). This constraint seems fit for accounting for the last-resort strategy necessary for resolving pronoun reference to COMPSET. But it is easy to see that there can be no ranking of AVOIDC, FORWD and EMPTINESS such that it accounts for the data. The reason is that if we rank EMPTINESS higher than AVOIDC we get no preference for REFSET when it is potentially empty but lacking contrast in predication. So would we now choose to rank the two constraints the other way around, then we predict COMPSET readings in cases like (3a), repeated here.

(3a) Most students went to the party. #They went to the beach instead.

The obvious problem is that EMPTINESS is only supposed to be decisive for COMPSET reference. As we have seen, we actually prefer referring to the possibly empty REFSET over reference to the guaranteed non-empty COMPSET. Thus the proper role of EMPTINESS appears not to be a general constraint on reference at all, but it appears to tell us which sets are potential antecedents and which are not.

Here we stumble on an important relation between emptiness and witness-hood. Both show us the logical accessibility of a set. EMPTINESS seems to exclude possibly empty reference sets. We also saw that the formalization of witness-hood predicted that REFSETs are always available as an antecedent. Let us therefore replace EMPTINESS with a new constraint, a reformulation of what it means to be a witness:

(15) WITNESS: Be a witness, i.e. as an antecedent of an anaphoric expression choose an accessible referent or choose a constructed referent which is guaranteed to be non-empty.

The intuition behind this constraint is that reference to an introduced referent is less dangerous than reference to a constructed one, since the properness of such a construction is not guaranteed. Kamp and Reyle proposed that in discourse representation theory subtraction is not a permissible operation for the forming of an

		WITNESS	AVOIDC	FORWD
Most(A)(B). They ¬B	REFSET		*	
Most(A)(B). They ¬B	COMPSET	*		*
Most(A)(B). They $C_{neutral}$	REFSET			
Most(A)(B). They $C_{neutral}$	COMPSET	*		*
Few of the (A)(B). They ¬B	REFSET		*	
Few of the (A)(B). They ¬B	COMPSET			*
Few of the (A)(B). They $C_{neutral}$	REFSET			
Few of the (A)(B). They $C_{neutral}$	COMPSET			*
Less than ten (A)(B). They ¬B	REFSET		*	
Less than ten (A)(B). They ¬B	COMPSET	*		*
Less than ten (A)(B). They $C_{neutral}$	REFSET			
Less than ten (A)(B). They $C_{neutral}$	COMPSET	*		*

Figure 13.1: Tableau for the paradigm

antecedent (see (Kamp and Reyle 1993):307). (Kibble 1997b) already suggested that this is a far too strong a constraint on antecedent formation and proposed his *witness* alternative. I propose to alter this again into a constraint which is more like Hendriks and de Hoop's emptiness, with the key difference that it is only applied to constructed plurals. Following the current constraint WITNESS and given some sort of ontology (like for instance the upper semi-lattice structure used by (Kamp and Reyle 1993)) summation of referents is always allowed, subtraction is only allowed once it is defined.

WITNESS is thus a powerful constraint on anaphoric reference. Ranking AVOIDC over FORWD accounts for the data in the previous section, i.e. only contradictive meaning can force us to interpret a pronoun as referring to a complement set. From (3a) it follows that if the complement set is not a suitable antecedent due to WITNESS we should accept any reading resulting from resolving the pronoun to REFSET, even a contradictive one. Hence, AVOIDC is ranked below witness-hood.[3] The effects of this ranking are illustrated in the tableau in 13.1.

WITNESS rules out all references to COMPSET where COMPSET is possibly empty. This only leaves the proportional downward entailing quantifiers. In general REFSET reference is preferred, but this can be overruled in the non-neutral cases, where the predication in the continuation contradicts the predication in the antecedent sentence.

Following a downward entailing cardinal quantifier, WITNESS is violated by COMPSET reference. This is because the set A is unknown and thus $D(A)(A)$ could be valid. In the introduction I mentioned that following these quantifiers COMPSET reference is not plainly out, but *bad*. The reason for this is that contrary to increas-

[3] See (de Hoop 2000) for more on unintelligibility, optimality theory and avoid contradiction.

ing quantificational structures there is an indirect way the complement set can still be a witness. Sometimes, there is an implicit domain available triggering a partitive reading and making the complement set into a negative witness. Consider for example (16), where the fact that we know how many nominees there are for an Oscar category (five) provides the implicit partitive reading.

(16) Less than two nominees for the Oscar for best actor wore a tuxedo. They simply wore a suit and tie.

Moreover, it seems to me that sometimes a domain can be accommodated. Consider (17):

(17) Less than ten students handed in their essay.
They had all sorts of excuses.

We do not know how many students there are, but there is no real reason to assume there only to be ten. The continuation seems to provide evidence for the existence of more than ten students. This allows the COMPSET to be a witness set for the first sentence.

Let us briefly consider the case of *the others*. Notice first how it differs from a complement anaphor in that it can occur following increasing environments.

(18) a. Most students went to the party. The others stayed in.
b. Less than ten students went to the party. The others stayed in.

It seems plausible to characterize *the others* as signaling a shift from the preferred reference to its complement. This behavior follows directly from our analysis once we assign *the others* a semantics like: $\lambda P.P(X - A)$, where X is a context set (the maximal set) and A is the antecedent for *the others*. Notice that WITNESS, since it is a constraint on antecedent-hood, already blocks COMPSET as a potential antecedent for X in increasing and cardinal environments. What is left to be explained is the fact that *the others* cannot refer to REFSET even in the remaining cases. One way to go would be to say that forward directionality (like WITNESS is and like emptiness was) is a constraint on antecedents, i.e. it prefers anaphoric expressions to select a forward directional antecedent. At first, this might seem a break with the original intention of FORWD, viz. reduction of topichood, but notice that such a break only occurs in case the anaphoric expression expresses a shift of some sort. Of course, that is exactly what we want. I leave it to the reader to check that the current proposal indeed covers all the data.

6 Conclusion, reflection, further research

Complement sets are marked antecedents for plural pronouns. Their construction is only permitted once they are guaranteed to be non-empty, but even in this case resolution to COMPSET only occurs once REFSET reference is out for independent reasons.

The interpretation of pronouns referring to a part of a quantificational domain now sums up to the following: interpret the pronoun as a referent or as a necessarily non-empty constructed referent; next –if there is a choice– interpret the pronoun in such a way that there arises no contradiction; finally, if there still is a choice, choose REFSET over all other alternatives.

Reflecting on this resolution strategy we could say that the task of optimality theory here is nothing but mediating between different modules of language use. WITNESS could thus be seen as a hard constraint saying that we should only refer to that which is available or safely inferable. Making sense is thus only possible once we have obeyed witness-hood. Finally, pragmatic preference for a certain type of information flow is only important once we have an intelligible interpretation.

I have remained quiet on one complicating issue. It concerns a problem with reference to the maximal set. In (19) we prefer a maximal set reading over a resolution to COMPSET.

(19) Not all of the animals in this zoo are dangerous.
More than half can be stroked.

Clearly, more than half of the animals can be stroked, not more than half of the non-dangerous ones. If we assume that the maximal set is present as a referent, then following our analysis presented here we predict MAXSET reference to be immune to WITNESS. This, however, poses the question how MAXSET and COMPSET are in competition in examples like (19). A possible solution would be to assume that in addition to forward directionality (violated by both COMPSET and MAXSET) there is a constraint which represents the costs of constructing a new referent out of existing ones. This would account for examples like (19). However, there is little data known involving maximal set reference, so the details will have to be part of further research.

Bibliography

Corblin, F. (1996). Quantification et anaphore discursive: la référence aux complémentaires. *Langages 123*, 51–74. Also as: *Peut-on anaphoriser le complémentaire d'un ensemble?* In: Référence temporelle et nominale. Actes du 3e cycle romand de Sciences du langage, Cluny, avril 1996. Moeschler, J. and Béguelin, M. (eds.).

de Hoop, H. (2000). The problem of unintelligibility in OT semantics. unpublished manuscript.

de Hoop, H. and P. Hendriks (2001, February). Optimalilty Theoretic Semantics. *Linguistics and Philosophy 24*(1), 1–32.

Geurts, B. (1997). Book review of Linda M. Moxey and Anthony J. Sanford. Communicating Quantities. 1993. *Journal of semantics 14*(1), 87–94.

Kamp, H. and U. Reyle (1993). *From Discourse to Logic*. Dordrecht: D. Reidel.

Kibble, R. (1997a). Complement anaphora and dynamic binding. In A. Lawson (Ed.), *SALT VII*, Ithaca, NY: Cornell University.

Kibble, R. (1997b). Complement anaphora and monotonicity. In G. Morrill, G.-J. Kruijff, and R. Oehrle (Eds.), *Formal Grammar*, pp. 125–136.

Klima, E. (1964). Negation in english. In J. Fodor and J. Katz (Eds.), *The structure of language*. Englewood Cliffs, New Jersey: Prentice Hall.

Moxey, L. and A. Sanford (1987). Quantifiers and focus. *Journal of semantics 5*, 189–206.

Nouwen, R. (2001). On the reality of complement set reference. Unpublished manuscript. UiL-OTS, Utrecht.

Sanford, T. and L. Moxey (1993). *Communicating quantities. A psychological perspective.* Laurence Erlbaum Associates.

Discourse Structure and Anaphoric Accessibility

MASSIMO POESIO
University of Essex and University of Edinburgh
Massimo.Poesio@ed.ac.uk

BARBARA DI EUGENIO
The University of Illinois at Chicago
bdieugen@eecs.uic.edu

ABSTRACT. We evaluated the predictions of Grosz and Sidner's theory of anaphoric accessibility using a corpus of tutorial dialogues whose discourse structure was annotated according to Relational Discourse (RDA) Analysis. We found support for Moser and Moore's proposal that only segments with an intentional core should be viewed as introducing new focus spaces; we also found that both embedded cores and embedded contributors should remain open as long as the RDA-segment in which they occur is open, and discuss the implications of this finding for Grosz and Sidner's theory.

1 Introduction

In this paper, we present the results of an empirical study of the relation between discourse structure and anaphoric accessibility. Seminal theories such as (Reichman 1985; Grosz and Sidner 1986; Fox 1987) have been around for about fifteen years. However, only now is it possible to subject them to rigorous empirical testing because of the recent improvements in annotation methodology leading to more reliable annotation techniques, and the resulting increased availability of corpora annotated for discourse structure (Carletta, Isard, Isard, Kowtko, Doherty-Sneddon, and Anderson 1997; Moser, Moore, and Glendening 1996; Marcu 1999).

Much of this recent effort on discourse structure annotation has been based on Rhetorical Structure Theory (RST) (Mann and Thompson 1988) or on similar theories that restrict the range of possible structures. One reason for this is that (each version of) RST specifies a fixed (although fairly large) repertoire of intentions to choose from. While it is still an open question whether it is possible to come up with an exhaustive list of discourse intentions (Grosz and Sidner 1986), in order to achieve reliable annotation it is essential to do so.

As a result, however, the corpora annotated in this way have mainly been used to test new theories of anaphoric accessibility that build directly on RST, such as

Paper presented at the ESSLLI 2001 Workshop on
Information Structure, Discourse Structure and Discourse Semantics
Edited by Ivana Kruijff-Korbayová and Mark Steedman.

the recently proposed Veins Theory (Cristea, Ide, and Romary 1998). Our own aim, at least at this stage, is not to develop a new theory of anaphoric accessibility, but to study the empirical validity of existing theories of anaphoric accessibility, and specifically of the best-known among these, Grosz and Sidner's theory (1986). This used to be a problem because although Grosz and Sidner's theory has originated a coding manual (Nakatani, Grosz, Ahn, and Hirschberg 1995) that has been used at least once (Nakatani 1996), as far as we know there is no sizable corpus coded accordingly. However, recent proposals concerning the mapping between rhetorical structure and intentional structure such as (Moser and Moore 1996b) produced annotated corpora that can be used to investigate the predictions of the theory.

The Sherlock corpus collected at the University of Pittsburgh is a case in point. This corpus is a collection of tutorial dialogues annotated according to Relational Discourse Analysis (RDA), a theory of discourse structure that attempts to merge RST with Grosz and Sidner's theory. One of the features of RDA is to distinguish rhetorical relations into intentional and informational (Moore and Pollack 1992; Moser and Moore 1996b): intentional relations pertain to the effects that the speaker intends his discourse actions to have on the hearer, whereas informational relations pertain to domain relations between the entities being talked about. For example, among RST relations, **evidence** is intentional, whereas **cause** is informational (Moore and Pollack 1992; Moser and Moore 1996b).

In this paper, we report on the results of a study of anaphoric accessibility using this corpus. We first briefly review both Grosz and Sidner's focus-space based theory of the attentional state and RDA. We then discuss how RDA structures can be used to specify opening and closing operations on focus spaces. In the evaluation part of the paper, we present first of all our results concerning the effect on anaphoric accessibility of the more distinctive aspect of RDA, the distinction between intentional and informational relations, and of Moser and Moore's proposal that only the former express DSPs. We then look at a few aspects of the relation between Grosz and Sidner's theory and RDA left open by Moser and Moore, in particular, how embedded segments of different types should be treated. Finally, we analyze the results and evaluate Grosz and Sidner's proposal.

2 Background

2.1 Grosz and Sidner's Theory

According to Grosz and Sidner (G&S), the structure of a discourse is determined by the intentions that the people producing it intend to convey, or DISCOURSE SEGMENT PURPOSES (DSPs). In a coherent discourse, the DSPs are related to form an INTENTIONAL STRUCTURE by either **dominance** relations (in case a DSP is interpreted as contributing to the satisfaction of another intention) or **satisfaction-precedes** relations (when the satisfaction of an intention is a precondition for the satisfaction of a second one).

Anaphoric accessibility of entities in a discourse is modeled by its ATTENTIONAL STRUCTURE, which, according to Grosz and Sidner, is a stack of FOCUS SPACES. G&S propose that when a segment is open, its corresponding focus space, which includes the discourse entities introduced in that segment, is pushed onto the focus stack; when the segment is closed, the focus space is popped, and the discourse entities associated with that focus space are not accessible any more. G&S also argue that the pushing and popping of focus spaces on the stack reflects the intentional structure, in the sense that a new focus space is pushed on the stack whenever the discourse introduces a new DSP subordinate to the present one, and the focus space of the current is popped whenever the associated DSP is satisfied.

This claim about anaphoric accessibility was illustrated in the original paper with a few examples; however, as far as we know, it has not been empirically tested. There are no sizable corpora annotated according to both G&S's treatment of discourse structure and anaphora. More in general, there are no guidelines about how to identify the DSPs in a discourse. Our purpose is therefore twofold: to use RDA to make more specific claims about the DSPs in discourses (of a certain genre and in a given domain), and then to test G&S's claims (within this genre).

2.2 Relational Discourse Analysis (RDA)

Relational Discourse Analysis (RDA) (Moore and Pollack 1992; Moser and Moore 1996b) synthesizes Grosz and Sidner's approach and RST. RDA owes to Grosz and Sidner the idea that discourse is hierarchically structured, and that discourse structure is determined by intentional structure; each RDA-segment originates with an intention of the speaker. But in RDA segments have additional internal structure: each segment consists of one CORE, i.e., that element that most directly expresses the speaker's intention, and any number of CONTRIBUTORS, the remaining constituents in the segment, each of which plays a role in serving the purpose expressed by the core. The notions of core and contributor derive of course from the notions of nucleus and satellite in Rhetorical Structure Theory (RST) (Mann and Thompson 1988), which claims that in each "segment" (text span, for RST) one component should be identified as the 'main' one, and the others as secondary. However, in RST there is a distinction between nucleus and satellite for (almost) all RST relations, whereas in RDA core and contributors are only identified if a segment purpose has been recognized. In this case, each contributor is linked to the core by one intentional relation, and one informational relation. This is unlike RST, in which only a single relation can obtain between nucleus and satellite — see (Moore and Pollack 1992).

In RDA, segment constituents may in turn be other embedded segments, or simpler functional elements: these elements may be either basic UNITS, i.e., descriptions of actions and states, or relational CLUSTERS. Clusters are spans that only involve constituents linked by informational relations; no *core:contributor* structure exists, but they can themselves be embedded.

Unlike G&S's theory and like RST, RDA is based on a fixed number of relations;

1.1 Before troubleshooting inside the test station,
1.2 it is always best to eliminate both the UUT and TP.
2.1 Since the test package is moved frequently,
2.2 it is prone to damage.
3.1 Also, testing the test package is much easier and faster
3.2 than opening up test station drawers.

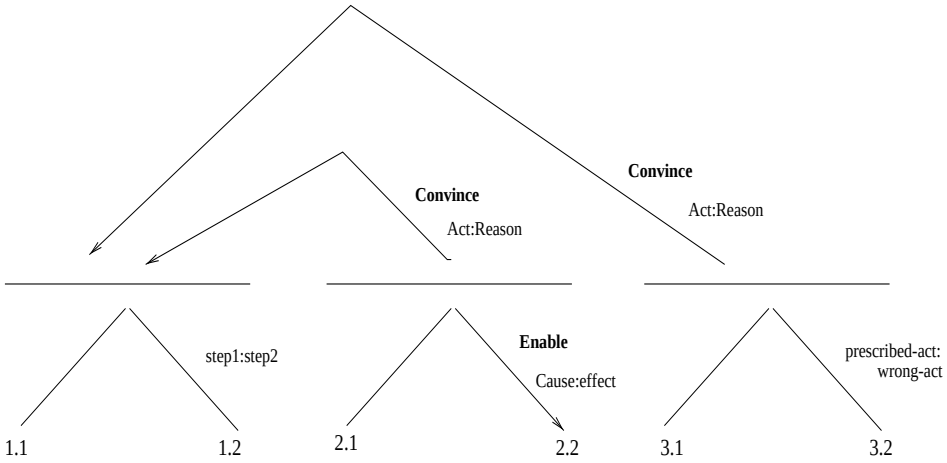


Figure 14.1: A tutorial excerpt and its RDA analysis

in particular, RDA assumes four intentional relations – **convince**, **enable**, **concede**, **joint**–and a larger set of informational relations; this latter set is expected to be domain dependent. In the Sherlock corpus, 23 informational relations are used, of which 13 pertain to causality (they express relations between two actions, or between actions and their conditions or effects) (Moser, Moore, and Glendening 1996).

Figure 14.1 shows a small excerpt from one of the dialogues in the Sherlock corpus, and its corresponding RDA analysis. The text is broken into clauses (UUT is "Unit under test", TP is "test package"). The analysis shows the text to be analyzed as an intentional segment whose core spans 1.1 and 1.2. This segment has two contributors, spanning 2.1 and 2.2, and 3.1 and 3.2 respectively. Graphically, the core is at the end of the arrow whose origin is the contributor; moreover, the link is marked by two relations, intentional (in bold), and informational. In this specific case, the two contributors carry the same intentional and informational relations to the core, but this doesn't need to be the case. The core and the two contributors are further analyzed. The core and the second contributor are analyzed as informational clusters, whereas the first contributor is recognized as having its own intentional structure.[1] Clusters are marked by one informational relation, but

[1]According to the manual used for the annotation (Moser, Moore, and Glendening 1996), an **enable** relation holds "if the contributor [2.1] provides information intended to increase the hearer's

not by intentional relations.

We will come back to the analysis of the text according to G&S in the next section. As regards the analysis of the text according to RST, the structure would presumably be the same, although no double relations would exist, and for every relation one relatum would be considered as the nucleus, the other(s) as its satellites.

3 Anaphoric Accessibility in RDA

In order to use an annotation based on RDA to test Grosz and Sidner's claims about anaphoric accessibility we have to specify a mapping from an RDA structure to focus spaces: which RDA constituents correspond to separate focus spaces, and which focus spaces should be open when a given anaphoric expression is encountered. This mapping is not entirely trivial, because the structure of a discourse according to RDA is much more detailed than the structure that would be assigned to that discourse by Grosz and Sidner. In RDA, each clause is treated as a distinct constituent, whereas in a G&S-style analysis, multiple sentences can be considered a single constituent. Furthermore, G&S make no distinction between cores and contributors, and only allow two relations between intentions, whereas in RDA many types of intentional relations are possible.

Moser and Moore partially specify a mapping between RDA notions and an intentional structure in the G&S sense based on the following principles:

- Every DSP must be associated with a core.
- Constituents of the RDA structure that do not include cores - i.e., clusters (see above) - do not introduce DSPs.

In terms of segments / focus spaces, these principles mean, first of all, that a segment in the G&S sense should always be a segment in the RDA sense (an RDA-SEGMENT): an element with a core and one or more contributors. I.e., no focus space should be pushed on the stack unless a core is recognized. (Moser and Moore leave open the question of whether the reverse should also be true, i.e., whether there should be a 1:1 mapping from G&S-segments to RDA-segments.) Secondly, informational relations do not affect the attentional state, unless associated with intentional relations.

The segment structure that–according to what we have seen so far–we can derive from the RDA analysis in Figure 14.1 should then be as in Figure 14.2. Notice that because informational relations don't give rise to intentional segments, the informational clusters 1.1-1.2 and 3.1-3.2 are not assigned separate segments. In particular, the unembedded core in 1.1-1.2 is not assigned a separate focus space, since it expresses (part of) the DSP associated with the overall RDA-segment.

understanding of the material presented in the core, or to increase the hearer's ability to perform the action presented in the core." (p. 6).

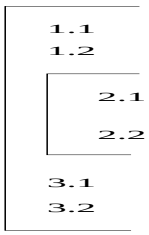


Figure 14.2: A G&S-style analysis for the text in Fig. 14.1

Even these first simple mapping principles already result in different predictions concerning accessibility than one would get from a pure RST analysis. In Fox's analysis, for example (see below), even if the constituent spanning 2.1 and 2.2 were a cluster, not an intentional segment, no antecedent introduced in these propositions would be accessible from 3.1 and 3.2. This is the first claim whose correctness we have to test.

However, the principles proposed by Moser and Moore leave a number of aspects of the mapping open. One question Moser and Moore themselves raise is whether embedded cores, i.e., cores that are themselves RDA-segments (whose possibility they consider, but do not analyze in detail) should be treated as embedded G&S-segments or as part of the same G&S-segment as their embedding RDA-segment. (Examples in which the antecedent of a pronoun is contained in an embedded nucleus are discussed by (Fox), p. 101.) A second question is how should embedded and not embedded contributors be treated: whether all of them should count as separate G&S-segments, or only the embedded ones.

Third, even if (some) contributors push focus spaces, when should these focus spaces be popped? Immediately, or only when the intentional relation is completed? E.g., in the example in Figures 14.1 and 14.2, should segment 2.1-2.2 be popped as soon as we are done processing it, or should it remain on the stack until the whole segment is over, given that it participates in the intentional relation that determines the superordinate segment? Fox's data seem to suggest that antecedents introduced by 'active' non-embedded satellites should be accessible; but even antecedents introduced by active embedded contributors might.

4 The Study

We tested all of these possible ways of using an RDA structure to guide the focus space construction mechanism using the Sherlock corpus. Our results show that the version in which both embedded core and embedded contributors remain on stack has the least number of reference violations.

4.1 Existing Data

The corpus we used is a collection of tutorial dialogues between a student and a tutor, collected within the Sherlock project (Lesgold, Lajoie, Bunzo, and Eggan 1992). The corpus includes seventeen dialogues between individual students and one of 3 expert human tutors, for a total of 313 turns (about 18 turns per dialogue), and 1333 clauses. The student solves an electronic troubleshooting problem interacting with the Sherlock system; then, Sherlock replays the student's solution step by step, schematically criticising each step. As Sherlock replays each step, the students can ask the human tutors for explanations. The student and tutor communicate in written form.

The Sherlock corpus had been previously annotated using RDA to study cue phrases generation (Moser and Moore 1996a; Di Eugenio, Moore, and Paolucci 1997). The research group which proposed RDA discusses the following reliability results (Moser and Moore 1996a). 25% of the corpus was doubly coded, and the κ coefficient of agreement was computed on segmentation in a stepwise fashion.[2] First, agreement at the highest level of segmentation was computed. After computing agreement at level 1, the coders resolved their disagreements, thus determining an agreed upon analysis at level 1. The coders then independently proceed to determine the subsegments at level 2, and so on. The deepest level of segmentation was level 5; the κ values were .90, .86, .83, 1, and 1 respectively (from level 1 to 5).

4.2 Our Methods

We annotated about half of the Sherlock corpus for anaphoric information, using a much simplified version of the annotation scheme developed by the GNOME project (Poesio 2000b). More specifically, we marked each NP in the corpus, specified its NP type (proper name, pronoun, the-np, indefinite NP, etc) and then we marked all 'direct' anaphors between these NPs (i.e., no bridges). This scheme has good results for agreement (Poesio 2000a) and has already been used for studying anaphoric accessibility (Poesio, Cheng, Henschel, Hitzeman, Kibble, and Stevenson 2000). We annotated a total of 1549 NPs, 507 of which were anaphoric.

One problem we had to address was that in the RDA annotation, only tutor turns had been annotated (because the students' questions are very short), but many of the antecedents of anaphoric references were discourse entities introduced in the preceding student turn asking the question. We followed Fox and made the first elements of adjacency pairs part of the accessibility space of anaphoric expressions in the second part. To do this, we enclosed each student turn in a special `student-turn` element, marked the NPs it contained, and made this turn a special focus space accessible from the entities in the tutor turn. We dealt with antecedents contained in 'tied' adjacency pairs (Fox 1987), i.e., in turns further away, by count-

[2] It is unknown to us whether κ was also computed on clusters, and on the specific informational relations used.

ing the anaphoric expressions whose antecedent was unaccessible for these reasons and factoring them out.

A second problem to solve was the fact that a large proportion of the anaphoric expressions whose antecedent is not on the stack are proper names. Because it can be argued that these expressions do not access the stack to find their antecedent, we also counted them separately.

We ran a script over the annotated corpus that simulates focus space construction under the several possible ways of mapping RDA structures into focus spaces that we considered, and attempts to find the antecedent for an anaphor in the focus space stack accessible to the anaphor according to each of these possibilities. The variants considered are:

1. **All:** Push a new focus space on the stack whenever a non-atomic RDA unit (both intentional segments and informational clusters) is encountered, and pop this focus space when the constituent ends. (E.g., in Figure 14.1, push a new focus space for all three constituents of the top segment: 1.1-1.2, 2.1-2.2, and 3.1-3.2.)
2. **Intentional Only / Imm Pop**: Only push a new focus space when an intentional segment is encountered; pop it as soon as the segment is completed. (E.g., in Figure 14.1, only push a new focus space for segment 2.1-2.2, and pop it as soon as that segment is completed. 1.1, 1.2, 3.1, and 3.2 are just added to the top focus space.)
3. **Intentional Only / Delay pop of cores**: Only push a new focus space when an intentional segment is encountered. Pop focus spaces introduced for contributor segments immediately; but only pop the focus space introduced for a core segment when popping the whole segment .
4. **Intentional Only / Partial delay pop of tribs** Like the previous version, but in addition keep a focus space introduced for a contributor on the stack as long as the segment in which it occurs is still on. (E.g., in Figure 14.1, do not pop the focus space for segment 2.1-2.2 before processing 3.1-3.2.)

5 Results

Table 14.1 illustrates the impact on anaphoric accessibility of the distinction between intentional and informational relations by showing the percentage of anaphoric antecedents which are on the stack according to the first two variants of the mapping algorithm. The line indicated as 'All' shows the results obtained by treating both informational and intentional relations as introducing new focus spaces: 'OK' indicates the number of anaphoric antecedents which are accessible, 'NO' indicates the number of antecedents which are not accessible, 'Out of AP' the cases in which the antecedent is not accessible because it's outside the current Adjacency Pair, and 'PN' the number of cases in which the antecedent is not accessible but the anaphoric expression is a proper name (which, presumably, can access it denotation through long term memory rather than through the stack). The table shows that separating intentional constituents (which introduce new focus spaces) from informational clusters (that don't) makes more antecedents accessible; the result is highly significant by the χ^2 Test ($\chi^2 = 29.47, p \leq 0.001$).

	OK	NO	Out of AP	PN
All:	199	74	63	158
Intentional only / Imm Pop:	280	20	63	131

Table 14.1: The informational / intentional distinction and accessibility

	OK	NO
Int. only / Imm pop of emb core and trib	280	20
Int. only / Delay pop of emb cores	287	16
Int. only / Delay pop of emb trib	310	8

Table 14.2: Effect of the differences in popping.

Table 14.2 shows the differences among the different ways of fixing the options left open by Moser and Moore (variants 2-4 above). The first line shows the results obtained if both embedded cores and embedded contributors were to push new focus spaces, popped as soon as these embedded constituents are completed; the second line the results if we keep embedded cores open until the end of the RDA-segment; the third line the results if we treat embedded contributors within an RDA-segment as remaining on the stack until the segment is closed off. In this table we have ignored both cases in which the antecedent is inaccessible but the anaphoric expression is a proper name, and the 63 cases in which the antecedent is inaccessible because it's not in the same adjacency pair (see discussion above). The differences are not so great in this case, but the correlation is still significant ($\chi^2 = 6.09, p \leq 0.05$) and in particular there is a clearly significant difference between the simplest possible treatment of embedded intentional relations, in which they are always popped, and the last model.

6 Discussion

As said above, we studied two separate issues: how best to use ideas from RDA to make G&S's theory more specific, and to evaluate G&S's proposals concerning anaphoric accessibility under this mapping.

6.1 Mapping RDA into Focus Spaces

The first goal of our work is to use RDA structures to gain a more detailed understanding of when focus spaces should be opened and closed. In this respect, our first result is a significantly better characterization of anaphoric accessibility if we assume that new focus spaces are only pushed on the stack when cores are recognized, as opposed to also being pushed when a purely informational structure is observed. This result is especially interesting when compared with Fox's proposal (discussed below). According to Fox, informational relations also affect accessibility. It is also interesting to contrast this result with the proposals of Veins Theory,

which also makes no distinction between informational and intentional relations.

The second interesting finding in this respect is that the best results concerning accessibility are obtained when embedded contributors, as well, are only popped when an RDA-segment is closed. This is not something that would be predicted on the basis of either Fox's work, or any obvious interpretation of the notion of RDA-segment; in fact, as we will see shortly, it's not immediately obvious how to account for this result in terms of Grosz and Sidner's theory, either.

6.2 An Evaluation of Grosz and Sidner's Theory

If we assume that contributors stay on the stack until the RDA-segment is completed, only 8 anaphoric antecedents are outside the currently open focus spaces. Of these, 5 are cases of definite descriptions that might be viewed as referring deictically to parts of the circuit, 1 is a cataphoric discourse deixis, one is a temporal deixis (*at this point*), and one a bit unclear. In other words, under the suggested rules for opening and closing focus spaces, virtually all anaphoric antecedents are accessible.

The question then is whether these rules are consistent with the G&S framework: the answer, we think, is that while the simplest treatments of contributors and cores probably are, the most complex ones probably aren't. In fact, we didn't even test what some people might think of as the most natural mapping - make the unembedded core of an RDA segment part of the focus space for that segment, but treat unembedded contributors as introducing subordinate focus spaces - since this would violate one of Moser and Moore's basic hypotheses (that only RDA-segments result in G&S-segments). The approach of leaving unembedded cores and contributors on the stack, but popping all embedded RDA-segments, can probably be still viewed as consistent with the theory;[3] but delaying the popping of embedded cores and contributors on the stack may be considered beyond its limits, whether we view the problem in algorithmic terms or in terms of intentional structure. Leaving these focus spaces open means that the attentional state cannot be properly seen as a stack anymore, but has to be seen as a list, since there is no guarantee that the focus space associated with the core will be the last element on the stack. Figure 14.3 shows an example in which it is necessary to leave the embedded contributor 24.13-24.14 on the stack, in order to solve *the "other" voltage* to *some other voltage*; at the end of the segment, this focus space has to be removed while leaving the focus space for the core on top of the stack. Furthermore, note that in the case of Figure 14.1 we need a 'discontinuous focus space' to make the embedded contributor accessible without eliminating the principle that new entities are always added to the focus space on top of the stack.

Looking at the problem from the point of view of intentional structure, these cases of accessibility can only be handled within the G&S framework by hypothe-

[3]This however means that the intentional relations contained in the segment - e.g., **convince** in Figure 14.1–should become part of the DSP for that segment, which again may or may not be consistent with G&S's ideas.

24.13a Since S52 puts a return (0 VDC) on it's outputs
24.13b when they are active,
24.14 the inactive state must be some other voltage.
24.15 So even though you may not know what the "other" voltage is,
24.16 you can test to ensure that
24.17a the active pins are 0 VDC
24.17b and all the inactive pins are not 0 VDC.

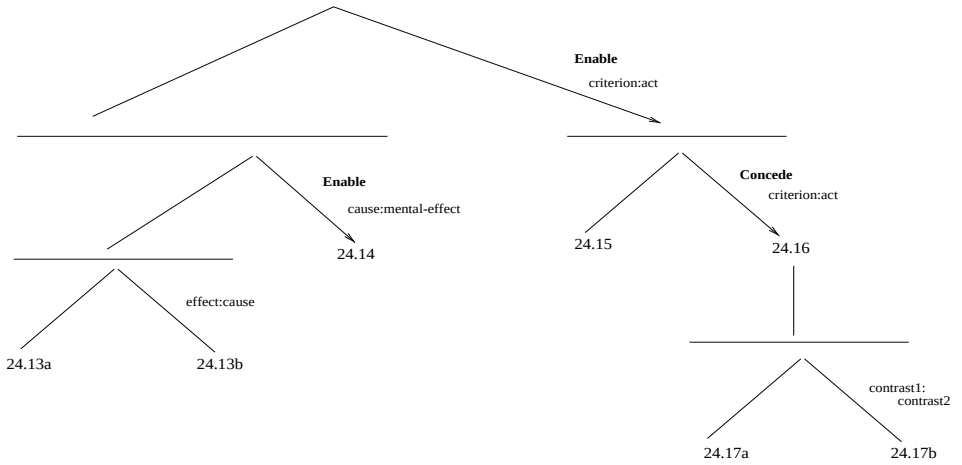


Figure 14.3: A contributor that precedes the core

sizing that the DSPs associated with embedded cores and contributors **satisfaction-precede** the core and the other contributors. More in general, one has to wonder whether Fox's (and before that, Reichman's) suggestion that the material introduced by 'active' propositions (i.e., the other constituents of a rhetorical scheme) remains accessible can really be formulated in terms of Grosz and Sidner's intentional structure.

6.3 Open Issues

Keeping more entities on the stack is obviously going to increase the likelihood that an antecedent will be found on the stack, but we have to be careful not to lose the crucial property of the attentional state, reducing search ambiguity (otherwise, we could just make every discourse entity accessible at all times). In order to properly compare models of anaphoric accessibility it is necessary to compute what we might call their 'perplexity'; unfortunately, in order to properly do this it is necessary to have an anaphoric resolution mechanism . We plan to do this evaluation next.

A major problem with any attempt at looking at the empirical import of dis-

course structure is that it's very hard to get researchers to agree on what the structure of a given discourse is. Work such as Marcu's has shown that in general we can expect people to roughly agree on the main segments of a discourse, but it is much harder to get agreement on the relative importance of the constituent segments, and even harder to get agreement on the particular relations. At the moment, the only way to achieve a reliable annotation is by extensive training and discussions, as done when constructing the Sherlock corpus. So, we do not expect everybody to agree on the particular analyses proposed here; we would want to point out however that these annotations do have a certain amount of objectivity in that they were not produced by us.

7 Related Work

7.1 Fox

(Fox), although only concerned with references to singular and human antecedents, is perhaps the most extensive study of the effects of discourse structure on anaphora in both spoken and written discourses. Fox uses different methods for analyzing the two genres: she uses concepts from Conversation Analysis, and in particular the notion of Adjacency Pair, for spoken conversations, and RST to analyze written texts. Her main proposal about written texts is as follows:

> A pronoun is used to refer to a person if there is a previous mention of that person in a proposition that is ACTIVE or CONTROLLING; otherwise a full NP is used.

(Where a proposition is ACTIVE if it's part of the same RST scheme as the proposition in which the pronoun occurs; whereas a proposition is CONTROLLING if it's part of a scheme which dominates the scheme in which the pronoun occurs.)

Fox's proposals concerning pronominalization apply less well to references to objects (and even in her corpus there are many references for which the hypothesis above would licence the use of a pronoun are actually realized by a definite NP, which she explains by arguing that the principle above is only one of many interacting principles that determine the realization of a NP); nevertheless, she makes a lot of compelling points about structure. In particular, she makes it very clear that active propositions should be accessible for as long as the scheme is open; and produces several examples showing that material introduced in active embedded nuclei is accessible. Fox didn't find references inside active embedded satellites (but then again none of these is made via a pronoun in our corpus). In addition, our study suggests that proper names behave differently from definite descriptions in that the former are much less sensitive to discourse structure than the latter, so the two classes should not be conflated like Fox does; and not separating informational relations from intentional ones restricts too much the range of accessible antecedents, even if it may be correct as far as pronominalization is concerned.

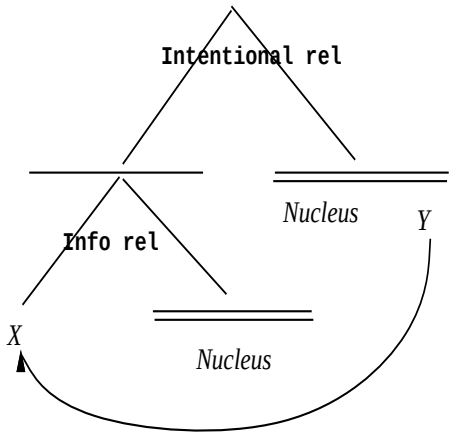


Figure 14.4: A tree distinguishing Veins Theory from the proposals discussed in this paper

7.2 Veins Theory

Veins Theory (VT) (Cristea, Ide, and Romary 1998; Cristea, Ide, Marcu, and Tablan 2000; Ide and Cristea 2000) is a recently proposed theory of the effect of discourse structure on anaphoric accessibility, which relies on RST for its definition of discourse structure, and whose predictions have been tested using an RST annotated corpus of newspaper texts (Cristea, Ide, Marcu, and Tablan 2000). The propositions accessible to an anaphoric expressions are computed by an algorithm that operates directly over an RST tree and involves two steps: a bottom up step in which the 'heads' of each node in the tree are computed (where the head of a non-terminal node is the concatenation of the heads of its nuclear daughters) followed by a top-down computation of the VEIN EXPRESSIONS. The crucial idea of VT is that material introduced in nuclear nodes 'percolates up' veins, where veins are paths in the tree all of whose arcs connect nuclear nodes; the antecedents introduced in any node along the vein are accessible from all the nodes of the subtree which has the top of the vein to which that node belongs as its root. The second idea of the theory is that antecedents introduced in a satellite node to the left of a nucleus remain accessible to all nodes controlled (in Fox's sense) by that nucleus.

In some respects, the proposal presented here can be viewed as a generalization of the proposals of VT: the material introduced in core constituents percolates up in a similar way, but we also allow antecedents introduced by embedded contributors to the *right* of the nucleus to be accessible as long as these contributors are active (in VT, only binary trees are considered).

The one point of contrast between the two theories is that in our proposal, we do not consider all nuclei, but only the nuclei of intentional relations. We have seen that treating informational relations as introducing focus spaces makes a big difference in terms of accessibility; the difference is significant even if we allow

these additional relations to remain open as long as the dominating relation is open. The crucial case distinguishing the two theories are trees with the structure in Fig 14.4: in our theory X would be available as an antecedent of Y, whereas in Veins Theory it wouldn't.

Bibliography

Carletta, J., A. Isard, S. Isard, J. Kowtko, G. Doherty-Sneddon, and A. H. Anderson (1997). The reliability of a dialogue structure coding scheme. *Computational Linguistics 23*(1), 13–32.

Cristea, D., N. Ide, D. Marcu, and V. Tablan (2000). Discourse structure and co-reference. In *Proc. of COLING*, pp. 208–214.

Cristea, D., N. Ide, and L. Romary (1998). Veins theory: A model of global discourse cohesion and coherence. In *Proc. of COLING*, pp. 281–285.

Di Eugenio, B., J. D. Moore, and M. Paolucci (1997). Learning features that predict cue usage. In *Proc. of the 35th ACL*, Madrid.

Fox, B. A. (1987). *Discourse Structure and Anaphora*. Cambridge.

Grosz, B. J. and C. L. Sidner (1986). Attention, intention, and the structure of discourse. *Computational Linguistics 12*(3), 175–204.

Ide, N. and D. Cristea (2000, October). A hierarchical account of referential accessibility. In *Proc. of ACL*, Hong Kong.

Lesgold, A., S. Lajoie, M. Bunzo, and G. Eggan (1992). SHERLOCK: A coached practice environment for an electronics troubleshooting job. In J. Larkin and R. Chabay (Eds.), *Computer assisted instruction and intelligent tutoring systems*, pp. 201–238. Hillsdale, NJ: Erlbaum.

Mann, W. C. and S. A. Thompson (1988). Rhetorical structure theory: Towards a functional theory of text organization. *Text 8*(3), 243–281.

Marcu, D. (1999, May). Instructions for manually annotating the discourse structures of texts. Unpublished manuscript, USC/ISI.

Moore, J. and M. Pollack (1992). A problem for RST: The need for multi-level discourse analysis. *Computational Linguistics 18*(4), 537–544.

Moser, M. and J. D. Moore (1996a). On the correlation of cues with discourse structure. Unpublished manuscript.

Moser, M. and J. D. Moore (1996b). Toward a synthesis of two accounts of discourse structure. *Computational Linguistics 22*(3), 409–419.

Moser, M., J. D. Moore, and E. Glendening (1996). Instructions for Coding Explanations: Identifying Segments, Relations and Minimal Units. Technical Report 96-17, University of Pittsburgh, Computer Science.

Nakatani, C. H. (1996). Discourse structural constraints on accent in narrative. In J. P. H. van Santen, R. W. Sproat, J. P. Olive, and J. Hirschberg (Eds.), *Progress in Speech Synthesis.* Springer Verlag.

Nakatani, C. H., B. J. Grosz, D. D. Ahn, and J. Hirschberg (1995). Instructions for annotating discourses. Technical Report TR-25-95, Harvard University Center for Research in Computing Technology.

Poesio, M. (2000a, May). Annotating a corpus to develop and evaluate discourse entity realization algorithms: issues and preliminary results. In *Proc. of the 2nd LREC*, Athens, pp. 211–218.

Poesio, M. (2000b, July). *The GNOME Annotation Scheme Manual*. University of Edinburgh, HCRC and Informatics. Available from `http://www.hcrc.ed.ac.uk/ ~ gnome`.

Poesio, M., H. Cheng, R. Henschel, J. M. Hitzeman, R. Kibble, and R. Stevenson (2000, October). Specifying the parameters of Centering Theory. In *Proc. of the 38th ACL*, Hong Kong.

Reichman, R. (1985). *Getting Computers to Talk Like You and Me*. Cambridge, MA: The MIT Press.

Topic Structure in Route Explanation Dialogues.

LAURENT PRÉVOT
IRIT- Université Paul Sabatier
118 Route de Narbonne
31062 Toulouse Cedex 04
France
prevot@irit.fr

ABSTRACT. This paper deals with the representation of dialogue, including the semantic aspects. We have been working on a real speech corpus of route explanation dialogues. The paper follows the tradition of dynamic semantics (Kamp and Reyle 1993), and we will use the SDRT of (Asher 1993) which proposes a treatment of the semantic-pragmatic interface. Dynamic semantic theories have traditionally focused on discourse issues, but a few studies have begun to look at dialogue within that framework in the past few years. (Asher and Lascarides 1998). To the rhetorical and intentional structure proposed by SDRT we add a "topic structure" closely related to the informational structure.

1 Introduction

We address here the problem of representing human dialogue[1]. We are interested in a specific kind of dialogue : *route explanation dialogue*[2]. We place ourselves in the tradition of dynamic semantics (Gronendjik and Stokhof 1991) (Kamp and Reyle 1993) (Asher 1993). These approaches, which have mainly focused on the problem of monologue, are now addressing the domain of dialogue (Ginzburg 1994) (Asher and Lascarides 1998). From another perspective, works also exist which enrich dialogue-act theories with semantic content, rhetorical structure and dialogue games (Poesio and Traum 1997).

Here, the chosen framework is SDRT (Segmented Discourse Representation Theory) (Asher 1993). Several works based on this theory have studied the interaction between semantic content and *Discourse Structure* (Asher, Aurnague, Bras, Sablayrolles, and Vieu 1995) (Asher, Busquets, and Vieu pear).

The notion of "topic" is used in this framework but it is not precisely defined. Recently we have proposed a treatment of dialogue global topic in route explanation dialogues (Prévot 2000). After the collection of a real corpus, our analysis has

[1] Dialogue is seen here as a discourse with two speakers.

[2] See section 2 for a more complete definition of this kind of dialogue.

Paper presented at the ESSLLI 2001 Workshop on
Information Structure, Discourse Structure and Discourse Semantics
Edited by Ivana Kruijff-Korbayová and Mark Steedman.

shown the necessity of taking *Information Structure* into account, as was predicted by (Asher and Txurruka 1995) (Engdahl, Larsson, and Ericsson 2000) (Engdahl 2000).

Now, in a prospective work, we want to study the interaction between *Information Structure* and *Discourse Structure* and between *Information Structure* and *Topic Global Structure*.

First, we briefly present the corpus that we have recorded and that we are studying (section 2). Then our analysis of some pieces of those dialogues will be presented in three points : first from the SDRT point of view, strictly speaking (section 3.1) ; then in the light of Txurruka's account of *Information Structure* and *Discourse Structure* (section 3.2); thirdly, taking global topic into account (section 3.3). Lastly, we propose a construction of a topic structure (section 4).

2 The Corpus

We introduce in this section the corpus we have recorded. At this stage, it is composed of 15 dialogues for a total time of 30 minutes. To ensure a large coverage of the kind of phenomena we plan to collect more dialogues. As we have said, it is a corpus of route explanation dialogues. This kind of dialogues involves two participants : a giver and a receiver. The giver explains to the receiver how to go from one location to an other one. We have chosen the domain of route description because a lot of work has already been done on it, especially in lexical semantics (motion verbs, spatio-temporal prepositions, ...) (Asher and Sablayrolles 1995) (Muller 1998). It give us a basis for the representation of the semantic content of utterances.

Many features are important in the complete definition of this dialogues (characteristics of places used, of the participants, ...)[3]. Here, for reasons we will explain below, we focus on two of these features : dialogue situation and participants' knowledge.

As far as the dialogue situation is concerned, we will focus on the following two questions :

- Are the speakers located in the "explanation place"?
- Are the speakers together, in a same place?

In our work, we want to focus on purely verbal linguistic phenomena. Therefore, we have to minimize the utterance context dependency (deictic uses, gestures, face expressions). Thus we have not placed the participants together in a same location

[3]See (Denis 1997) for a classification of route explanation.

and have recorded telephone conversations.

During a previous corpus collection we had noticed that if the receiver didn't know anything about the global location including the route (for example in a town) there was no exchange between participants. As we wanted precisely to study interactions between the participants mental states, we chose subjects who had some knowledge about the location.

The analysis of our corpus is directed along three lines :

1. Study the Discourse Structure / Information Structure interaction.
2. Within each structure, study the interaction of sentential and discursive level.
3. Study the Task Structure / Linguistic Structure interaction.

Below is a transcript of a representative example from our corpus.

A	1	A: Allons-y! donc tu pars d'où? *Let's go ! so where are you going to come from ?*
	2	B: Je pars donc de chez Laurent, c'est dans la grande rue St-Michel euh à peu près 200 mètres avant la place St-Michel. *I will come from Laurent's place, it is in St-Michel Street er... about 200m before St-Michel Plaza*
	3	A: 200 mètres avant la place St-Michel. Donc cté plutt vers le boulevard des Recollets, en partant.... *200m before St-Michel Plaza. Therefore, rather on the side of Recollets Boulevard, leaving from....*
	4	B: Ouais. *yeah.*
	5	A: D'accord. Donc pour venir chez Philippe c'est assez simple. Tu descends la grande rue St-Michel. Tu traverses le carrefour. *Ok. So to come to Philippe's, it's rather simple. You go down St-Michel street. You cross the junction.*
	6	B: Ouais. *yeah.*
	7	A: Tu... euh... Tu prends la rue Pharaon. *you ... er... you take Pharaon Street*
A	8	B: La rue Pharaon? *Pharaon Street ?*
	9	A: Oui. *Yes.*
	10	B: Je ne connais pas cette rue. *I don't know that street*

11 A: Tu ne connais pas la rue Pharaon.
You don't know that street.

12 B: C'est après le carrefour, euh?
It's after the junction, ..?

13 A: Dans le prolongement euh oui c'est après le carrefour.
continuing ... er yes it's after the junction.

14 B: Ouais.
yeah.

15 A: Dans le prolongement de la grande rue St-Michel Tu as la rue qui s'appelle je sais plus rue du Languedoc ou la rue Ozenne.
Continuing St-Michel street, you've got that street called i don't know , Languedoc Street or Ozenne Street

16 B: Ouais d'accord.
yeah, okay

17 A: Bon euh 100 mètres après le carrefour, le grand carrefour de St-Michel...euh... il y a une fourche, la rue du languedoc ou la rue Ozenne je sais plus comment elle s'appelle oblique vers la droite.
well er 100m after the junction, after the big junction at St-Michel er... there's a fork, Languedoc St or Ozenne St, i don't remember its name, it turns right.

18 B: Ouais.
yeah.

19 A: et tu as devant toi une petite rue à sens unique... qui s'appelle la rue Pharaon.
and you've got a small one-way street in front of you... called Pharaon Street

20 B: Ah ouais d'accord. OK. Je situe.
Oh, yeah, okay. I see.

21 A: Donc, tu parcours toute cette rue... euh... Tu vas arriver à la place des Carmes.
So, you take this street and go all the way, er... you will arrive at Carmes Place

22 B: muhm... muhm...
...

A 23 A: Arrivé à la place des Carmes euh... tu tournes à gauche dans la rue des Polinaires, c'est une rue qui débouche sur la place des Carmes... Arrivé au 27 de la rue des Polinaires... sur ta droite tu trouveras...euh le 27 où habite Philippe Muller.
When you are a the Carmes Place,... you turn left in Polinaires Street, it's a street that goes on Carmes Place. when you are at the 27 of Polinaires street, on your right you will see er... the 27 where PM lives.

24 B: OK. Donc, rue des Pharaons jusqu'à la rue des Carmes et après...
Ok. So, Pharaon Street up to Carmes Street and then...

25 A: jusqu'à la place des Carmes.
up to Carmes Place

26 B: Oui place des Carmes... et à gauche.
yes, Carmes Place,... and to the left

27 A: Alors la première à gauche, c'est pas la bonne quand tu débouches sur la place des Carmes.
then the first on the left is not the right one when you arrive at Carmes Place

28 B: Ouais.
Yeah.

29 A: la seconde à gauche c'est la bonne. C'est au bout de la place des carmes en fait. Est-ce que tu vois la rue des Filatiers?
The second one on the left is the right one. It's at the end of Carmes Place actually. Do you see Filatiers Street ?

30 B: Ouais.
yeah.

31 A: Eh bien tu vas jusqu'à la rue des Filatiers, tu t'engages pas dedans. Tu regardes à ta gauche et tu verras la rue des Polinaires.
well, you go to Filatiers Street, you don't take it. You look on your left and you will see Polinaires Street

32 B: D'accord.
Okay.

33 A: C'est bon?
All right ?

34 B: Ouais c'est bon.
yeah, all right.

3 Elements of dialogue analysis

3.1 Dialogue structures as Discourse Structures

We will now give an SDRT-oriented analysis of a piece of this dialogue, but first we will give a very rough outline of this theory. Asher's theory assumes that a dialogue is a kind of discourse involving two participants. SDRT also assumes, as RST (Mann and Thompson 1987), that a discourse can be seen as a set of segments linked with (rhetorical or intentional) relations. These relations can be hierarchical or not. Hierarchical relations between segments induce a tree structure which imposes constraints on the interpretation of current utterances (e.g. anaphora resolution can be limited to certain segments of the current interpretation). More generally, constraints based on the structure induced by the segmentation determine what is a coherent dialogue. New utterances will be attached to some segment incrementally within an already existing dialogue structure as they come, by taking

into account lexical semantics, world knowledge and semantic-pragmatic rules selecting an appropriate relation in context (see (Asher and Lascarides 1998)). These aspects are often taken for granted in theories taking as primitives propositions and speech acts related to these propositions. Our goal is to account for the construction of such structures within an incremental interpretation process of the dialogue seen as a public construction of a common ground between participants. We will make use of topics in that perspective, focusing here on questions and answers sequences. Our approach is thus rather "top-down" compared to (Kruijff-Korbayova and Kruijff 1997) who places her analysis in the same discursive perspective but closer to the sentence level. That being said, even though we use SDRT as our framework, we depart somewhat from some of its choices made in the context of dialogues e.g. in (Asher and Lascarides 1998). We will keep the following principles, while leaving aside the presently less stable aspects of the still in progress theory :

- the global representation of a dialogue is composed of a set of labelled speech acts (SA) and rhetorical relations between these occurrences of SA. A speech act will be of the form $\langle Speaker, Mode, Content \rangle$, where the mode can be interrogative, declarative or imperative. Content can be any part of the representation of the dialogue or a sub-dialogue. A basic semantic content will be a DRS, that is a set of linguistic referents and of conditions (predicates) on those referents. Speech acts, seen in other frameworks as having an intentional content, take on an intentional dimension only when linked together by relations bearing an intentional content.

- In order to deal with dialogue, SDRT had to be enriched with new relations, The list of relations needed by a dialogue version of SDRT is still very unstable in the prospective work of (Asher and Lascarides 1998). We have divided them into the following categories :

  — **discursive relations**[4]: *narration, elaboration, explanation, continuation, background, parallel, contrast, comment, result, consequence* ;

  — **dialogic relations**: we depart from the work of Asher by considering the following relations, dividing question elaborations into *precision request*[5] and *confirmation request*[6]. (At this time we don't make any differences between direct and indirect partial answer.[7]. *no-answer* (linking replies, such

[4]The set presented in (Asher, Busquets, and Vieu pear).

[5]Divided between Yes/No questions and WH-questions.

[6]This kind of request doesn't introduce any information, it only ask a confirmation of already introduced informations. See B_8.

[7]Probably these distinctions are necessary for the evaluation of the "answerness" of an utterance. But for the moment, within the application of our rules, we only consider that there are utterances which give an answer (direct, indirect or partial) and some which don't. In technical terms, the relation *answer* recovers *QAP* (Question Answer Pair), *IQAP* (Indirect Question Answer Pair), and *PQAP* (Partial Question Answer Pair) SDRT relations.

as, “I don’t know”), *acknowledge*, and *correction.* For instance $answer(\alpha, \beta)$ relates a turn β to its corresponding question α.

confirm-r — answer

B 7 A: Tu...sss...Tu prends la rue Pharaon. π_7
You ...sss... You take Pharaon street.

8 B: La rue Pharaon? π_8
Pharaon Street?

9 A: Oui. π_9
Yes.

It should be noted that intonation is crucial in this example. The interrogative mode, that we use for the inference of the *confirmation-request* relation, is given by an interrogative intonation. This kind of questions is precisely identified by Reis (cited by (Engdahl 2000)) as “echo deciding question”. Here if we don’t have the rising intonation, (8) could be an acknowledgement of (7).

3.2 Information Structure, Asher and Txurruka analysis

In this section we are going to present the augmented SDRT of (Asher and Txurruka 1995).

As was said before, dialogue analysis has to take into account *informational structure*. The question is how to explain why a single propositional content may have different realizations in dialogue (Lambrecht 1994) (Vallduví and Engdahl 1996). The partitioning of the content of a sentence into two parts –sentence topic and sentence focus– structures the sentence in terms of informativeness. The partition criteria varies among authors : given/in discussion, presupposed/asserted, ... [8] and these different definitions raise different denomination : *topic/comment, topic/focus, ground/focus.* In spite of these heterogeneity in the sentence topic definition, we will use the common notion of topic emphasis. This *Informational Structure* is assumed to be given by the syntactic and intonational analysis (Lambrecht 1994) (van Kuppevelt 1995).

Now we focus on Asher and Txurruka’s propositions. They study the interaction between Discourse Relations and Informational Structure [9] of utterances. They assume, as their main hypothesis, a double interaction between *Discourse*

[8] For a complete taxonomy of topic/focus partition, see (van Kuppevelt 1998).

[9] They speak of informational *partition* with the aim to underline the two separate parts of the sentence.

Structure and *Informational Structure*. The informational structure will be influenced by the previous piece of discourse, but the discourse relation which links two segments helps to disambiguate the informational structure of the second one.

They claim for a systematic study of context impact on the determining of informational structure.

Very roughly, we can say that in their theory :

- Discourse relations force constraints on *Informational Structure.* (DS $\rightarrow$ IS)
- The constraint satisfaction is an important clue for the establishment of the discourse relation. (IS $\rightarrow$ DS)

Now we are going to consider our example from Asher and Txurruka's point of view. We now look more precisely at (25) turn of (A) (a correction). The corrected utterance will be called the target (here 24). Asher and Txurruka say that a correction introduces an inconsistent proposition with some previous discourse proposition. In the example, there is no inconsistency in a strict sense. (Asher 1998) extends this definition and proposes a semantic for the *correction* relation. Here we have only studied one kind of correction. In the future we will have to look at the other ones.

¿From Asher and Txurruka's point of view the interaction between the discourse relation of *correction* and the *Information Structure* of (25) is fundamental for dialogue coherence. In particular, the distinction between the linguistic material shared by correction (25) and target (24) *(until the Carmes)* and contradicting information *(place/square)* is given by *Information structure*. The inconsistent information (related to a shared situation) will be in focus whereas shared information will appear as sentence topic element.

In our example, it is the intonation which stresses the focus. The *Information Structure* is crucial here; indeed if the stressed element was not only the word "place", the discourse relation would be different.

C 24 B: OK. Donc, rue des Pharaons jusqu'à la rue des Carmes et après...
ok so Pharaon street up to Carmes street and after...

25 A:...jusqu'à la PLACE des Carmes.
...up to Carmes SQUARE[10]

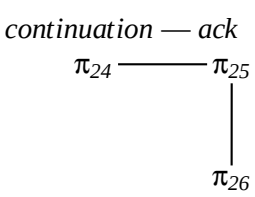
26 B: Oui, place des Carmes.
Yes, Carmes square.

The (24-25) pair satisfies the correction constraint[11]: The correction (25) and the target (24) differs in one point. The focus of the correction does not match

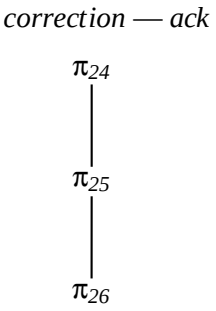
[10] We write in capitals the stressed part of utterances.

[11] See (Asher and Txurruka 1995), (Asher 1998) for a more complete account on *correction* in SDRT.

with the corresponding element in the target. Here the stressed element "place" (square) does not match the "rue" (street). If we didn't have this intonational stress on "place", (25) would be a continuation of the explanation of the route (figure 1.) instead of a correction (figure 2.). We notice that here, we don't have an inconsistency[12] between the correction (25) and the target (24). So, a strictly semantic analysis cannot provide us with the correction. Here *Information Structure* is not only a cue in the analysis but really the only way to obtain the right interpretation.



(*Fig*1.)



(*Fig*2.)

3.3 Global Topics

Dialogues of our corpus are all task-oriented. This kind of dialogue are studied by a lot of researchers, mainly in a human-machine interaction perspective. A widespread idea about these dialogues is that they obey a general purpose, which constrains the structure of the whole dialogue. (Grosz and Sidner 1986) speak about the *Discourse Purpose* distinguished from the *Discourse Segment Purpose*. (van Kuppevelt 1995) enriches this view with his definition of topic in term of *topic constituting questions.* He claims that in a Task Oriented Dialogue, as in a narrative discourse, "the utterances belonging to the main structure form together an answer to a single topic-constituting question". We follow this idea that there is a topic which arches over the other topics (which become sub-topics). In this section we study the effect of global topic, not only the overarching one but also the subtopics corresponding to sub-dialogues. From a general point of view, the dialogue topic is given by the context of dialogue (linguistic or not; verbalised or not).

In our corpus, the overarching topic is a *route explanation.* We have to represent the route components related to the explanation : there is a starting point, and

[12] In fact, if there was a Carmes street leading to a Carmes square, there would be no contradiction between (24) and (25). It is the intonation that points out the correction nature of (25). So Asher and Txurruka's proposition for the constraints on *correction* relation must be extended.

end point, and a prescription sequence (Denis 1997), introducing recursive sub-itinerary descriptions. The two participants are supposed to fill in an incomplete structure made of these elements. This presupposed structure gives us a canvas for the construction of the final, shared structure. We call these elements the **dialogue topics (DT)**. It will be used in a **top-down** manner. We give the following form :

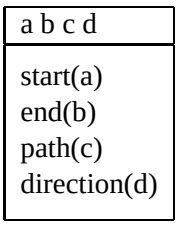


*Fig*3.

We have to explain what is a "path" in this context. There are several ways of defining paths in theories of space[13]. But in the route context, a path is usually seen as an action sequence. Each action of the sequence is a move between two places. We are interested in the linguistic aspect of routes, so we focus on the prescriptions of these actions. Participants (A and B) are going to establish a sequence of prescription of the actions that B has to do in order to follow the path (c) in a direction (d) from (a) to (b).

At the beginning of the dialogue (turns 1 to 5), there is a first sub-dialog to establish the starting point. It is a subtopic elaborating the overarching one. It is related to the referent "a". This sequence is a specification of "a" which fills out the canvas. We don't give here the dialogue analysis of this establishment of the starting point, but we give the representation for this dialogue segment with an approximate translation in English (*Fig*4.)[14][15].

[13] See (Muller 1998) for a presentation of these definitions.

[14] The "d1" direction is presupposed by the "avant" use. It is not the same with "d" (from overarching global topic) which comes from the preliminary context.

[15] For a more precise account of the semantic representation of a route, see (Krause 2000).

e f d1 g h
Laurent(e) chez(e,a) rue(f) nom(f,"St-michel") dans(a,f) place(g) nom(g,"St-michel") direction(d1) avant(a,g,d1,c) distance(a,g,200m) rue(h) nom(h,"Recollets") du-cote-vers(a,g)

e f d1 g h
Laurent(e) a is Laurent's place street(f) name(f,"St-michel") in/on(a,f) square(g) name(g,"St-michel") direction(d1) before(a,g,d1,c) distance(a,g,200m) street(h) name(h,"Recollets") near(a,g)

*Fig*4.

All the informations exchanged in this sub-dialog concerns the starting point. This fact is crucial information in the analysis of the dialogue.

The (8-20) sequence is the establishment of Pharaon street introduced by (7). We do not study here the construction of this establishment. At the end of the sequence, Pharaon street is grounded and the shared information is only the taking of the street by B (*Fig*5.) .

g e1
street(g) name(g,"Pharaon") take(B,g,e1)

*Fig*5.

We can see in the final route structure that some of the information introduced in the dialogue is preserved and some is left out. Typically the piece of information kept is what is shared (i.e. accepted) by both participants.

The building of this shared information is mainly guided by the rhetorical structure of the dialogue. However, trying to process it as a shared plan within the rhetorical representation of the dialogue, as in (Asher and Lascarides 1998), yields clumsy representations of the dynamics of the process itself (e.g. for correction, no-answers,...), and obscures the difference between the course of a dialogue and its underlying objectives. We therefore propose to create another structure, based on the notion of topic, for integrating this information.

As we have seen before, in task oriented dialogues (in which route explanations can be included) these "topics" comes from two sides : sentence level (see 3.2) and discourse level (see 3.3). We think that we must integrate these two sides of Information structure to contribute to the construction of a hierarchic topic structure.

4 Constructing topic and rhetorical structures in parallel

Our proposal is to enrich the rhetorical structure with a parallel structure keeping track of topic constructions. For this we consider topics as propositional contents linked by specific relations, called *topical relations*, either *topic sequence* or *topic elaboration.* In classical SDRT, these relations are somewhat related to the plan relations of SDRT (e.g. there is a plan-elaboration relation), but are not separated from rhetorical relations, while we think it is important to distinguish the course of a dialogue with the construction of a common-ground (which is often the sharing or communication of a plan in task-oriented dialogues). Besides, the topics of some parts of the dialogues are just these parts of the task to be achieved. In this sense we join the (Traum and Nakatani 1999) point of view on the multi-level approach of dialogue. Our global topic recovers the meso and macro levels of Traum and and Nakatani. Nonetheless topic relations intend to go beyond plans in task-oriented dialogue, by introducing a more general information structure.

An important notion for defining the topic structure is the resolvedness of a topic. The content of a resolved topic is shared between the participants: this kind of topic is in the common ground of the conversation. By contrast, the content of an unresolved topic is still under discussion, it can be pulled out of the topic structure. We will mark an unresolved topic with "*" and we remove "*" when the topic become resolved.

We now give a set of construction rules for the topic structure. Topic structure construction is done in parallel with semantic interpretation and establishment of rhetorical relations. For every rhetoric relation we have to trigger off a topic update rule.

- **Confirmation request(α,β) :** unresolves the α associated topic.
- **Precision request(α,β) :** introduces a new unresolved topic, attached to the topic associated to α with a Topic-Elaboration relation. The actual content of this topic will depend on the nature of the question (Y/N, WH).
- **Answer(α,β) :**
  — to a confirmation request : resolves the topic associated to α.
  — to a precision request : Resolves the topic associated to α and destroys unresolved sub-topic elaborating it.[16] Moreover we have to add new focused (information) from β in the DT.
- **No Answer :** This rhetoric relation has no effects on the topic structure.
- **Acknowledge(α,β) :** An acknowledgement resolves the topic associated to α.

[16] We destroy the unresolved topic because they are note established in the common ground and so they don't participate in the construction of the global topic structure

- **Continuation(α,β) :** A continuation resolves the topic associated to α and introduces an unresolved new topic attached to the topic associated to α by a topic sequence relation.
- **Correction(α,β) :** A correction replaces in the topic associated to α the corrected element by the focus of the correction (β)[17].

Considering that we have the *correction* relation in the rhetorical structure (inferred with an IS account); considering the preceding rules; we give the following construction for the (24-26) turns. This structure recovers the meso and macro levels of (Traum and Nakatani 1999). In our structure, the meso-level is down and the macro-level is up.

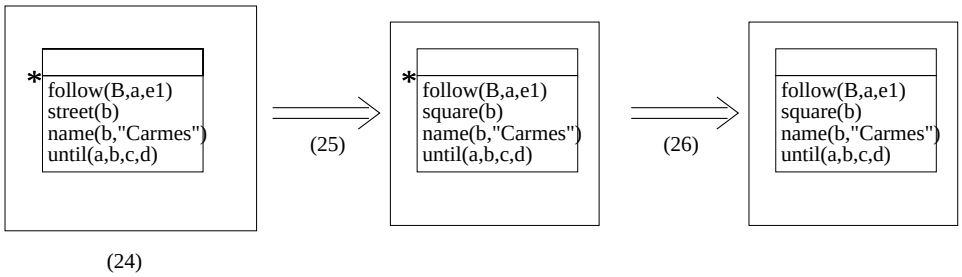


*Fig*6.

5 Conclusion

We have proposed here a way of structuring the analysis of a dialogue by taking into account works on topics in a semantic-pragmatic framework. We adopted a still-open theory, the Asher's SDRT, showing how to modify it to build in parallel a dialogue structure and a topic structure reflecting the construction of a common ground between two participants in a task-oriented dialogue. Our approach has remained global (ignoring the specific task of building a dialogue topic from a dialogue turn) and to a certain extent prospective. We claim that topics have to be considered in a systematic manner at any segment of a dialogue, and not only occasionally to solve specific local problems.

One of our preoccupations is to stay as close as possible to the real speech corpus. Indeed, if we want to take *Information Structure* into account we must be precise and integrate prosodic and intonational components. We don't want to avoid the problems related to the oral nature of the corpus and we think that on this subject, we have to do a lot of work. In the future we plan to modify our

[17]The *correction* rule that we give here concerns a specific use of correction. The are a lot of different types of corrections (*counterevidence, contradiction*) which are mentioned in (Asher 1998) (Prévot 2000).

transcripts. Instead of using the "written" punctuation, we are going to use specific speech notations, developed in speech corpus domain (Blanche-Benveniste. 1997) (Valli and Véronis 1999) for the speech phenomenon (pauses, overlaps, truncated words).

Many open problems remain to be fully investigated in that perspective. First sentence topic and sentence focus have to be more precisely linked to the discourse topic, to see how they contribute to the whole structure. Then, we have to see in which contexts the computation of discourse topics interact with the determination and acceptability of rhetorical relations. We have so far only looked at the cases where selecting these relations precedes constructing the topics.

Bibliography

Asher, N. (1993). *Reference to Abstract Objects in Discourse*. Kluwer.

Asher, N. (1998). Common ground, corrections and coordination. *submitted to Journal of Semantics.*

Asher, N., M. Aurnague, M. Bras, P. Sablayrolles, and L. Vieu (1995). De l'espace-temps dans l'analyse du discours. *Sémiotique* (9), 11–63.

Asher, N., J. Busquets, and L. Vieu (to appear). La SDRT : une approche de la cohérence du discours dans la tradition de la sémantique dynamique. *Verbum*.

Asher, N. and A. Lascarides (1998). Questions in dialogue. *Linguistics and Philosophy 21*, 237–309.

Asher, N. and P. Sablayrolles (1995). A typology and discourse semantics for motion verbs and spatial pps in french. *Journal of semantics 1*(12), 163–209.

Asher, N. and I. G. Txurruka (1995). Extending SDRT to integrate a theory of the informational partition. Technical report, ILCLI.

Blanche-Benveniste., C. (1997). *Approches de la langue parlée en français.* Gap: Ophrys.

Denis, M. (1997). The description of routes : a cognitive approach to the production of spatial discourse. *CPC 4*(16), 409–458.

Engdahl, E. (2000). The role of syntactic features in the analysis of dialogue. Technical report, INDI.

Engdahl, E., S. Larsson, and S. Ericsson (2000). Focus-ground articulation and parallelism in a dynamic model of dialogue. Technical report, TRINDI.

Ginzburg, J. (1994). An update semantics for dialogue. In *Proceedings of the international workshop on computational semantics*, Tilburg.

Gronendjik, J. and M. Stokhof (1991). Dynamic predicate logic. *Linguistics and Philosophy 14*, 39–100.

Grosz, B. and C. L. Sidner (1986, July-September). Attention, intentions, and the structure of discourse. *Computational Linguistics 12*(3), 175–204.

Kamp, H. and U. Reyle (1993). *From Discourse to Logic*. Kluwer Academic Publishers.

Krause, P. (2000). A cooperative conventionnal language game for identifying intended referent. Technical report, IMS.

Kruijff-Korbayova, I. and G. Kruijff (1997). Topic-focus articulation in DRT. In *Proceedings of the 11th Amsterdam Colloguium, December 17-20*.

Lambrecht, K. (1994). *Information structure and sentence form*. CUP.

Mann, W. and S. Thompson (1987). Rhetorical structure theory : a theory of text organization. Technical report, Information Science Institute.

Muller, P. (1998). *Éléments d'une théorie spatiale du mouvement pour la formalisation du raisonnement spatio-temporel de sens commun*. Ph. D. thesis, Université Paul Sabatier, Toulouse.

Poesio, M. and D. R. Traum (1997). Conversational actions and discourse situations. *Computaional Intelligence 13*(3), 309–347.

Prévot, L. (2000). Les dialogues d'orientation, une approche par la sémantique du discours. Master's thesis, Université Paul Sabatier, Toulouse.

Traum, D. and C. Nakatani (1999). A two-level approach to coding dialogue for discourse structure : activities of the 1998 dri working group on higher-level structures. In *Proceedings of the ACL'99 Workshop Towards Standards and Tools for Discourse Tagging*, pp. 101–108.

Vallduv, E. and E. Engdahl (1996). The linguistics realization of information packaging. *Linguistics 34*, 459–519.

Valli, A. and J. Véronis (1999). Étiquetage grammatical des corpus de parole: problèmes et perspectives. *Revue française de linguistique appliquée. 4*(2), 113–133.

van Kuppevelt, J. (1995). Discourse structure, topicality and questionning. *J. Linguistics 31*, 109–147.

van Kuppevelt, J. (1998). *Topic and comment.*, pp. 191–197. Elsevier Science, Oxford.

Presupposition or Abstract Object Anaphora?: Constraints on Choice of Factive Complements in Spoken Discourse

JENNIFER SPENADER
Computational Linguistics, Stockholm University
jennifer@ling.su.se

ABSTRACT. Using results from an empirical study of factives in spoken English, a comparison is made between factive verbs and their presupposed complements, both bound and accommodated, and their non-presupposing alternatives: abstract object anaphora that derive an antecedent from a discourse-given linguistic expression. The role of discourse structural constraints and grounding in choice of expression and the contexts in which they are used is discussed. Finally, the results are related to theoretical issues in presupposition theory having to do with the difference between anaphora and presuppositions, the ability of presuppositions to accommodate and the nature of accommodation.

Introduction

The aim of this work is to give a better description of the distribution and function in spoken dialogue of one expression type in one specific construction, complements of factive verbs that can be used to refer to an abstract object.

Factive verbs presuppose their sentential complements. For introducing the same information the non-presupposing alternative is to assert the information, making it part of the discourse record, and then refer to this information with an abstract object anaphoric expression such as a pronoun or definite noun phrase. An example may make things clearer. Here factive verbs are marked with **bold** and abstract object anaphora are written in CAPITALS.

Example (1)
a. Computational linguists are in demand.
b. Students apply for our program because they **know** that (computational linguists are in demand).
c. Students apply for our program because they **know** THIS.
d. Students apply for our program because they have **noticed** THIS TREND/THIS FACT.

Paper presented at the ESSLLI 2001 Workshop on
Information Structure, Discourse Structure and Discourse Semantics
Edited by Ivana Kruijff-Korbayová and Mark Steedman.

Sequence ab illustrates a bound presupposition usage with the presupposition in parentheses, ac illustrates abstract object anaphoric reference, as does ad, though here a full NP is used. It is possible to communicate the same information by using b alone, in which case the presupposition would be considered to be accommodated. In order to learn more about the use of these different expressions in natural spoken discourse, a corpus study was done.

1 Background

This first section explains the anaphoric theory of presupposition, and how it functions with the factive verbs studied here. The second section presents abstract objects, and how they can be referred to as well as proposed constraints on their usage based on discourse structure.

1.1 Presuppositions as Anaphora

The anaphoric theory of presupposition, developed by van der Sandt (1992), argues that presuppositions can be treated just as anaphora are treated in DRT (Kamp & Reyle 1993). Presupposition resolution involves examining the previous discourse context for an antecedent. If an antecedent is found then the presuppositions is bound to it. If an antecedent cannot be found then the presupposition is accommodated, and the ability to accommodate is what distinguishes presuppositions from other anaphoric expressions.

Factive verbs presuppose full propositions, and it is not immediately evident how binding should be identified here. The potential antecedent must be sufficiently similar to the presupposed proposition so that the latter can truly be considered to function as an antecedent. Examples where the potential antecedent is identical with the presupposed proposition, as in Example (1) ab, are probably rare. Speakers tend to vary their speech and avoid uninformative repetitions, and this means that there are potentially great difficulties in identifying propositional presuppositional binding.

The ability of presuppositions to accommodate is suggested in van der Sandt (1992) to be related to their greater semantic content and internal structure. Other anaphoric expressions are said not to be able to accommodate.

Despite the popularity of the anaphoric theory of presupposition, there has not been much work that studies presuppositional usage from the same perspectives that are considered central to characterizing the use of anaphoric expressions, such as the choice between a presupposition or a non-presupposing form, e.g., parallel to the choice between a pronoun or a name, or constraints on accessibility[1] of potential antecedents. Generally, in theories of anaphora accessibility, the number of potential antecedents in the previous context decreases as the semantic content of

[1]Here I mean accessibility in terms of salience, attentional state, etc., and not the structural accessibility of discourse referents within embedded DRSs as the term is used in DRT.

the anaphoric expression increases. This in turn allows a greater distance between anaphor and antecedent with semantically highly specified anaphora because the antecedent should be easy to identify even with a certain amount of distance. Because the presuppositions of factives are semantically very rich, we can begin with the hypothesis that they will allow longer distances between antecedent and presupposed propositions than abstract object anaphora and the linguistic expressions from which their antecedents are derived.

A puzzling question though, is why speakers would ever choose to use a factive verb with a presupposed sentential complement if it will get bound, i.e. why repeat an already given proposition? This would seem to break the informativity constraint on discourse. Repeatedly using definite NPs, for example, to refer to already introduced entities seems more natural because we often need to refer to the same discourse individuals or objects, in order to describe them in new situations and relationships. But propositions would seem to be unnecessary to repeat.

However, Walker (1996) has shown that informationally redundant utterances, or IRUs, are not at all infrequent in spoken discourse, citing that they made up about 12 % of the utterances in her corpus study of collaborative task dialogues. She gives three categories of communicative functions based on the corpus study, these were 1. *Attitude* : which seems to be the use of an IRU for explicitly grounding, 2. *Attention* : use of an IRU to make a propositions salient, or 3. *Consequence* : an IRU is used "to augment the evidence supporting beliefs that certain inferences are licensed" (Walker, 1996, p. 5).

Can bound propositional presuppositions be considered to be informationally redundant utterances? I think their status is somewhat different. Examine the formal definition (taken from Walker (1996), p. 5): An utterance u_i is INFORMATIONALLY REDUNDANT in a discourse situation **S**

> (i) if u_i expresses a proposition p_i, and another utterance u_j that entails p_i has already been said in S.
> (ii) if u_i expresses a proposition p_i and another utterance u_j that presupposes or implicate p_i has already been said in S.

Presuppositions differ from IRUs in two important ways. The first is that they have been presupposed. The speaker is therefore explicitly signalling that the information is known, given and should be accepted as backgrounded by the hearer. The second difference is that presupposed propositions occur in utterances that also have an assertional component, in which case the utterance itself is not informationally redundant, but the presupposed part is.

Still, even if the use of given or bound presuppositions is distinct from the use of IRUs, the functions of propositional presuppositions that are bound may be similar. Hopefully empirical study can answer some of these questions.

1.2 Abstract Object Anaphora and Discourse Structure

The type of higher order entity anaphora being referred to here as abstract object anaphora has been called many different things in the literature, e.g. discourse deixis (Webber 1991, Eckert & Strube, 2001), situation anaphora (Fraurud, 1992), and finally, as abstract object anaphora, by Asher (1993).

All these terms are used to refer to anaphoric reference to situations, factualities, eventualities, propositions, speech acts, as well as more deictic functions referring to the reference of a discourse segment, though many of these distinctive uses are often difficult to distinguish in practice. We can often find examples of different aspects of the same antecedent being referred to by the same type of anaphoric expression, and where the actual type of the referent is coded in the predication of the sentence in which the anaphoric expression is used. This observation is made by Asher(1993), Dahl & Hellman (1995) and is an important part of Eckert & Strube (2001). The type of entity also may be indicated by the NP used, as in Example (1) `d.` The continuation of Example (1) below illustrates some other possibilities. The sequence `ae` indicates that we should consider the type of referent a proposition, because usually only propositions can be considered true or false. Continuing with the sequence `af` means that the type will be a situation, as coded in the noun phrase. Eckert & Strube (2001) use this information to distinguish between individual and abstract object anaphoric reference which then helps guide their resolution algorithm.

Example (1) `continued`

```
e.  Everyone at the university knows that THIS is true.
f.  THIS SITUATION has led to an alarmingly high drop-out rate
as students leave school to take industry jobs.
```

Dahl & Hellman (1995) discuss abstract object reference as a type of anaphor that instigates a process of reference-creation, or reference coercion; that is, using an abstract object anaphor signals to the interpreter to look for established information of the appropriate type in the discourse record and create a discourse referent that can function as an antecedent from this linguistic information. Dahl & Hellman (1995) further list three referent creating operations, 1) Summation and Complex Creation, 2) 'Type-coercion' and 3) Abstraction and Substitution. The process of referent-creation has been compared to that of accommodation by Eckert & Strube (2001), though it is unclear if this should be considered accommodation in the same sense as it is used within the anaphoric theory of presupposition. For presuppositional accommodation, new information is added to the discourse record, but in referent-coercion, already given information is considered in a new way.[2]

[2]Some researchers have argued that there is only accommodation when the information is new, and information that is derivable from an already given representation is technically not new. Here I think it is unclear what the status of information that is available in the discourse record in different forms actually is. Developing a distinction similar to that in computer science between Information

Abstract objects have been shown to be severely constrained in their ability to be accessed by anaphoric means, and these limitations seem to be due to discourse structural constraints. For example, Webber (1991) has argued that abstract object anaphora can only access discourse segments on the right frontier of the discourse structure for referents; discourse segments on the right frontier are also those in focus, or salient.

Fraurud (1992) has argued that Webber's description of constraints on accessible antecedents is for the most part correct, but that in addition to unfocussed discourse segments, propositions that are subordinate to a main proposition also need to be excluded from supplying an antecedent for abstract object anaphora. She gives examples that show that the simple tree structures used by Webber to represent discourse structure don't correctly capture the full range of distinctions that will be relevant because the right frontier constraint doesn't give a means by which to factor out subordinated or modifying propositions that are conjoined with a main proposition as part of the right frontier. Making a distinction between main and subordinating propositions would solve this problem, though a different discourse structure with different accessibility rules would be necessary.

As this earlier work indicates, discourse structure is generally agreed to constrain accessibility for abstract anaphoric reference. In order to be able to discuss possible discourse structural constraints on accessibility, it is necessary to define what units of discourse structure will be considered relevant, i.e. the discourse segmentation. So far, most work on discourse structure (including that mentioned above) has focussed on written discourse, and work on spoken discourse has looked mainly at utterances and how they form adjacency pairs, only sometimes considering a more hierarchical structure. It is not clear what levels of structure are present in spoken dialogue, nor what units will be relevant to the discourse participants and discourse interpretation. Generally, the levels that have been discussed in the literature include a speech act or discourse move level. These are usually then discussed as adjacency pairs or as instances of dialogue games, e.g. as in Carletta et al.(1997), or called discourse units. Finally, a higher level is sometimes postulated that in Carletta et al.'s coding scheme has to do with domain specific goals in the discourse. Which, if any of these levels, is relevant for characterizing anaphoric (and presuppositional) accessibility is still unclear.

Eckert & Strube (2001) is therefore particularly relevant to the work here in that they have done empirical work on the special problems of anaphoric resolution in spoken discourse. Their work is a corpus study of the SWITCHBOARD corpus, a spoken corpus of telephone conversations between two participants who were unaquainted with each other before their conversation.

Eckert & Strube use a very simplified discourse structure but seem to get reliable results in using this simplified structure to delimit potential antecedents for

Extraction and Data Mining, i.e. between retrieving information you knew existed and gleaning new information that you weren't aware of from an existing knowledge representation, might be helpfully related to accommodation, and discourse interpretation in general.

personal pronouns as well as for abstract anaphoric pronominal reference. They analyzed each utterance as being a dialogue act of one of three types based on the top dialogue moves in Carletta et al.(1998): Initiation, **I,** Acknowledgment, **A** and **A/I** for those utterances which served both an initiating and an acknowledging function. These dialogue acts are then paired into what are called synchronizing units (**SU'**s). Certain dialogue acts don't require acknowledgement, and these may be represented by as single **I** and still be interpreted as an **SU**. Central to their work is the idea that grounding constrains the accessibility of antecedents. Grounding is the process by which discourse participants signal that information introduced in the discourse is part of the common ground (Traum 1994), by some sort of acknowledgment. **A** 's (Acknowledgments) are grounding acts. Discourse referents or information introduced in utterances that are not acknowledged, that is, not grounded, are not available for anaphoric reference. Note however that while acknowledgement is a signal of grounding, lack of acknowledgment is not necessarily as sign of lack of grounding, and the lack of a protest, or allowing the speaker to continue to have the floor, can be considered a type of implicit grounding. Eckert & Strube (2001) also give an example where a clear indication that a Speaker's utterance was not grounded by the other participant means that the discourse segment and referents within that segment are not added to the common ground, and therefore are not available for anaphoric reference of any kind. In this way they define what is salient or in focus as that which is in common ground, and the need to clearly identify grounding is also reflected in how they choose to code discourse structure.

2 Empirical Data and Method

Examples of factive verbs in context were excerpted from the London-Lund Corpus of Spoken English (LLC)[3]. 50 multi-speaker dialogues were used, which contained roughly 233,000 words. The following factive verbs were excerpted:

> **Factive verbs: subject complements** *count, make sense, suffice, amuse, bother, matter*
> **Factive verbs: object complements** *discover, find out, know*[4]*, notice, realize, regret, resent, see*

Examples without a complement, or with an NP-object, non-abstract object complement, were discarded. For each example, the relation between the potential antecedent, if there was one,[5] and discourse structure was examined, noting speaker

[3]Information on obtaining this corpus can be be found on the ICAME website at http://www.hd.uib.no/icame.html

[4]"know" with a sentential complement not marked by *that* is not included in the analysis due to the great number of false hits found because the corpus is untagged, meaning that these examples must gone through by hand. The author plans to do this at a later date.

[5]Presupposed propositions without antecedents are accommodated for that very reason.

TYPE	TOTAL	ACCOM PRESUPP	BOUND PRESUPP	ABS-OBJ ANA
object comp	75	48	6	20
subject comp	13	0	0	14
TOTALS	88	48	6	34

Table 16.1: Type of Abstract Object Complement

shifts between antecedent and anaphor, distance, in utterances, and any other relevant aspects. The analysis was done from the perspective of the annotator.

The dialogues studied seemed to be more complicated than those studied by Eckert & Strube in that many were between more than two participants. Because they take place in person so turn-taking is less precise and there are many cases of overlapping speech. Additionally, the participants often know each other well, which also seems to support interruptions and overlaps.

To adapt Eckert & Strube's coding of discourse structure to dialogues between more than two participants, the following guidelines were used: An utterance may be grounded by more than one speaker and this was often the case - so a synchronizing unit (**SU**) can be made up of a sequence of multiple **A**'s as long as they are acknowledging the same **I.** Sometimes one speaker will clearly acknowledge another speaker while after a third speaker has contributed an **I,** which means that **SU**'s must be able to overlap (cross structures) (e.g. we can have Speaker A: I_1 , Speaker B: I_2 , Speaker C: A_1 , Speaker A or C: A_2). Segmentation was done making each new turn a new utterance and splitting a turn into more than once act if part of the turn clearly only has an acknowledgement function. Of course- if an utterance with an abstract anaphoric reference overlaps with another utterance this cannot be a potential antecedent as it is not yet part of the discourse, and for this reason overlaps must be taken into consideration.

3 Results

Table 1 presents the results. By far the most frequent usage was a presupposed proposition that had to be accommodated (48 examples).[6] Looking at an example wouldn't really contribute to the discussion here so because of space limitations no examples will be given here.

The next largest group found were abstract object anaphoric reference. Here particular attention was paid to the discourse structure and whether or not utterances were grounded, using the modified coding system based on the one given in Eckert & Strube (2001) and described above.

[6]Note that accommodation can occur globally (39 examples), or intermediately and locally (9 examples) according to van der Sandt's (1992) theory.

Example (2) Same speaker, previous discourse segment, chain of references

Speaker A: So that it's the faculty of arts,	I	↖	
or the faculty of economics or both that'll			**SU**
be putting him forward (1)			
Speaker B: Mmm. (2)	A	↙	
But they can put it forward for any title			
that they like apparently.(3)	I		
I didn't realize THIS. (4)	I	↖	**SU**
1 to 2 sylls. so this		↙	
Speaker A: No, I didn't know THAT. (5)	A		

Here, utterances were considered to be discourse segments. In the example above it seems clear from the speaker's use of "No", in "No, I didn't know that," that he is referring to B's last utterance, and not to the entire informational content of what B has just said. THIS (4) refers to an abstract object derived from the linguistic expression in discourse segment (3). THAT in (5) refers either to the same thing as THIS in (4), making this a chain of abstract reference.

Example (3)[7] Overlapping speech, source of abstract object could be synthesis of several utterances made by different speakers

Speaker C: University of the Air (1)	I		
Speaker D: *that would be S* (2)	n.g.	↖	
Speaker C: *are doing a series* on various	I		
sorts of +communication which struck me			**SU**
immediately as **disparates** (3)		↙	
Speaker B: +disparates, surely+ ? (4)	A		
Speaker D: **<but but>** but it would be,			
um it would be non-surreptitious	I		
wouldn't it ? (5)		↖	
Speaker C: *presumably* (6)	A		
Speaker B: *yes -*, it would be			**SU**
+non-surreptitious+ (7)	A	↙	
Speaker A: Yes, (8)	A		
but THAT wouldn't matter. (9)	A/I		

In the above example, four different speakers take part, and identifying synchronizing units was not totally straightforward. Eckert & Strube's coding system is expanded to allow different speakers to each ground the same utterance (here (6), (7), and (8), grounding (5)) and still consider it to be one SU. Utterance (2) is labelled as "n.g." for "not grounded" and this discourse segment does not introduce any referents available for anaphoric reference. Note that it overlaps with part of

[7]Note that the diacritic marks that encapsulate parts of the utterances mark where the speech of speaker's overlapped. Here, for example,"that would be S" and "are doing a series" overlapped.

(3). Here I have split Speaker A's utterance into two discourse segments, (8) and (9), because (8) seems to solely have a grounding function, whereas (9) seems to be informative, though it is not clear if this is the correct segmentation. Speaker A's abstract anaphoric reference in (9) seems to refer to the immediately proceeding grounded SU (5-8), or could perhaps be considered to referring to only the grounded portion, (6-8). It is not clear what analysis would be correct, though in both cases the abstract object anaphor derives its antecedent from an immediately preceding discourse segment, within the same SU, and it is impossible that the anaphora could be referring to the previous SU (1-4).

The forms of abstract object anaphora used may also be of interest. There were 8 cases of *it*, 2 cases of *this*, 18 cases of *that* and 3 cases of *zero anaphora* and 2 cases of *definite noun phrases*. For almost all examples a linguistic expression that could be a source for an abstract object could be found in the previous SU, or in the same SU, though there were a few exceptions. In 19 cases the same speaker who uttered the abstract object anaphor also had said the linguistic expression from which its antecedent can be derived; in 12 remaining cases the speakers were different, and in 3 cases it is impossible to pinpoint exactly what utterance(s) provided the antecedent are (cf. Eckert & Strube who also found a great number of vague or difficult to identify abstract anaphoric occurrences), though it is clear that it is part of the previous context (e.g. previous or same SU), it is often (as illustrated in example (3)) a question of determining how much of the previous context is intended as the antecedent. It also seemed that the simple analysis of discourse structure modified from Eckert & Strube was helpful and adequate to organize and understand the data.

The next example illustrates something that could potentially be considered presuppositional binding. In determining presuppositional binding the entire discourse record up until the use of the factive was taken into consideration. The criteria used to determine if binding was a potential analysis was the authors own intuitions as to whether the presupposed information was new, or had already been given in some form in the discourse.

Example (4) Presupposed propositions - bound
Speaker A: It was lethal to expectant mothers
with small children. (1) (38 intervening lines of text).
Speaker A: After all, I mean you can't go down and shop if you
KNOW that you're going to knock out an expectant mother ...
it was some violent streptococcus that he'd got (2)

Here the presupposition is that "you" (or anyone, in a generic sense) would knock out an expectant mother if having been infected with the streptococcus referred to by "it" in utterance (1). Here it is arguable that the information presupposed in (2) is not really new, in that utterance (1) refers to the same situation, though in a different way and with some other conclusions thrown in. Note that there really seems to be no point in doing a discourse segment analysis here, because the

intervening 38 lines means that the presupposed proposition is very far removed from utterance (1). Another one of the examples of potential binding also had a 900 line gap between potential antecedent and presupposition.

4 Discussion

To summarize the results: Factive verbs, for the uses studied here, overwhelmingly tend to occur with full propositional presuppositions that need to be accommodated, i.e. they presuppose discourse new information. Factive verbs also appear with abstract anaphoric complements that then refer to abstract objects. The use of factive verbs with presupposed information that is already part of the discourse record, e.g. presuppositional binding, is minimal (only 6 examples!).

The binding examples are most interesting to discuss first as one interesting research question was why speakers would choose to use full presuppositions when the presupposed information can be considered to be already part of the discourse context. The answer seems to be that they seldom do, but when they do, the function of the presupposition in the discourse seems tentatively to be one of the following:

(1) The bound presupposition has a *summation* function Full sentential complements are used to explicitly express as a whole an idea that was present only in bits and pieces in the earlier discourse, perhaps even contributed by one or more speakers. This usage seems to relate to the referent-creating operation of *Summation* proposed in Dahl & Hellman (1995), and may also be the effect of the multi-speaker discourse setting in that information is being contributed from so many different directions.

(2) The bound presupposition states some kind of conclusion that is deductable or inferrable from the discourse record Fully inferrable information is considered to be known information. If a strict definition of accommodation that limits its application to new material is used, then the presupposed proposition is known and must bound because it must be true in each attentive speakers information state. However, it is being expressed in the discourse explicitly for the first time here. All conclusions may or may not have been realized by *all* discourse participants. Multi-speaker dialogue may need to use presuppositions this way because establishing information as mutually known is a more complex task when several discourse participants are involved than when only two participants are involved. In fact, it would be strange if we would not need to explicitly conclude things on occasion, as a form of grounding.

(3) The bound presupposition has another pragmatic function than the original, asserted usage. There were two examples where this could be considered to have been the case. In one example the presupposed information was a near repetition of the other speakers immediately proceeding statement and had the function of showing agreement. In the other example, the repetition was used to relate what another speaker had said, so should perhaps not be considered as

very good example of a presupposition. [8]

(4) The linguistic expression from which the abstract object could be derived is in a discourse segment that is no-longer accessible for reference. This was my original hypothesis about when speaker would chose to use a bound presupposition, given the great number of constraints on abstract object anaphoric reference. Indeed, for five of the six examples found here, it would be awkward, if not impossible, to refer to the same information with an abstract object anaphor.

Two of the communicative functions of IRUs identified by Walker (1996) (see subsection 1.1) were quite similar to the usages identified above. Category (1) seems very close to the *Consequence* communicative function and *Attitude* seems similar to category (2), a kind of grounding function. *Attention*, would have most likely fallen under category (4), but there were no examples of category (4) found.

In light of the infrequent cases of potential binding, and the dubious ways in which the potential antecedents often were related to the presupposed proposition, is it defensible to consider the cases found as actual binding?

If we take the idea of presuppositions as anaphora seriously, then perhaps we should also seriously entertain the idea that binding presupposed propositional information may also be constrained by the same discourse structural constraints as abstract object anaphora. Until now I have been working under the naive assumption that presupposed propositions, because of their greater semantic content, will freely allow reference from almost any position in the discourse.[9] Usually the discourse structural constraint on accessibility of anaphora can be tested by creating example sentences and considering what interpretation the anaphoric expression will get. Unfortunately we cannot test for constraints on the accessibility for presupposed propositions in this way because we can't really distinguish between our own processes of binding or accommodation.

While we can't really test accessibility in the way we can for abstract anaphoric reference, we can however make conclusions about whether or not the presupposed information has entered into the common ground of the discourse participants, the most important criteria for anaphoric reference accessibility in Eckert & Strube's study. Indeed, the explanations for the use of a bound presupposition given in (1), and (2) above could both be described as cases where the speaker was unsure if the presupposed information had entered into the common ground, and that his/her usage of a presupposition with this information was in a sense a way to ground it. Presupposing as an information presentation device may be quite apt in that the information presented is uncontroversial information that may already be known to some discourse participants.

The analysis of the examples of binding given above can then be revised. Parallel to constraints on abstract object reference, we could consider that these exam-

[8] In this example the speaker first described a situation with a proposition, and then related a story about another person's reaction to the same situation. When reporting on the other person's reaction, she used the same proposition, but as the complement of a factive verb

[9] Excluding of course structurally inaccessible positions due to context created by logical operators, such as modal contexts, modal subordinating contexts, and belief contexts

ples represent information that is not considered by the participants to have truly been established in the common ground, and therefore is unavailable for binding. We could also disallow binding for some examples based solely on the idea the great distances between potential antecedent and presupposition make reference unacceptable. In an interesting twist, taking the analysis of presuppositions as anaphora full out, and applying the same types of constraints to them, actually gives us an analysis where presuppositions and anaphoric alternatives are used in functionally complementary distribution, but not as first thought. Presuppositions present new information or ground unclear information, abstract object anaphora in the same context (here, as the complement of factives) refer to already given information.

If the cases earlier identified as binding are now categorized as accommodation, then all examples found in the data of presupposed propositions were accommodated. What does this tell us about the use of factives in particular, and the nature of presupposition accommodation in general?

Remember, van der Sandt's (1992) theory tells us that presuppositions should theoretically be able to both bind and accommodate, and that the ability to accommodate has something to do with the semantic content of the presupposition. The conclusions reached here suggest that in practice not all triggers are used to do both. And for factives particularly, binding does not seem to be a normal usage. Factive verbs with a presupposed complement are then used primarily to introduce discourse new information.

In terms of how much semantic information is being presupposed, factive verbs, because they presuppose entire propositions, must be one of the richest triggers. Because anaphoric expressions and presuppositions with lesser semantic content are considered not to be able to accommodate, or to accommodate badly, it is tempting to interpret the results here as evidence that the reverse is also true: Presupposition triggers with a rich semantic content can not only accommodate when necessary, but this is their preferred usage.

All this has relevance to our view of the nature of accommodation. There are two competing views of accommodation: that it is a repair strategy, and something that should be avoided and the view that it is a normal method of communication. Most work on presuppositions seems to have taken the former view. Of course, in presupposition resolution binding must be preferred over accommodation, hence the preferences in van der Sandt (1992) and in Blutner (2000)'s constraint AVOID ACCOMMODATION in his bi-directional OT treatment of presupposition. Zeevat (2001), building on Blutner (2000) has argued that it is not the richness of semantic content that determines what expressions can accommodate, but the availability of non-presupposing alternatives. If an alternative is available, AVOID ACCOMMODATION will inhibit, or even block the speaker's use of a trigger to be accommodated because there is a simpler alternative. The speaker should choose to use the simpler alternative, and non-presupposing alternatives are by their very nature simpler. For factives, there *is* an expression alternative: assert and then refer with an abstract object anaphora, but this alternative has been shown to not always be

available, and therefore it may be incorrect to consider it a true blocking alternative.

This view of accommodation is still one that considers it to be a repair strategy by the hearer when a non-presupposing alternative was available to the speaker. Accommodation is considered to be a costly method of communication that could lead to misinterpretation, and something that both speakers and hearers should avoid.

Again, I think the results here suggest a different view. Accommodation is both an exploitable[10] communication strategy and a repair strategy, but it depends on the trigger involved and the context. Sometimes asking your listener to accommodate is the best means by which to express your idea and when the semantic content of the trigger is rich enough that accommodation can proceed without the danger of misinterpretation on the part of the hearer, then it *is* the most optimal way to communicate, and that is why factives with presupposed, accommodated complements are the norm.

So the presupposition triggers that are best equipped for accommodation, e.g. those with rich semantic content and structure, will also be exploited by speakers to the fullest as this will be the most effective and economical way to introduce information. In fact, more effective than non-presupposing alternatives in certain cases. Those triggers that are worst equipped for accommodation, e.g. those whose meaning is underspecified to the degree that accommodation is a strain on the listener, and a real potential source of confusion, e.g. most pronominal anaphora, will not be exploited by speakers, though these will be able to be interpreted by listeners by accommodation when necessary - and then it is *being used* as a repair strategy. Accommodation should be considered to be an available option for hearers both for anaphora and presuppositions, but speakers will tend to limit their exploitation of the hearer's ability to accommodate to semantically rich anaphoric and presuppositional expressions.

5 Future research

Differences between other presuppositions and their non-presupposing alternatives should be looked at more carefully. The results here should also be compared with written discourse, in particular because written discourse seems to have a more hierarchical discourse structure than dialogue, and it would be interesting to see how this would affect the choice to use presupposed complements.

Acknowledgment

I would like to thank Bart Geurts, Emiel Krahmer and Sofia Gustafson-Capková for discussing this work with me. Thanks also to the anonymous reviewers for

[10] Actually Lewis's (1979) original paper on accommodation seems to characterize it as a type of 'exploitation' of the hearer's ability to make certain inferences by the speaker, hence the master-slave analogy he uses.

helpful comments. All errors are my own.

Bibliography

Asher, Nicholas (1993). Reference to Abstract Objects in Discourse, Kluwer Academic Publishers.

Blutner, Reinhard (2000). Some Aspect of Optimality in Natural Language Interpretation, In Optimality Workshop, Utrecht, January 2000

Eckert, Miriam & Strube, Michael (2001). Dialogue Acts, Synchronising Units and Anaphora Resolution, to Appear in Journal of Semantics.

Carletta, Jean, Isard, Amy, Isard, Stephan, Kowtoko, Jacqueline, Doherty-Sneddon & Anne Anderson (1997). The reliability of a dialogue structure coding scheme. Computational Linguistics, 23: 13-32

Dahl, Östen, & Hellman, Christina. (1995). What happens when we use an anaphor?. Presentation at the XVth Scandinavian Conference of Linguistics, Oslo, Norway.

Fraurud, Kari (1992). Situation Reference, (What does 'it' refer to?). GAP Working Paper No. 24. Hamburg: Fachbereich Informatik, Universität Hamburg.

Kamp, Hans and Reyle, Uwe (1993). From Discourse to Logic, Dordrecht, Kluwer Academic Publishers

Lewis, David (1976). Scorekeeping in a Language Game, Journal of Philosophical Language, vol 8.,D. Reidel Publishing Company.

van der Sandt, Rob (1992). Presupposition Projection as Anaphora Resolution, Journal of Semantics, Vol 9, pp 333-377.

Traum, David (1994). A Computational Theory of Grounding in Natural Language Conversation, Ph.D. Thesis, Computer Science Dept., U. Rochester, December, 1994

Walker, Marilyn A. (1996). The Effect of Resource Limits and Task Complexity on Collaborative Planning in Dialogue, Artificial Intelligence Journal, 85:1-2.

Webber, Bonnie Lynn (1991). Structure and Ostension in the Interpretation of Discourse Deixis. Language and Cognitive Processes. 6:107-135

Zeevat, Henk (forthcoming). Explaining Presupposition Triggers. In *Information Sharing: Givenness and Novelty in Information Processing,* Kees van Deemter & Rodger Kibble, (Eds.)

Contrast and Contrastive Topic

CARLA UMBACH
University of Leipzig
umbach@cs.tu-berlin.de

ABSTRACT. The starting point of this paper is the observation that in a question-answer dialog the use of *but* instead of *and* is obligatory if the answer is overinformative in that it includes an additional topic. A focus-semantic analysis of *but* is presented showing that (a) *but* is focus-sensitive and (b) *but* requires a denial with respect to the appropriate quaestio. This analysis provides a uniform basis for explaining the different uses, e.g. semantic opposition, denial-of-expectation, and the topic change use of *but*. Beyond that it gives some insight into the interaction between information structure and discourse relations in constructing the discourse.

1 Introduction

Consider the question-answer dialogs in (1)-(3). Due to the contrastive accents in the topic the answers in each of (1)-(3) have to comprise at least two conjuncts, otherwise Adam would be inclined to ask for a continuation: *"And / but what ...?"* In (1) Adam asks about all of the children, and Ben addresses one part of the children in the first conjunct and the other part in second conjunct. In (2), though Adam asks about the small children only, Ben first refers to the bigger ones, and Adam has to wait for the second conjunct to get the required information. In (3) it is the other way around: Adam's question is already answered by the first conjunct and the second conjunct gives information Adam did not ask for. Anyway, in each of the examples in (1)-(3) Adam's question is completely answered in the end.

(1) a. Adam:
What did the children do today?

b. Ben:
The **small** children stayed at HOME and/but the **bigger** ones went to the ZOO.[1]

(2) a. Adam:
What did the small children do today?

b. Ben:
The **bigger** children went to the ZOO, but/*****and the **small** ones stayed at HOME.

[1] **Boldface** type denotes a contrastive topic accent and CAPS denote a focus accent.

Paper presented at the ESSLLI 2001 Workshop on
Information Structure, Discourse Structure and Discourse Semantics
Edited by Ivana Kruijff-Korbayová and Mark Steedman.

(3) a. Adam:
What did the small children do today?

b. Ben:
The **small** children stayed at HOME, but/*****and the **bigger** ones went to the ZOO.

It is commonly assumed that in a coherent question-answer dialog the answer has to refer to the subject matter of the question only. In (2) and (3) information about an additional topic is provided without rendering the answers unacceptable. But comparing (1) and (2)/(3) we observe that in the latter case the use of *but* instead of *and* is obligatory. The use of *but* to indicate a topic change has been mentioned in the literature. Yet there is no explanation why a contrast can be used that way: Why does the use of *but* instead of *and* render an over-informative answer acceptable?

To address this question, first, I will briefly consider the notion of contrastive topic. Then I will present the outlines of a focus-semantic analysis of *but* inferring the different uses of *but* from a uniform semantic basis. (For a comprehensive discussion see Umbach in prep.). It will turn out that by using *but* instead of *and* the speaker presents the additional topic as being closely related to the original one, thus minimizing the deviation.

2 Contrastive Topic

Following e.g. Eckard (1996) and Vallduvi/Villkuna (1998), I assume two kinds of foci, sentence focus and contrastive focus. Sentence focus is expressed by a sentence default accent and partitions the sentence into a topic and a comment. Contrastive focus is due to, e.g., focus-sensitive operators and wh-questions. For contrastive focus, I adopt the idea of Alternative Semantics (cf. Rooth 1992):[2] A contrastive focus triggers the presupposition that there exists at least one proper alternative, i.e. an element differing from the ordinary meaning of the focussed phrase with respect to the accented item. For example, the contrastive focus in *the **small** children* triggers the presupposition that there exist other (groups of the aforementioned) children in addition to the small children. Following the presupposition-as-anaphor theory (van der Sandt 1992) the proper alternative triggered by the contrastive focus is regarded as an anaphor which has to be bound (or be accommodated).

A contrastive focus may occur in any position in a sentence. If, however, it occurs in the topic part, it represents a contrastive topic and will typically be marked

[2] I will not opt for a particular framework here because it's not relevant for the point I want to make in this paper. In fact, the semantics of *but* proposed in the next section will need access to both focus and background of the conjuncts thus requiring a fairly expressive framework, cf. Krifka (1999b).

with a rising accent.[3] Being in the topic position, a contrastive topic refers to an entity the speaker wants to talk about. Moreover, due to its contrastiveness, it indicates that there exist alternatives the speaker wants to talk about, too. This intuition is, e.g., captured by the partial-answer account suggested in Krifka (1999): Assuming that a sentence is an answer to some question, the role of a contrastive topic consists in indicating that the answer is a partial one. (Roughly, a sentence is a partial congruent answer to a question if it is entailed by some proposition p in the question meaning Q, but it is not a (complete) congruent answer entailing some p in Q.) In (1b), for example, the answer given in the first conjunct is partial with respect to the question in (1a) since it is entailed by the entire answer.

There are two notorious problems with contrastive topics: First, in a sequence of answers the last answer completes the requested information, so, intuitively, it is not partial any longer. Second, in the additional-topic answers in (2b) and (3b) one of the conjuncts is a complete answer and the other one is not even congruent with respect to the question. In Krifka (1999), the first problem is handled by requiring each answer in a sequence to be partial in isolation. The second problem, however, is not discussed.[4]

I will suggest a solution for the additional-topic problem distinguishing between the overt question posed in a dialog and the implicit "quaestio" addressed in the answer. The quaestio of an utterance is supposed to be a question which is a posteriori reconstructed from the utterance. In the unmarked case in (1), for example, the quaestio reconstructed from Ben's answer is equal to Adam's question. But in (2) and (3) the quaestio of the answer is slightly different from the overt question. Ignoring the connective for the moment the quaestio related to (2b) could be "*What did the small children do, and what did the bigger ones do?*". Being a posteriori reconstructed the quaestio shows which question is actually answered by an utterance, even if there is a deviation from the question that has been posed.[5] Depending on the specific interest the quaestio may be reconstructed either as a constituent question or as a polarity question.

Adopting the notion of the quaestio, the role of the contrastive topic can be defined as indicating that the answer is partial with respect to the quaestio reconstructed from the entire conjunction. Congruence then has to refer the quaestio, too: An answer is congruent if the respective quaestio entails the question. This accounts for the acceptability of the dialogs in (2) and (3). But we have to be careful not to throw the baby out with the bath water: Bringing in an additional topic obviously requires some extra effort, for example, using *but* instead of *and*. So the

[3] In terms of Steedman (2000) a contrastive topic is called theme-focus, and a contrastive focus in the comment part is called rheme-focus. Steedman's notion of focus matches with the notion of contrastive focus employed here.

[4] The open-question/strategy account of Büring (1998) handles both problems. However, it fails to cover "crossed" contrastive topics, as e.g. in (12a).

[5] It has been suggested to view the quaestio as the "question under discussion" QUD (cf. Ginzburg 1996). But then we would need a more liberal protocol for querying. According to Ginzburg, when Adam poses a question, Ben can either accept it as the top most QUD or reject it. In (2) and (3), however, Ben just slightly deviates from Adam's question neither rejecting it nor (fully) accepting it.

question is why the use of *but* facilitates a topic change.

3 Standard accounts of *but*

The semantics and pragmatics of *but* has been the topic of a lively discussion starting with Lakoff's seminal paper in 1971. Lakoff distinguished between two uses of *but*, semantic opposition (*John is tall, but Bill is short*) and denial-of-expectation (*John is tall, but he's no good at basketball*). Since then there has been a host of investigations pointing out further uses of *but*, e.g. for topic-change, and generalizing the analysis to other contrastive connectives. Few approaches, however, have tried to examine the underlying notion of contrast and trace the various uses of *but* to a uniform semantic meaning.

Recent accounts of the meaning *but* mainly draw on default knowledge. Asher (1993), for example, assumes *but* to be licenced by different polarities in the conjuncts where the polarities are due to linguistic and common world knowledge. Winter and Rimon (1994) use a default implication interpreted in possible world semantics to capture the notion of contrast. Gaerdenfors (1994) presents a semantics of *but* within his general framework of reasoning with expectations. Common to these analyses is the idea that the use of *but* indicates a denial-of-expectation, the expectation being due to default world knowledge.

On the other hand, it is easy to show that common world knowledge cannot be decisive for the use of *but*. Suppose, for example, you are not versed in botany and you don't know what loosestrife is. Nevertheless, you will interpret (4) as denying the expectation that loosestrife is found in July. This expectation, however, cannot belong to your common world knowledge, simply because you cannot have any knowledge about an entity or kind you aren't acquainted with.[6]

(4) It was July but we couldn't find any loosestrife

The example in (5) refers to the film "The English Patient".[7] The situation is this: Lord Almasy has an affair with Katherine. Katherine's husband Jeffrey has to pick up Lord Almasy by plane from somewhere in the desert. Katherine will be on the plane, too. Jeffrey, knowing about the affair, decides to crash the plane on the ground and kill them all. (5a)-(5d) tell the outcome of his plan, describing exactly the same situation. Nevertheless, in responding to different questions, the sentences differ with respect to the contrast they involve.

(5) a. (What happened?)
Jeffrey is dead, Katherine is seriously injured, and Almasy is unhurt.

b. (Did Jeffrey succeed in killing them all?)
Jeffrey is dead, but Almasy is unhurt and Katherine is alive, too.

[6] More than you ever wanted to know about loosestrife, thanks to Kathryn Bock: loosestrife = Lysimachia; invasive perennial, can in some varieties displace native plants.

[7] adopted from Braußе (1999).

c. (Have all of the participants been affected by the accident?)
Jeffrey is dead and Katherine is seriously injured, but Almasy is unhurt.

d. (Do all of the participants need a doctor?)
Jeffrey is dead and Almasy is unhurt, but Katherine is seriously injured.

The "loosestrife" example demonstrates that the expectation denied by the use of *but* need not be given by common world knowledge and is therefore not a prerequisite for the interpretation of the sentence. Instead, it is triggered by the interpretation of *but*, comparable to a presupposition or (conversational) implicature. Taking the expectation as a presupposition would allow for accommodation, which would work fine for the "loosestrife" example: Accommodate that, normally, loosestrife is found in July. But what to accommodate in the "English Patient" examples? Since the situation is the same in each of (5b)-(5d), world knowledge cannot trigger different expectations. The expectations seem to be evoked, instead, by the questions. But do we really want to accommodate, e.g. in the case of (5c), that normally, if Jeffrey/someone succeeds in killing himself, then he succeeds in killing the others, too? This is clearly absurd. The expectations induced by the questions in (5b)-(5d) are by far too ad hoc to be captured by way of accomodation.

Both examples make it plain that a contrastive relation is neither given by the meaning of the conjuncts nor induced by common world knowledge. The expectation denied by the use of *but* is obviously due to a question explicitly or implicitly posed by the preceeding discourse. So the use of *but* primarily has to comply with a question posed by the preceeding discourse. Instead of readily accommodating ad hoc expectations we will investigate the role of these questions and try to find out how they relate to the *but*-sentences and why they reflect an expectation.

4 Two novel observations

The analysis of *but* proposed in this paper takes its starting point from two characteristics which have up to now been neglected in considering the meaning of *but*: First, *but* is focus-sensitive. This is evident when you compare (6a) and (6b). In (6a) the verb phrase is focussed whereas in (6b) the subject is focussed. Due to the focus we expect different contrasts: In (6a) washing the dishes has to be contrasted with some other activity. In (6b) Bill has to be contrasted with a different person. This suggests that we should examine the alternatives induced by the focussed expressions and take the respective sets of alternatives into account.

(6) a. ... but Bill has washed the DISHES.
b. ... but BILL has washed the dishes.

The second observation relates to the questions answered by a *but*-conjunction. If the question in (7) is answered by confirming both conjuncts, the use of *but*

instead of *and* is unacceptable, cf. (8a), (8b). If the answer denies both conjuncts *but* is equally unacceptable, cf. (8c). If, however, one part of the question is confirmed and the other part denied, the use of *but* is perfect (and the use of *and* is at least marked), cf. (8d)-(8f). Denial, by the way, does not hinge on the presence of an explicit negation, cf. (8e). So, obviously, if a *but*-sentence is an appropriate answer to a question comprising two conjuncts, one of them will be confirmed and the other one will be denied.

(7) Adam: Did John clear up his room and wash the dishes?

(8) Ben:

a. [yes]John cleared up his room and [yes] he washed the dishes.

b. # [yes] John cleared up his room, but [yes] he washed the dishes.

c. # [no] John didn't clear up his room, but [no] he didn't wash the dishes.

d. [yes] John cleared up his room, but [no] he didn't wash the dishes.

e. [yes] John cleared up his room, but [no] he skipped the washing-up.

f. [no] John didn't clear up his room, but [yes] he did the washing-up.

For the quaestio of a *but*-sentence to reflect the confirm+deny characteristics it has to comprise polarity question conjuncts instead of constituent questions. The polarity question conjuncts will relate to the alternatives contrasted by *but* asking whether both alternatives apply simultaneously. Hence, in accordance with the confirm+deny characteristics one of the conjuncts of the quaestio will be confirmed by the corresponding *but*-sentence and the other one will be denied.

5 The focus-semantic analysis of *but*

The focus-semantic analysis of *but* makes use both of its focus-sensitivity and its confirm+deny characteristics. The basic idea is as follows: In a *but*-conjunction there are two corresponding foci (in the first and in the second conjunct, respectively) which establish alternatives with respect to each other.[8] The semantics of *but*, beyond being a mere conjunction, requires that one of the alternatives renders a true proposition and the other one is denied with respect to the first alternative's background. In short: *but* excludes an alternative. This, by the way, doesn't mean that *but* introduces a negation (*but* is not a "nand"!). Instead, *but* requires a negation, in the same way a verb selects an argument of a certain type. If there is no overt negation in one of the conjuncts, then the hearer is requested to reconstruct it. The fact that the quaestio has to be answered by "*Yes,... but, no, ...*" reflects the required negation.

To show that this idea applies to *but*-conjunctions in general we have to distinguish four cases:[9] Either the subject of the conjuncts is the same and the predicates

[8] Thus they are contrastive foci.

[9] I assume that commas may be substituted by full stops. Due to limitations of space correction cases are left out.

differ from each other (A), or the predicates are the same and the subjects differ from each other (B), or both subjects and predicates are different yet comparable (C), or subjects and predicates are not comparable to each other, i.e. we have to compare the entire propositions (D). To simplify matters let us assume that if there is an overt negation it occurs in the second conjunct (i.e. take only yes-no sequences into account) and consider briefly the four cases.

The A-case is illustrated in (9a) and (9b). Normally, the predicates will be focussed.[10] In (9a) the negation is overt inducing the quaestio in (9c). In (9b) there is no overt negation. The question in (9d), however, would not be the appropriate quaestio because (9b) cannot be an answer to (9d). Instead, the quaestio has to be the same as in the negated example indicating that there is an implicit negation to be reconstructed from the complement of the predicate (i.e. *skip the washing up* is supposed to be the set complement of *wash the dishes*). So the relevant alternatives in (9a) and (9b) are the same: *clear up the room* as against *wash the dishes.*

(9) a. [yes] John cleared up his ROOM, but [no] he didn't wash the DISHES.
b. [yes] John cleared up his ROOM, but [no] he skipped the WASHING-UP.
c. Did John both clear up his room and wash the dishes?
d. Did John both clear up his room and skip the washing-up?

The B-case is given in (10a): The subjects are focussed and establish alternatives with respect to each other. The quaestio is given in (10b). Contrary to the A-case examples explicit negation is obligatory, cf. (10c). The reason for this is easy to see: Individuals, as opposed to predicates, don't have complements (there is no "non-John"). However, B-case examples will be acceptable without explicit negation if the particle *too* is added, cf. (11a). Note that, when adding *too*, the quaestio will be a different one, cf. (11b). Similarly, B-case examples with a negation in both conjuncts will be acceptable if the particle *either* is added (cf. (11c) and the quaestio in (11d).

(10) a. [yes]JOHN cleared up his room, but [no] BILL didn't.
b. Did both John and Bill clear up their rooms?
c. *John cleared up his room, but Bill did.

(11) a. John cleared up his room, but Bill did, too.
b. Did John clear up his room, and was he the only one who did?
c. John didn't cleared up his room, but Bill didn't, either.
d. Did John leave his room in a mess, and was he the only one who did?

The C-case is more complex because we have to consider two foci in each of the conjuncts, one of them being a contrastive topic. The contrastive topics may

[10]There may be an additional focus on the negation, then the predicate in second conjunct has to be regarded as a contrastive topic. In German, in this case, the word order will be reversed: *John hat* AUFGERÄUMT, *aber* ***abgewaschen*** *hat er* NICHT.

be "parallel" comprising either the subjects or the predicates, or they are "crossed" comprising the subject of the first conjunct and the predicate of the second conjunct, or vice versa. (12a) presents one of the crossed variants. The it-cleft paraphrase in (12b) makes it clear that the focus (in the comment part of the sentence) is on *Bill*, and *doing the dishes* represents a contrastive topic.[11] For this reason in German the word order is reversed, cf. (12c). In (12a), as in all of the C-case examples, there is no explicit negation in either of the conjuncts. Nevertheless, there is a denial, because (12a) clearly entails that John did not wash the dishes. Taking the entailment into account, the quaestio in (12d) is again partly confirmed and partly denied, as demonstrated in (12e). Note, that the quaestio is exactly the same as the one in (9c). In fact, (12a) and (9a) are very much alike both conveying the information that John cleared up his room and did not do the dishes. But in (12a) we additionally learn who finally did the washing up.

(12) a. **John** cleared up his ROOM, but BILL did the **dishes**.
b. **John** cleared up his ROOM, but it was BILL who did the **dishes**.
c. **John** hat AUFGERÄUMT, aber **abgewaschen** hat BILL.
d. Did John both clear up his room and wash the dishes?
e. [yes] John cleared up his room, but [no, John did not do the dishes] the dishes were washed by Bill.

Let us skip over the second crossed variant and look at one of the parallel variants. In (13a) the contrastive topics are parallel being the subjects in both of the conjuncts. The quaestio is given in (13b). The denial of part of the quaestio is entailed by telling what Bill did instead of clearing up the room, cf. (13c).

(13) a. **John** cleared up his ROOM, but **Bill** did the DISHES.
b. Did both John and Bill clear up their room?
c. [yes] John cleared up his room, but [no, Bill did not clear up his room] Bill did the dishes.

The last of the four cases concerns *but*-sentences with wide foci in the conjuncts, cf.(14a). In these cases the entire propositions have to be regarded as being alternatives with respect to each other. If there is no explicit negation in one of the conjuncts, e.g. (14b), it has to be reconstructed. Note that the appropriate quaestio for (14b) has to be (14c) instead of (14d).

(14) a. $[\text{It is raining}]_F$, but $[\text{we are not going to stay at home}]_F$.
b. $[\text{It is raining}]_F$, but $[\text{we are going to go for a walk}]_F$.
c. Is it raining, and are we going to stay at home?
d. Is it raining, and are we going to go for a walk?

[11] In English the it-cleft variant is clearly preferred over (12a). In German the reversed word order variant in (12c) will be the preferred one.

To sum up these findings:[12] First, in the second conjunct of a *but*-sentence there is a **focus associated with *but*** (F_{but}). It is either a contrastive topic, if there is one, or a regular focus. The focus associated with *but* represents the **expected alternative** (EA), that is, the alternative denied with respect to the quaestio.[13] Second, in the first conjunct there has to be a **corresponding focus** (F_{corr}) that contains the expected alternative in its set of alternatives. Third, there is a **denial condition**, stating that the proposition resulting from substituting the expected alternative for the corresponding focus is false, i.e $\neg[...F_{corr}/EA...]_{C1}$.[14] Suppose the meaning of the first conjunct is given by $[...F_{corr}\ ...]_{C1}$, and the meaning of the second conjunct is given by $[...F_{but}\ ...]_{C2}$. Then the meaning of a *but*-conjunction "C1 *but* C2" is given by:

$$[...F_{corr}...]_{C1} \wedge [...F_{but}...]_{C2} \wedge \neg[...F_{corr}/EA...]_{C1}$$

The crucial point in the semantics of *but*, which distinguishes *but* from a mere conjunction, is the denial condition. In the A-case, with an overt negation in the second conjunct, the denial condition is trivially satiesfied, because it is given by the second conjunct. For example, in (9a), repeated in (15), the focus associated with *but* is *wash the dishes*, and this is also the expected alternative. The corresponding focus is *clear up the room*. So the denial condition is "It's not the case that John washed the dishes", which is equivalent to the second conjunct.

(15) (= 9a) John [cleared up the ROOM]$_{Fcorr}$, but he didn't [wash the DISHES]$_{Fbut}$.

If an A-case example occurs without overt negation the negation has to be reconstructed using the predicate's complement, cf. (9b). This time, the expected alternative is given by the predicate's complement instead of being directly given by F_{but}. The denial condition will then be entailed by the meaning of the second conjunct.

With B-case examples and D-case examples satisfaction of the denial condition is similarly trivial. In the C-case examples, however, the denial condition is not given by one of the conjuncts. For example, in (12b), repeated in (16), the focus associated with *but* is a contrastive topic, i.e.*wash the dishes*, which is also the expected alternative. The corresponding focus is *clear up the room* matching with the type of the expected alternative. Hence the denial condition is the same as above: "It's not the case that John washed the dishes". But this time, the denial condition is an entailment resulting from the additional information concerning who/what instead of the expected alternative satisfies the proposition.

[12] As said in the beginning, the presentation in this paper is restricted to confirm+deny sequences. To include deny+confirm sequences the definitions have to be extended.

[13] In cases likes (9b) and (14b) where the negation has to be reconstructed by means of the complement of the focussed expression, the expected alternative is given by the complement.

[14] Dots indicate the part of the conjunct which is not subsumed by F_{corr} or F_{but}, i.e. either the background or another focus. F_{corr}/EA means the substitution of EA for F_{corr}.

(16) (= 12b) **John** [cleared up his ROOM]$_{Fcorr}$, but it was BILL who [washed the **dishes**]$_{Fbut}$.

Since it is the denial condition which distinguishes *but* from a mere conjunction, expectations denied by a *but*-sentence are obviously induced by the denial condition. This is in accordance with a well-known property of negated sentences in general: Negated sentences commonly trigger the implicature that the speaker expects (or assumes the hearer to expect) that the affirmative proposition holds (cf. Givon 1978). This is the reason why we reconstruct the quaestio asking whether both of the alternatives hold simultaneously, e.g., whether John did both, clear up his room and also wash the dishes. Note that, due to this quaestio there is an expectation that both of the alternatives do hold simultaneously: If John cleared up his room, he will have washed the dishes, too. So, finally, the focus-semantic analysis confirms the idea that there is an expectation denied by the use of *but*. However, contrary to what is said in the literature, the expectation is not given by common world knowledge. Instead, it is triggered by the denial condition inducing the special form of the quaestio. The quaestio, of course, has to be in line with the previous context. But that is a general problem of discourse construction, not restricted to the use of *but*.

6 ”concessive *but*”?

In some contexts a concessive marker apparently can be added or even be substituted for *but* without affecting the meaning of the sentence. From that it has been concluded that there is a concessive use of *but* (e.g. Grote et al. 1997). However, regarding *but* as being interchangeable with a concession in these contexts presupposes that a concession is interpreted as indicating a denial of expectation. As opposed to that, König (1991) convincingly argues that a concession expresses ”incausality”, thus accounting for the close relationsship between causal and concessive statements. Following König’s incausality analysis it is easy to show that a concession is not a special case of a contrast: First, according to the incausality interpretation (17a) has to be paraphrased as (17b). Second, it is well-known that *and*-conjunctions may be interpreted in many different ways, e.g. as temporal or as causal relations, cf. (18a), (18b). But this is an overinterpretation by the hearer, not included in the meaning of *and* (cf. Posner 1980). Third, due to the focus-semantic analysis interpreting (19a) requires to reconstruct a negation (analogous to (14b)). Now, (19b) may be overinterpreted in a causal way, too, cf. (20). Due to the negation causal overinterpretation results in incausality, compare (17b) and (20).

(17) (incausality analysis of concessives)

a. Although it is raining Mary is happy.

b. It it raining, and it is not the case that Mary is not happy because of that.

(18) (causal overinterpretation of *and*)

a. It is raining, and Mary is happy

b. It is raining, and Mary is happy because of that.

(19) (reconstruction of the negation)

a. It is raining, but Mary is happy.

b. It is raining, and it is not the case that Mary is not happy.

(20) (causal overinterpretation of *but*)
It is raining and it is not the case that Mary is not happy because of that.

Hence, there is no ”concessive *but*”, just as there is no ”causal *and*”— interpreting *but* as a concession is due to overinterpretation. At the same time, a contrast is perfectly compatible with a concession, just as a causal relation is perfectly compatible with a conjunction. This suggests that contrast and conjunction, on the one hand, and concession and causality, on the other hand, are different types of discourse relations exploiting different features of the discourse. Concession and causality represent relations between propositions, or states of affairs, i.e. semantic/external relations (cf. Mann, Thompson 1988). Thus they may be realized by an adverbial containing a propositional anaphor (*because of that, in spite of that* etc.) and establish an anaphoric link (cf. Webber et al. 1999).

A contrast, on the other hand, is based on the information structure of the sentence combining subsequent foci. Making use of the information structure, a contrast is a genuine structural relation, i.e. no semantic/external relation. Yet it should not be subsumed under the notion of pragmatic/internal relations because the latter are usually tied to the illocutionary aspects of their arguments (Sanders et al. 1992). The relation of contrast given by *but* seems to resist the standard classification, which may help to clarify the ongoing discussion about the types of discourse relations (cf. e.g. Moore and Pollack 1992, Bateman and Rondhuis 1997)

7 Topic change

Let us finally come back to the dialogs in (1)-(3), repeated in (21)-(23). The starting point of this paper was the question why in (22b) and (23b) the use of *but* instead of *and* is obligatory. Consider the quaestiones given below.[15] In the unmarked case in (21b) both *and* and *but* are acceptable because Ben may intend his answer as either referring to the question (21c) or to the one in (21d). Note, however, that there is a crucial difference: The *but*-quaestio but not the *and*-quaestio triggers

[15] For ease of comparison the quaestio is given here consisting of a polarity and a constituent question conjunct, assuming that answering a constituent question simultaneously confirms the respective polarity question, so this form of quaestio is also answered by a confirm+deny sequence.

the expectation that the bigger children did the same thing as the small ones did. By using *but* in (22b) and (23b) Ben deliberately conveys this expectation. In this way, although actually deviating from the original topic of Adam's question, Ben presents the additional topic as being closely related to the original one. Thus, by using *but* Ben suggests that the additional topic is relevant, too, and the deviation is reasonable.

(21) a. Adam:
What did the children do today?

b. Ben:
The **small** children stayed at HOME and/but the **bigger** ones went to the ZOO.

c. Ben's quaestio when using *and*:
What did the small children do and what did the bigger ones do?

d. Ben's quaestio when using *but*:
What did the small children do, and did the bigger ones do the same?

(22) a. Adam:
What did the small children do today?

b. Ben:
The **bigger** children went to the ZOO, but the **small** ones stayed at HOME.

c. Ben's quaestio:
What did the bigger children do, and did the small ones do the same?

(23) a. Adam:
What did the small children do today?

b. Ben:
The **small** children stayed at HOME, but the **bigger** ones went to the ZOO.

c. Ben's quaestio:
What did the small children do, and did the bigger ones do the same?

To conclude, the dialogs in (22) and (23) clearly demonstrate that an answer need not refer to only the topic of the question. This suggests that a natural language dialog should not be conceived as a server-client relation where B has to answer all and only A's questions. Partners in a dialog seem to be ”peer-to-peer”: They are entitled to introduce an additional topic, but they are bound to relate the additional topic to the original one, thus minimizing the deviation. One way to do this is by using the conjunction *but*.

Bibliography

[1] Asher, N. *Reference to Abstract Objects in Discourse*. Kluver, 1993.

[2] Bateman, J., Rondhuis, K. *Coherence Relations: Towards a General Specification.* Discourse Processes 24, 1997.

[3] Braußе, U *Was ist Adversativität?*ABER *oder* UND. Deutsche Sprache 2, 1999.

[4] Büring, D. *Focus and Topic in a Complex Model of Discourse*. In: Geurts, Krifka, van der Sandt (1999) Focus and presupposition in a multi-speaker discourse. ESSLI-99 Workshop Reader, 1999.

[5] Eckardt R. *Intonation und Predication.* Arbeitspapiere des SFB 340. Universität Düsseldorf, 1996.

[6] Foolen, A. *Polyfunctionality and the semantics of adversative conjunctions.* Multilingua 10-1, 1991.

[7] Gaerdenfors, P. *The role of expectations in reasoning.* In: Masuch, M., Plos, L. (eds) Knowledge Representation and Reasoning Under Uncertainty. Springer, Berlin/New York, 1994.

[8] Ginzburg, J. *Interrogatives: Questions, Facts and Dialogue*. In Lappin, S. (ed.) The Handbook of Contemporary Semantic Theory, Blackwell, Oxford, 365-422, 1996.

[9] Givon, T. *Negation in Language: Pragmatics, Function, Ontology*. In Cole, P. (ed) Syntax and Semantics: Pragmatics. Vol. 9, Academic Press, New York, 1978.

[10] Grote, B., Lenke, N., Stede, M. *Ma(r)king Concessions in English and German*. Discourse Processes 24, 87-117, 1997.

[11] König, E. *Concessive relations as the dual of causal relations.* In Zaefferer, D. (Ed.) Semantic universals and universal semantics. Groningen-Amsterdam studies in semantics 12, Foris, 1991.

[12] Krifka, M. *Additive particles under stress*. Proceedings of SALT 8. Cornell, CLC Publications, 1999.

[13] Krifka, M. *Frameworks for association with focus*. Workshop "Intonation and Interpretation", ESSLLI 99, Utrecht, 1999.

[14] Lakoff, R. *If's, and's and but's about conjunction.* In Fillmore, Langendoen (eds.): Studies in linguistic semantics. de Gruyter, New York, 1971.

[15] Mann, W., Thompson, S. *Rhetorical structure theory: Towards a functional theory of text organization.* TEXT 8 (2), 1988.

[16] Moore, J.D., Pollack, M.E. *A Problem for RST: The need for multi-level discourse analysis.* Computational Linguistics, 18(4), 1992.

[17] Posner, R. *Semantics and pragmatics of sentence connectives.* In Searle, Kiefer, Bierwisch (eds.) Speech act theory and pragmatics. Reidel, Dordrecht, 1980.

[18] Rooth, M. *A Theory of Focus Interpretation.* Natural Language Semantics 1, 1992.

[19] Sanders, T., Spooren, W., Noordman, L. *Toward a Taxonomy of Coherence Relations.* Discourse Processes 15, 1992.

[20] van der Sandt, R. *Presupposition Projection as Anaphora Resolution.* Journal of Semantics 9, 1992.

[21] Steedman, M. *Information Structure and the Syntax-Phonology Interface.* Linguistic Inquiry 31, 4, 649-689, 2000.

[22] Vallduví, E., Vilkuna, M *On Rheme and Kontrast.* Syntax and Semantics, Vol.29. Academic Press, 1998.

[23] Webber, B., Knott, A., Joshi, A. *Multiple discourse connectives in a lexicalized grammar for discourse.* Third International Workshop on Computational Semantics, Tilburg, 1999.

[24] Winter, Y., Rimon, M. *Contrast and Implication in Natural Language.* Journal of Semantics 11, 1994.

Intonational Phrasing and Discourse Segmentation

KLAUS VON HEUSINGER
Universität Konstanz
klaus.heusinger@uni-konstanz.de

ABSTRACT.

Theories that relate intonational structure and discourse structure concentrate on intonational and informational units that either correspond to a clause (proposition) or to a single referent. However, intermediate phrases of the intonational phrasing often segment units that are smaller than a clause, but that do not introduce a referent. Even though they are acknowledged (cf. Selkirk's "sense units"), there are no approaches that account for their role in building the discourse structure. The paper presents a new approach that describes the representation of such intermediate phrases and accounts for their functions in the construction of a more fine-grained discourse structure.

1 Introduction

Intonation contours are represented by phonologists as a sequence of abstract tones consisting of pitch accents and boundary tones. Pierrehumbert and Hirschberg (1990, 308) assign discourse functions to the particular tones: "Pitch accents convey information about the status of discourse referents (...). Phrase accents [= boundary tones of intermediate phrases] convey information about the relatedness of intermediate phrases (...). Boundary tones convey information about the directionality of interpretation for the current intonational phrase (...)." The status of discourse referents can be accounted for in terms of given vs. new; the boundary tones of intonational phrases indicate how the proposition expressed by the whole phrase is integrated into the discourse. Similarly, boundary tones of intermediate (or phonological) phrases that correspond to a full proposition indicate the way these propositions are interpreted with respect to the linguistic context, as illustrated in (1) and (2). However, in this view there is no account of treating intermediate phrases that correspond to units below the clause level and above the level of simple discourse referents, such as the modification "im achtzehnten Jahrhundert" ("in the eighteenth century") or the unsaturated phrase "lebte in Frankreich" ("lived in France") in example (3):

Paper presented at the ESSLLI 2001 Workshop on
Information Structure, Discourse Structure and Discourse Semantics
Edited by Ivana Kruijff-Korbayová and Mark Steedman.

(1) (George ate chicken soup) (and got sick)
L LL%

(2) (George ate chicken soup) (and got sick)
H LL%

(3) [(Im achtzehnten Jahrhundert) ∼(lebte in Frankreich)] | [(ein Mann,) (∼der zu den genialsten und abscheulichsten Gestalten dieser an genialen und abscheulichen Gestalten nicht armen Epoche gehörte.)]

"In the eighteen century France there lived a man who was one of the most gifted and abominable personages in an era that knew no lack of gifted and abominable personages"

In order to account for these intermediate units and their functions in building a discourse structure, I modify the general assumptions of discourse theories, namely that the discourse structure is exclusively realized between propositions (Hobbs 1990, Roberts 1996, Büring 2000 among others), and the view of DRT (Kamp and Reyle 1993) that discourse structure only relates discourse referents to each other. I extend Asher's (1993, 1999) segmented DRT (SDRT) by assuming that we can also create discourse relations between sub-clausal elements such as modifications or unsaturated clauses. The remainder of the paper is organized as follows: In section 2, I discuss different approaches to the function of intonational contours. I show that no attention is given to intermediate phrases and their discourse function. In section 3, I present one family of approaches to discourse structure: DRT, a theory that focuses on anaphoric relation in terms of accessibility relations between discourse referents. In section 4, I introduce Asher's segmented DRT, a modification of classical DRT that allows to express relations between propositions in a discourse. In section 5, I modify Asher's discourse model in order to implement even smaller units those units that correspond to the intermediate phrases of the intonational structure.

2 The meaning of intonational structure

Pierrehumbert (1980) represents an intonational contour by a tune consisting of abstract tones (4), which are generated by a finite state grammar (Figure 18.1) that combines the tones listed into legal tunes:

(4) *Phonological tones* (Pierrehumbert 1980)

a. Each phrase requires at least one pitch accent (for English: H*, L*, or bitonal as H*+L, H+L*, L*+H, L+H*, and H*+H)

b. Each phrase receives a phrase accent (H^-, L^-) at the end of the word that is associated with the last pitch accent

c. Each phrase ends with a boundary tone (H%, L%)

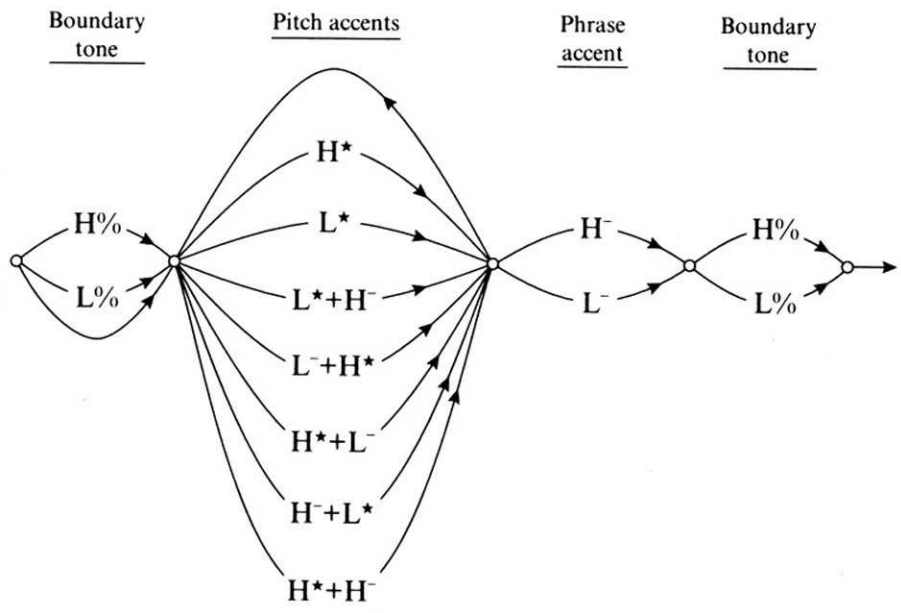


Figure 18.1: The finite state grammar of Pierrehumbert (1980)

This grammar accounts for the legal tune and therefore describes the possible composition of tones to tunes. However, Pierrehumbert (1980) does not account for the relations between the elements in the tune. There have been different suggestions how to interpret the relation between the elements of the tune and their functions: (i) The most classical description is in terms of phonological words, a tradition going back to Liberman's view of the contour as 'ideophonic word'. (ii) Steedman (1991, 2000) describes the pitch accent as functions that require an argument in order to yield a tune. His approach reflects the reconstruction of linguistic configurations in categorial grammar. (iii) Pierrehumbert and Hirschberg (1990) propose that the different tones have independent functions ranging over different domains.

2.1 The tune describing an ideophonic word

Hayes and Lahiri (1991) execute the program of the intonational lexicon proposed by Liberman (1975) in the most comprehensive way. They account for the different tunes in Bengali and decompose the tunes into *intonational stems*, *suffixes* and *prefixes* (they use *I* for intonational phrase and *P* for phonological or intermediate phrase (*ip*)):

(5) The intonational lexicon of Bengali Hayes and Lahiri (1991)

(a) Accents	L*	question accent
("stems")	H*	declarative accent
	L^*H_P	focus accent
(b) Boundary Tones	L_I	neutral
("suffixes")	L_I H_I	continuation rise
	H_I	offering
	H_I L_I	yes/no
(c) Prefix	L+	finality marker
		(form L+H* when attached to H*)

These tones combine to the following tunes, assuming that there are three pitch accents, and optional T_P and one obligatory T_P (Hayes and Lahiri 1991, 72). Note that a focused phrase is marked by a L*-pitch accent and an obligatory H_p boundary tone (of an intermediate phrase)

(6) Possible tunes in Bengali

L^*H_I	Offering
$L^*H_IL_I$	Yes/no question
$L^*H_PH_I$	Focus
$L^*H_PL_ILH_I$	Focus with continuation rise
H^*L_I	Declarative
$H^*L_IH_I$	Declarative with continuation rise
$L+H^*L_I$	Downstep
$L+H^*L_IH_I$	Downstep with continuation rise

This view is influenced by the phonological tradition of analyzing units like the syllable or the phonological word, which consist of one prominent part and other more or less important parts. Here, the pitch accent constitutes the central part and the boundary tones additional affixes.

2.2 Tune representing the functional contribution to the utterance

Steedman (1991, 2000) executes Halliday's thematic structure in terms of combinatory categorial grammar (CCG). This can be illustrated with the following example which receives the thematic structure in *theme-rheme.* Both thematic units are further divided into given material and new material; the latter is associated with the pitch accent.

(7) Q: I know that Mary's FIRST degree is in PHYSICS.
But what is the subject of her DOCTORATE?

A: (Mary's DOCTORATE) (is in CHEMISTRY)
L+H*LH% H* LL%
Given *New* *Given* *New*
theme *rheme*

The basic units are the theme and the utterance. All other parts are defined with respect to these basic elements. For example, the rheme is a function that takes the theme as argument yielding the utterance (this is of course, the instantiation of the subject-predicate structure in terms of functional application). Steedman now defines the function of the pitch accent L+H* as theme that misses a boundary tone, i.e. as a function that needs a boundary tone to yield a theme. Analogously, the pitch accent H* indicates a function that needs a boundary tone in order to yield a rheme. Thus in the description of tones, Steedman assumes the boundary tones and the whole tune as the primary units.

(8) Categorial functions of tones (Steedman 1991)

LH%	boundary tone	simple argument
LL%	boundary tone	simple argument
L+H*	pitch accent	function from boundary tone into theme
H*	pitch accent	function from boundary tones into rheme
L+H*LH%	contour	simple argument: theme
H* LL%	contour	function from themes into utterance

Steedman uses the terms *theme* and *rheme* as well as *given* and *new*. The first pair can be defined with respect to the sentence under analysis. Yet the second pair can only be defined by the discourse in which the sentence is embedded.

2.3 Tones representing different discourse functions

Pierrehumbert and Hirschberg (1990) give a list of functions of pitch accents and boundary tones. The latter indicate whether the phrase to which the boundary tone is associated should be interpreted with respect to the preceding discourse or to the following discourse. Pierrehumbert and Hirschberg (1990, 304) illustrate this point on the following contrast between (9) and (10). The low boundary tone L% in (9) indicates that this sentence is related to the discourse by its own, while the high boundary tone H% in (10a) indicates that it is to be interpreted with respect to the following sentence. This difference influences the choice of the antecedent of the pronoun *it*: In (9) it refers to the following proposition *I spent two hours figuring out how to use the jack*, while in (10) it refers back to the *new car manual* (see also examples (1) and (1) above):

(9) a. My new car manual is almost unreadable
LL%

b. It's quite annoying
LH%

c. I spent two hours figuring out how to use the jack
LL%

(10) a. My new car manual is almost unreadable
LH%

b. It's quite annoying
LL%

c. I spent two hours figuring out how to use the jack
LL%

Pierrehumbert and Hirschberg (1990, 308) assign the following discourse functions to the particular tones:

> Pitch accents convey information about the status of discourse referents, modifiers, predicates, and relationships specified by accented lexical items. Phrase accents convey information about the relatedness of intermediate phrasesin particular, whether (the propositional content of) one intermediate phrase is to form part of a larger interpretative unit with another. Boundary tones convey information about the directionality of interpretation for the current intonational phrasewhether it is "forward-looking" or not.

In explaining the function of intermediate boundaries tones (their "phrase accents"), they refer to the "propositional content" of the corresponding phrase. However, not all intermediate phrases express a propositional content, some might only refer to modifications such as "im achtzehnten Jahrhundert" ("in the eighteenth century") or the unsaturated phrase "lebte in Frankreich" ("lived in France") of example (3), repeated as (11).

(11) [(Im achtzehnten Jahrhundert) ∼(lebte in Frankreich)] | [(ein Mann,) (∼der zu den genialsten und abscheulichsten Gestalten dieser an genialen und abscheulichen Gestalten nicht armen Epoche gehörte.)]

"In the eighteen century France there lived a man who was one of the most gifted and abominable personages in an era that knew no lack of gifted and abominable personages"

(11) is the first sentence in Patrick Süskind's novel "Das Parfum". ("Perfume") (the data are from Braunschweiler and Fitzpatrick and Lahiri 1998ff). Intermediate phrases in intonational structure not always correspond to propositions or to simple discourse referents. Therefore, we need a more fine-grained discourse structure that allows to construct corresponding discourse segments.

3 Discourse structure

There are two main families of approaches to discourse structure. According to one family, discourse structure is understood as realizing relations between propositions. Here the structure is represented as a tree of propositions (e.g. Hobbs 1990, Roberts 1996, Büring 2000) as illustrated in Figure 18.2. In the other view the discourse is incrementally (re)constructed, as illustrated in Figure 18.3. This program is executed in DRT (Kamp and Reyle 1993), which will be presented below.

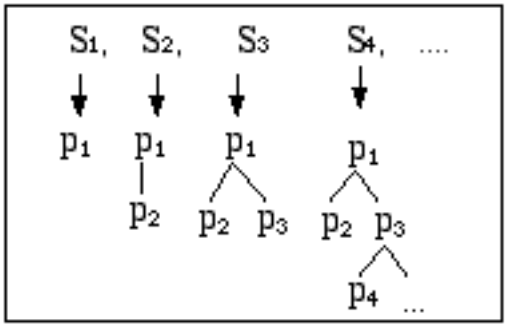


Figure 18.2: Discourse structure as a tree

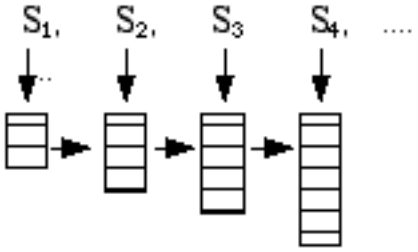


Figure 18.3: Incremental discourse structure

The initial problem that motivated discourse representation theories is the interpretation of nominal and temporal anaphora in discourse. The phenomenon of cross-sentential anaphora forces a semantics to extend its limits from the sentence to the discourse. The key idea in the approach to semantics of discourse, exemplified in (Heim 1982) and (Kamp 1981), is each new sentence or phrase is interpreted as an addition or 'update' of the context in which it is used. This update often involves connections between elements from the sentence or phrase with elements from the context. Informally described, a sequence of sentences S_1, S_2, S_3, S_4 is interpreted by incrementally constructing a discourse representation as in Figure 18.3 above. Anaphoric relations and definite expressions are captured by links between objects in this representation. In order to derive the truth condition of the sentence, the representation is embedded into a model. The DRS in (12b) graphically describes a discourse representation structure (in short DRS) with two parts.

One part is called the universe of the DRS, the other its condition set. A DRS is an ordered pair consisting of its *universe* and *condition set*, written as $<U_K, Con_K>$. The DRS in (12b) has as its universe one discourse referent x and as its condition a set of properties that are ascribed to the *discourse referents* in the universe. In (12b) the property of being a man and of walking is ascribed to the discourse referent x.

(12) a. A man walks
b. {x | man(x) and walk(x)}

The sequence or conjunction of two sentences as in (13a), receives a DRS incrementally. We start with the already established DRS for the first conjunct in (13b), then a new discourse referent for the pronoun *he* and a condition for the predicate *whistle* is added in (13c). The anaphoric link of the pronoun is graphically represented as y=?, indicating that the pronoun is still unresolved. The discourse referent which stands for an anaphoric expression must be identified with another accessible discourse referent in the universe, here the *y* is identified with the *x*, as in (13d).

(13) a. A man walks. He whistles
b. {x | man(x) & walk(x)}
c. {x,y | man(x) & walk(x) & y=? whistle(y)}
d. {x,y | man(x) & walk(x) & y=x whistle(y)}

The new discourse referent introduced by the pronoun must be linked or identified with an already established and accessible discourse referent. DRT defines accessibility in terms of structural relations, i.e., the discourse referent must be in the same (or a higher) universe. With this concept of accessibility, the contrast between (14) and (15) can be described by the difference in the set of discourse referents that are accessible for the discourse referent *u* of the pronoun *it*. The construction rule for the negation in 15 creates an embedded discourse universe with the discourse referent *y* and the conditions *donkey(y)* and *x owns y*. The anaphoric pronoun *it* in the second sentence cannot find a suitable discourse referent since it has no access to the embedded discourse universe with the only fitting discourse referent *y*.

(14) a. Pedro owns a donkey. He beats it.
b. {x,y,z,u | Pedro(x) & donkey(y) & x owns y z=x & u=y & z beats u}

(15) a. John does not own a donkey. *He beats it.
b. {x,z,u | John(x) & {y | & x owns y} z=x & u=? & z beats u}

4 Segmented DRT

Asher (1993; 1999) develops the extension "segmented DRT" (=SDRT) that is not confined to the incremental composition of DRSs, but also captures discourse relations between the sentences in the discourse. He revises the classical DRT of

Kamp (1981) and Kamp and Reyle (1993). Since the meaning of each sentence is construed as a function from truth conditions to truth conditions, the truth conditional content of the whole discourse is reconstructed by the sequential application of these functions. Asher (1993, 256) notes that "the notion of semantic updating in the original DRT fragment of Kamp (1981) (...) is extremely simple, except for the procedures for resolving pronouns and temporal elements, which the original theory did not spell out. To build a DRS for the discourse as a whole and thus to determine its truth conditions, one simply adds the DRS constructed for each constituent sentence to what one already had. (...) This procedure is hopelessly inadequate, if one wants to build a theory of discourse structure and discourse segmentation." In SDRT, each sentence S_i is first represented as a particular SDRS for that sentence. The SDRS can than interact with the already established DRS reconstructing a discourse relations *R*, such as causation, explanation, coherence, elaboration, continuation, as informally sketched in Figure 18.4. Only in a second step the representation is integrated into the already established representation.

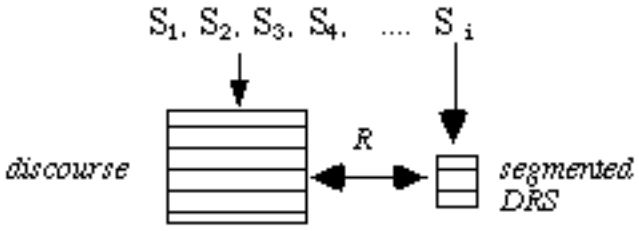


Figure 18.4: Construction of a segmented DRS in SDRT

To summarize this very short presentation of DRT: The discourse structure of DRT provides not only a new structure but also introduces new semantic objects, the discourse referents, the conditions and the discourse domains ("boxes"). DRT explains semantic categories such as definiteness and anaphora in terms of interaction between these representations. Furthermore, the extension to SDRT allows expressions of discourse relations between whole propositions as well.

5 Phrasing and Segmentation

The basic assumption of this paper is that intonational structure reflects discourse structure. Thus each element of the intonational structure must be assigned a function in constructing the discourse. Pitch accents introduce or refer to discourse referents, boundary tones of intonational phrases relate the content of these phrases, a proposition, to other propositions. However, there has been no function for phrases that do not express a propositions. The investigation of Hayes and Lahiri (1991) have shown that such boundary do function in discourse structure, e.g. as right edge of a focused phrase (cf. (5) and (6)). However, there is no general approach to these boundary tones in terms of discourse semantics. Here I sketch a new analy-

sis of intermediate intonational phrases and the corresponding discourse segments:

I introduce a mapping relation from intermediate intonational phrasing to the string (surface structure) of the sentence under discussion. For example, the short pauses indicated by "∼", and the long ones ("|") in (3), repeated as (16), yield the segmentation in Figure 18.5. Second, each segment receives a discourse representation. Third, we can establish relations between the discourse representation of the sub-clausal units and the of the established discourse structure, as in Figure 18.6. I assume the following two relation for modifications: (i) modifying an already existing discourse referent; (ii) setting the stage for a discourse referent to be introduced. Modificational relation allows that the information is integrated into the already established discourse. Fourth, after having established these relations, the rest of the sentence can be analyzed, and finally the representation of the whole sentence is integrated into the discourse structure, as in Figure 18.7:

(16) [(Im achtzehnten Jahrhundert) ∼(lebte in Frankreich)] | [(ein Mann,) (∼der zu den genialsten und abscheulichsten Gestalten dieser an genialen und abscheulichen Gestalten nicht armen Epoche gehörte.)]

"In the eighteen century France there lived a man who was one of the most gifted and abominable personages in an era that knew no lack of gifted and abominable personages"

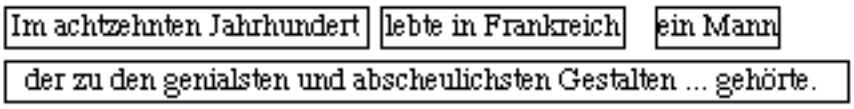


Figure 18.5: Discourse Segmentation

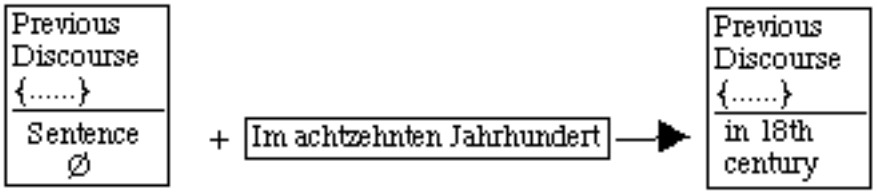


Figure 18.6: Discourse Relation "setting the stage"

I have proposed an extension of Asher's SDRT with smaller discourse representations and new relations between sub-clausal discourse representations. This allows to assign discourse functions to intonational phrasing, including boundary tones for intermediate phrases.

Figure 18.7: Adding the representation for the whole sentence

Bibliography

Asher, N. (1993). *Reference to Abstract Objects in Discourse*. Kluwer Academic, Dordrecht.

Asher, N. (1999). Discourse and focus/background distinction. In P. Bosch and R. van der Sandt (Eds.), *Focus. Linguistic, Cognitive, and Computational Perspectives.*, pp. 247–267. Cambridge: Cambridge University Press.

Braunschweiler, N., J. Fitzpatrick, and A. Lahiri (1998ff.). The konstanz intonation database: German, swiss german, american english, british english, east bengali, west bengali. University of Konstanz.

Büring, D. (2000). Focus and topic in a complex model of discourse, ms. ucsc.

Hayes, B. and A. Lahiri (1991). Bengali intonational phonology. *Natural Language and Linguistic Theory* (9), 47–96.

Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. Ph. D. thesis, University of Massachusetts. Amherst. Ann Arbor: University Microfilms.

Hobbs, J. (1990). The pierrehumbert-hirschberg theory of intonational meaning made simple. comments on pierrehumbert and hirschbert. In P. R. Cohen, J. Morgan, and M. E. Pollack (Eds.), *Intentions in Communication.*, pp. 313–323. Cambridge/Mass.: MIT Press.

Kamp, H. (1981). A theory of truth and semantic interpretation. In J. Groenendijk, T. Janssen, and M. Stokhof (Eds.), *Formal Methods in the Study of Language.*, pp. 277–322. Amsterdam: Amsterdam Center.

Kamp, H. and U. Reyle (1993). *From Discourse to Logic. Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory.* Dordrecht: Kluwer.

Liberman, M. (1975). *The Intonational System of English*. Ph. D. thesis, MIT. Cambridge/Mass.

Pierrehumbert, J. (1980). *The Phonology and Phonetics of English Intonation.* Ph. D. thesis, MIT. Cambridge/Mass.

Pierrehumbert, J. and J. Hirschberg (1990). The meaning of intonational contours in the interpretation of discourse. In P. R. Cohen, J. Morgan, and M. E. Pollack (Eds.), *Intentions in Communication.*, pp. 271–311. Cambridge/Mass.: MIT Press.

Roberts, C. (1996). Information structure in discourse. towards an integrated formal theory of pragmatics. In J.-H. Yoon and A. Kathol (Eds.), *Ohio State University [=OSU] Working Papers in Linguistics*, Volume 49, pp. 91–136.

Steedman, M. (1991). Structure and intonation. *Language 67*, 260–296.

Steedman, M. (2000). *The Syntactic Process.* Cambridge: MIT Press.