



Modeling Information Structure for Computational Discourse and Dialog Processing

Ivana Kruijff-Korbayová

korbay@coli.uni-sb.de

<http://www.coli.uni-sb.de/~korbay/esslli04/>

ESLLI 2004 Advanced Course

Nancy, 16-20 August 2004

Lecture 5 Outline

- Comparison of IS approaches, aligning terminologies
- Practical evaluation
- Corpus annotation

Reading:

- Course Reader: Section 2.1: Two Dimensions of IS
- For further reading suggestions see course website



Aligning the Approaches to IS

Aligning IS Terminologies

- Theme-Rheme (Mathesius, Firbas, Daneš, Steedman)
Theme-Rheme (Halliday)
Topic-Comment (Chomsky)
Topic-Focus (Sgall&Hajičová et al.)
Ground-Focus (Vallduví)
- Given-New within information units (Halliday)
Tail-Link within Ground (Vallduví)
Background-Focus within Theme and Rheme (Steedman)
- CB-NB (Sgall&Hajičová et al.)
- Presupposition-Focus (Chomsky, Jackendoff, Karttunen, Krifka, Rooth, etc.)
- Contrastive Topic, Focus Proper (Sgall&Hajičová et al.)
Kontrast (Vallduví and Vilkuna)

Mathesius, Firbas, Daneš	Theme	vs.	Rheme
Sgall et al.	Topic (CB) topic proper vs. contrastive topic communicative dynamism	vs.	Focus (NB) focus proper
Jackendoff, Krifka Rooth	Preupposition	vs.	Focus
Vallduví	Ground Tail vs. Link (Kontrast)	vs.	Focus Kontrast
Steedman	Theme Background vs. Focus	vs.	Rheme Background vs. Focus
Halliday	Given vs. New		Given vs. New

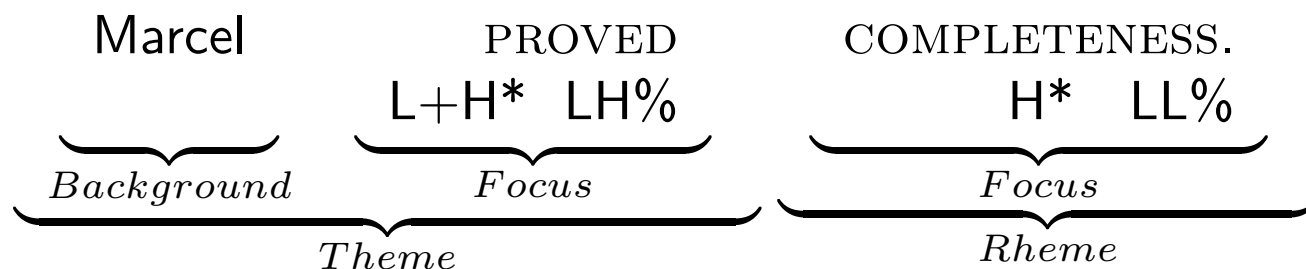
Aligning IS Terminologies

But, be aware of differences concerning:

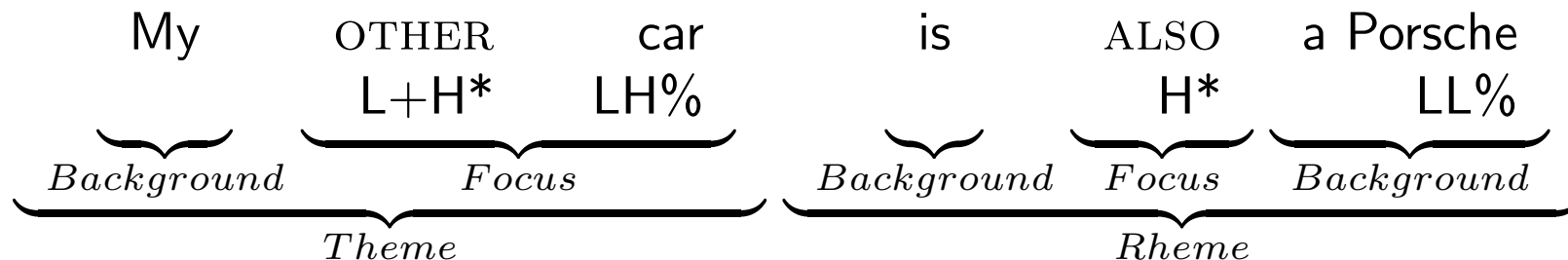
- level(s) at which IS distinctions are made, e.g., surface vs. deep structure, logical form
- flexible vs. fixed syntactic constituents, and how do IS components correspond to them
- multiple “foci”, discontinuity of IS components
- IS-boundary at main clause level vs. “deeper”
- focus projection
- degree of recursivity of IS notions (if any)
- IS in complex sentences
-

“Flexible Constituents”

- (1) I know which result Marcel predicted.
But which result did Marcel prove?

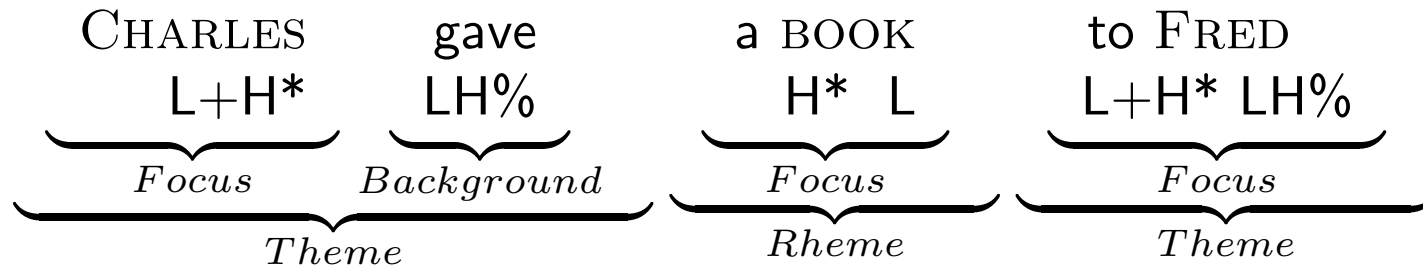


- (2) I know that THIS car is a PORSCHE.
But what is the make of your OTHER car?

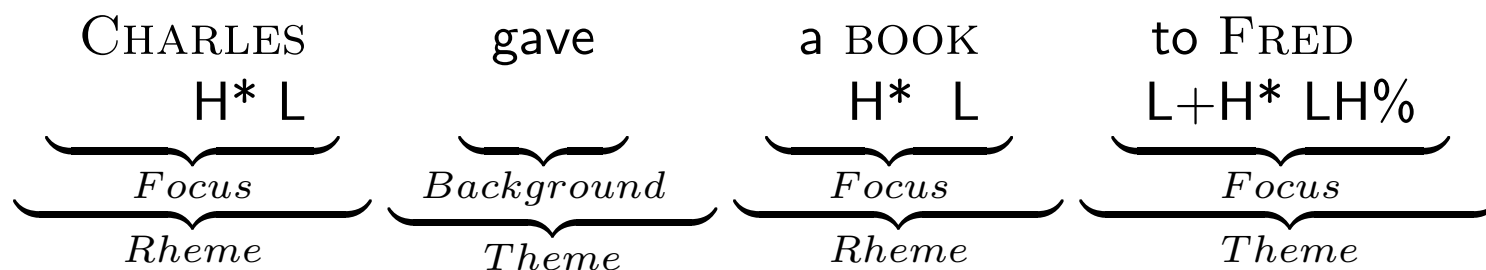


Multiple Foci

- (6) I know what Marcel gave to Harry. But what did CHARLES give to FRED?



- (7) I know what Marcel gave to Harry. But who gave what to FRED?



Deeper Embedded IS Boundary?

(Sgall et al., 1986), (Hajičová et al., 1998)(p.135):

(8) Which teacher did you meet yesterday?

(9) $\underbrace{(\text{Yesterday})_{cb} (\text{I})_{cb} (\text{met})_{cb} (\text{the teacher})_{cb}}_{\text{Topic}} \underbrace{(\text{of CHEMISTRY})_{nb}}_{\text{Focus}}.$

Steedman:

(10) $(\text{I read a book about})_{\text{Theme}} (\text{COMPLETENESS})_{\text{Rheme}} (\text{Steedman, 2000a})[\text{p.678}]$

(11) $\underbrace{\text{Yesterday I met the teacher of}}_{\text{Theme}} \underbrace{\text{CHEMISTRY}}_{\text{Focus}} \underbrace{\text{H* LL\%}}_{\text{Rheme}} .$

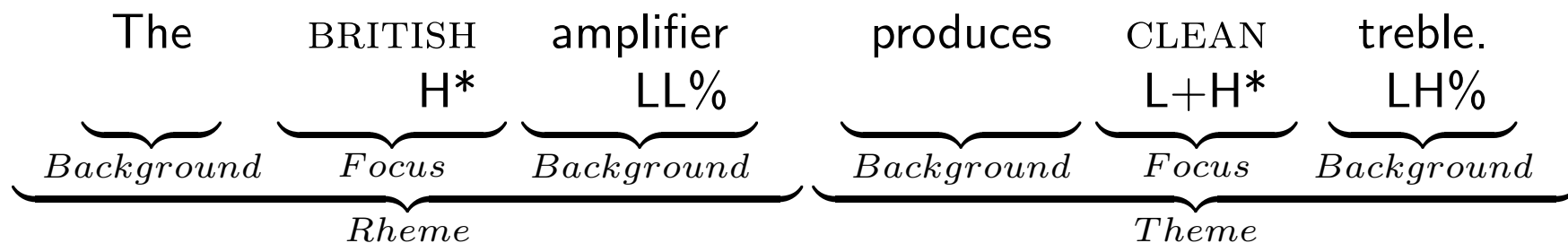
Deeper Embedded IS Boundary?

(Prevost, 1995; Prevost, 1996)

(12) I know the AMERICAN amplifier produces MUDDY treble.

(But WHAT) (produces CLEAN treble)?

L+H* LH% H* LL%



(13) I know the AMERICAN amplifier produces MUDDY treble.

(But WHICH amplifier) (produces CLEAN treble)?

L+H* LH% H* LL%

Deeper Embedded IS Boundary?

(14) Which amplifier did you buy? I bought a BRITISH amplifier.

H* LL%

(15) $\underbrace{(I)_{cb} \text{ (bought)}_{cb} \text{ (a)}}_{Topic} \underbrace{\text{ (BRITISH)}_{nb}}_{Focus} \underbrace{\text{ (amplifier)}_{cb}}_{Topic}.$

(16) $\underbrace{\text{ (Amplifier)}_{cb} \text{ (I)}_{cb} \text{ (bought)}_{cb}}_{Topic} \underbrace{\text{ (a BRITISH one)}_{nb}}_{Focus}.$

(17) $\underbrace{\text{ (Zesilovač)}_{cb} \text{ (∅)}_{cb} \text{ (jsem koupila)}_{cb}}_{Topic} \underbrace{\text{ (BRITSKÝ)}_{nb}}_{Focus}$

cf.

(18) $\underbrace{\text{ I bought an amplifier }}_{Theme} \underbrace{\text{ from BRITAIN }}_{Rheme}$

Recursivity

What is the domain of IS partitioning?

- Sentence?
- Utterance?
- Clause?
- “Basic” proposition?
- Is IS fully recursive? (e.g., (Partee, 1995))

cf., e.g., (Partee et al., 1998), (Kruijff, 2001), (Komagata 2003)

Recursivity and IS in Complex Sentences

(19) (Where do you buy wine if it's Sunday?)

If it's Sunday, we buy wine over the STATE LINE.

L+H*LH%

H*LL%

(20) Although Clyde married BERTHA, he is not HAPPY.

(21) Clyde isn't HAPPY, although he married BERTA.

what's Focus/Kontrast and what's Rheme? Any explicit indicators?

- particles, e.g., Japanese (Komagata, 1999), (Komagata 2003)
- intonation, e.g., (Steedman, 2000a)
- word order in German subordinate clause, e.g., (Günthner, 1996)
- placement of clitics in Czech
- etc.



(22) Do you see that old boat next to the *Amsterdam*?

(23) $(I)_{cb}$ $(\text{can see})_{nb}$ $(a \text{ (VERY)}_{nb} \text{ (old)}_{cb} \text{ TALLSHIP})_{nb}$ $(\text{next to the } Amsterdam)_{cb}$
 Topic Focus Topic

(24) I can see a VERY old TALLSHIP next to the *Amsterdam*
 Theme Rheme Theme
 Background Focus Background Focus Background
 H* L*+H LH% LL%

GJ:

I can see a VERY old TALLSHIP next to the *Amsterdam*
 Theme Rheme Theme
 Background Focus Background Focus Background
 H* L+H*

Summary

- There are different proposals in the literature
- How to decide?
- Do the differences actually matter?
- i.e., do they result in different predictions?
- When yes, we can test the predictions! Can't we?

- Empirical evaluation:
 - psycholinguistic experiments
 - corpus-based experiments
 - experiments with practical systems (indirect evaluation)
 - * direct evaluation: output-quality judgements
 - * indirect, task-based evaluation, e.g., success rate in dialogue system



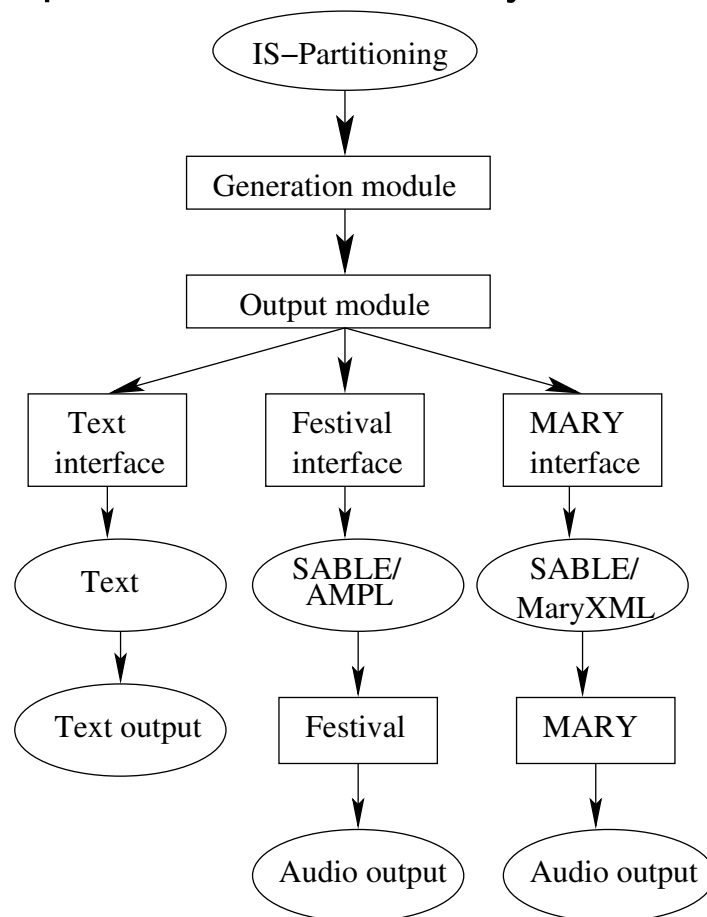
Example Practical System Evaluation

Controlling Intonation of Spoken Dialog System Output

(Kruijff-Korbayová et al., 2003)

- Within the Gothenburg Dialogue System (GoDiS), experimental dialogue system built using the TrindiKit (TRINDI, D'Homme, SIRIDUS projects)
- Determination of Theme/Rheme partitioning according to the QUD (QudTR rule): if QUD corresponds to the result of λ -abstracting over a part of the content, this part becomes the Rheme
- Determination of Focus/Background partitioning within each Theme and Rheme by determining alternatives, i.e., semantically parallel but not identical elements, w.r.t. shared commitments and/or domain knowledge

Producing synthesized output with contextually varied intonation in GoDiS



Experimental Implementation Evaluation

- Using the German TTS system Mary (Schröder and Trouvain, 2001) with various ways of intonation annotation
- Administered through a website (www.coli.uni-sb.de/cl/projects/siridus)
- Test of concept

- Experiment 1: default vs. controlled intonation using GToBI or SABLE
 - Dialogue fragments displayed on screen
 - Several turns provide context for target utterance
 - Target utterance synthesized in different versions
 - Subjects judge appropriateness of intonation in the given context
- Experiment 2: only default vs. GToBI controlled intonation
 - Subjects judge intonation without context
 - Subjects judge appropriateness of intonation in the given context

Experimental Evaluation Results

Although the results are not significant, observed tendencies correspond to expectations:

- overall average judgments worse for default than for controlled intonation
- average judgments per IS pattern also worse for default than for controlled intonation
(not much difference across patterns, though one would expect it!)
- judgments of default intonation in isolation closer to those where the context is matching with this, then to those where the context does not match
- roughly same results whether looking at absolute values of judgments or taking differences between values in isolation and in context, per subject

Experimental Evaluation Experience

- Proper (standard) evaluation methodology is lacking
- Indirect evaluation through task success / completion time does not seem suitable, because of accumulation of effects through dialogue (moreover, it would have to be Wizard of Oz, because of coverage and robustness issues)
- Direct evaluation is hard to design as a proper experiment:
 - Do subjects really take context into account?
 - Are they judging contextual appropriateness of the intonation pattern and not the quality of the synthesized output as such?
 - * Absolute judgments allow comparison of judgments across dialogues
 - * Comparative judgments could neutralize synthesis quality



Annotation of IS in Corpora

Why IS Annotation?

- Diverse and under-formalized terminology
- Lack of intuitions about the interpretation of complex IS, e.g. interaction with clause complexity, quantifiers, discourse relations, thematic structure in texts, . . .
- Can we compare and test the theories against corpus data?

IS Annotation: Problems

- Existing annotations are scarce, disparate, and theory-specific
- Existing theories are empirically inadequate: both too vague and too detailed, i.e., theoretically defined concepts are too brittle to apply to real-life data
- Lack of annotation methodology

We hope IS annotation can contribute to develop better intuitions, identifying critical issues obtain explanatorily more adequate perspectives on the realization & interpretation of IS

Desiderata for Annotation Methodology

- Theory-neutral notions
- Robustness
- Cross-linguistic applicability
- Genre/Register independence
- Multiple layers

IS annotation

- Theory-specific annotation
 - various small corpora, typically not available for reuse
 - Prague Dependency Treebank of Czech (Hajičová et al., 2003)
- Theory-neutral annotation: more basic, IS-relevant features
 - Anaphoric relations: various
 - Familiarity status at Edinburgh Uni (Nissim et al., 2004)
 - Syntactic and semantic features of referring expressions: GNOME project at Edinburgh Uni (Poesio, 2004)
 - MULI project at Saarland Uni (Baumann et al., 2004a; Baumann et al., 2004b)
 - Penn Discourse Treebank Project at UPenn (Miltsakaki et al. 2004)

TFA Annotation in PDT

Annotation of the TFA concepts in the Prague Dependency Treebank (Buráňová et al., 2000; Hajičová et al., 2003)

- data from the Czech National Corpus
- TFA is annotated in the dependency structures at the tectogrammatical level
- ordering of nodes represents communicative dynamism (deep, underlying order)
- each node is annotated with the TFA attribute:
 - T** contextually bound
 - F** contextually non-bound
 - C** contrastively bound (Partee et al., 1998)
- guidelines in Czech (cca 50 pages)

- PDT version 2.0 with several thousand sentences annotated with TFA as well as coreference will be released in the fall of 2004

Anaphora Annotation

- DRAMA (Passoneau, 1996) (available through the MATE project website)
- MUC-6 and MUC-7 (MUC Coreference Specification)
- the DRI guidelines (Carletta et al., 1997), see (www.dfki.de/dri)
- the MATE project (Poesio et al., 1999), see (www.cogsci.ed.ac.uk/~poesio/MATE/coreference.html or mate.mip.ou.dk)
- bridging references (Poesio and Vieira, 1998)
- coreference and bridging in the Heidelberg Text Corpus, MMAX tool (Müller and Strube, 2001a)

Familiarity Status Annotation

(Nissim et al., 2004)

- data: the Switchboard section of the Penn Treebank (dialog)
- annotation of referring expressions with familiarity status (Prince, 1981; Prince, 1992)
 - *brand new*: create a new discourse referent for a previously unknown entity
 - *unused*: create a new discourse referent for a known entity
 - *inferable*: create a new discourse referent for an inferable entity
 - *evoked* (textually or situationally): access an available discourse referent

	Discourse-new	Discourse-old
Hearer-new	brand new	inferable
Hearer-old	unused	evoked

- tool: the NITE workbench

NP Annotation in the GNOME Project

(Poesio, 2004)

- data from various sources
- Annotation of nominal referring expressions with syntactic and semantic features relevant for NP generation
 - the semantic attributes include, e.g., animacy, ontological status, countability, quantification and generic vs. specific reference
- detailed guidelines available from the GNOME project website

Penn Discourse Treebank Project

(Miltsakaki et al. 2004)

- data: the Penn Treebank
- annotation of explicit and implicit discourse connectives and their arguments (Webber and Joshi, 1998)
- tool: WordFreak

Multilayer IS Annotation in the MULI project

(Baumann et al., 2004a; Baumann et al., 2004b)

- data from the WSJ section of the PTB (English) and from the Negra/TIGER Treebank (German)
- annotation of markables at multiple layers, with layer-specific features (syntax, discourse, prosody)
- English: 7k words/320 sentences; German 3.5k words/250 sentences
- Multi-layer stand-off annotation with shared base data; layer-specific tools, translation into shared XML format

MULI: Syntax Layer Annotation

- IS-relevant aspects of realization:
 - Positioning
 - Ordering (words, phrases, clauses)
 - Marked syntactic constructions
 - Morphological marking (in some languages)
 - Definiteness marking (in some languages)
- Treebank data already available
- annotation of additional features: presence of marked syntactic constructions in clauses
- tool: XMLspy

MULI: Layer Annotation

- recording of read text, just one speaker, only the German corpus
- annotation of intonation following GToBI (Grice et al., to appear)
 - Position and type of boundary tones
 - Position and type of accents
 - Position and size of phrase boundaries
- tool: EMU

MULI: Discourse Layer Annotation

- Annotation of expressions with discourse referent properties
 - Semantic type and sort
 - Denotation characteristics:
 - * Extensional vs. intensional reference
 - * Unique, existential or variable determination
 - * Countability and quantification
 - Familiarity status (Prince, 1981)
 - Anaphoric relations: Coreference and various types of bridging
- guidelines (cca 30 pages)
- tool: MMAX (Müller and Strube, 2001b)

How Much Work is it?

- It's hard: . . .
 - Also designing annotation methodology and schemes
- And quite time-consuming
 - Intonation: 30 min/sent. (incl. discussion)
 - Discourse: 10 min/sent. (create markables and links, assign properties)
 - Syntax: 5 min/sent. –only addt'l annotation (clause segmentation, classification); Treebank available



Wrapping Up

Summary and Conclusions

- IS is an important aspect of meaning at the interface between utterance and discourse
- a theory relating IS and DS is essential for accurate NL processing
- formal accounts start to emerge, some embodied into practical systems
- many questions concerning IS partitioning and its realization in different languages still open
- Further research topics:
 - further systematization of terminologies
 - formalization and computational modeling&testing
 - empirical and corpus-based studies
 - cross-linguistic investigations and multilingual applications

Summary

Day 1 Information Structure as an Inherent Aspect of Sentence Meaning.

Day 2 The Praguian Topic-Focus Articulation. Saliency. IS-sensitive Saliency Modeling in Analysis and Generation.

Day 3 Vallduví's Information Packaging. File-Change Semantics of IS. IS in Word Order Generation. Halliday's Thematic- vs. Information-Structure. Daneš's thematic sequences.

Day 4 Steedman's Two Dimensions of IS. Alternative-set Semantics of IS. IS and Intonation. IS and Turn-Taking, Gesture and Gaze in Multimodal Dialog.

Day 5 Wrapping Up and Looking Out. Aligning IS-Approaches. Need for Empirical Studies to Test Theories. Practical Evaluation. Corpus Annotation.



That's it folks!



References

- Stefan Baumann, Caren Brinckmann, Silvia Hansen-Schirra, Geert-Jan Kruijff, Ivana Kruijff-Korbayova, Stella Neumann, Erich Steiner, Elke Teich, and Hans Uszkoreit. 2004a. The muli project: Annotation and analysis of information structure in german and english. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics, College Park MD, June ()*, pages 41–48.
- Stefan Baumann, Caren Brinckmann, Geert-Jan Kruijff Silvia Hansen-Schirra, Ivana Kruijff-Korbayova, Stella Neumann, and Elke Teich. 2004b. Multi-dimensional annotation of linguistic corpora for investigating information structure. In *Frontiers in Corpus Annotation 2004, NAACL/HLT Conference Workshop, Boston*.
- Eva Buráňová, Eva Hajičová, and Petr Sgall. 2000. Tagging of very large corpora: Topic-focus articulation. In *Proceedings of COLING'2000*, pages 139–144.
- Jean Carletta, Nils Dahlbäck, Norbert Reithinger, and Marylin A. Walker. 1997. Standards for dialogue coding in natural language processing. Report on the dagstuhl seminar, Discourse Resource Initiative, February 3–7.
- M. Grice, S. Baumann, and R. Benz Müller. to appear. German intonation in autosegmental phonology. In S.-A. Jun, editor, *Prosodic typology*. OUP.
- Susanne Günthner. 1996. From subordination to coordination? verb-second position in German causal and concessive constructions. *Pragmatics*, 6(3):323–356.



- Eva Hajičová, Barbara H. Partee, and Petr Sgall. 1998. *Topic-Focus Articulation, Tripartite Structures, and Semantic Context*. Studies in Language and Philosophy. Kluwer Academic Publishers, Dordrecht, Boston, London.
- Eva Hajičová, Petr Sgall, and Eva Buráňová. 2003. Topic-focus articulation and degrees of salience in the prague dependency treebank. In *Formal Approaches to Function in Grammar. In honor of Eloise Jelinek, Arizona*, pages 165–177. John Benjamins.
- Nobo Komagata. 1999. *A Computational Analysis of Information Structure Using Parallel Expository Texts in English and Japanese*. Ph.D. dissertation, also as IRCS Technical Report 99-07 and UMI 9937745, University of Pennsylvania, Philadelphia.
- Ivana Kruijff-Korbayová, Stina Ericsson, Kepa Joseba Rodríguez, and Elena Karagjosova. 2003. Producing contextually appropriate intonation is an information-state based dialogue system. In *Proceedings of the 10th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 227–234. ACL.
- Geert-Jan M. Kruijff. 2001. *A Categorical-Modal Logical Architecture of Informativity: Dependency Grammar Logic and Information Structure*. Ph.d. dissertation, Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic.
- Christoph Müller and Michael Strube. 2001a. Annotating anaphoric and bridging relations with MMAX. In *Proceedings of the 2nd SIGdial Workshop on Discourse and Dialogue*, pages 90–95, Aalborg, Denmark, 1–2 September.
- Christoph Müller and Michael Strube. 2001b. MMAX: A tool for the annotation of multi-modal corpora. In *Proceedings of the 2nd IJCAI Workshop on Knowledge and Reasoning in Practical Dialogue*, pages 45–50, Seattle, Washington, 5 August.
- Malvina Nissim, Shipra Dingare, Jean Carletta, and Mark Steedman. 2004. An annotation scheme for information



status in dialog. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics, College Park MD, June ()*, pages 41–48.

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

Barbara Partee. 1995. Allegation and local accommodation. In Barbara H. Partee and Petr Sgall, editors, *Discourse and Meaning*, pages 65–86. John Benjamins Publishing Company.

Rebecca Passoneau. 1996. Instructions for applying discourse reference annotation for multiple applications (DRAMA). draft, December 20.

Massimo Poesio and Renata Vieira. 1998. A corpus-based investigation of definite description use. *Computational Linguistics*, 24(2):183–216. url.

Massimo Poesio, Florence Bruneseaux, Sarah Davies, and Laurent Romary. 1999. The MATE meta-scheme for coreference in dialogue in multiple languages. In Marilyn Walker, editor, *Proceedings of the workshops on "Towards Standards and Tools for Discourse Tagging" at the 37th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 65–74, University of Maryland, June.

Massimo Poesio. 2004. Discourse annotation and semantic annotation in the GNOME corpus. In *Proceedings of the ACL Workshop on Discourse Annotation*, Barcelona, Spain.

Scott Prevost. 1995. *A Semantics of Contrast and Information Structure for Specifying Intonation in Spoken Language Generation*. Ph.D. dissertation, IRCS TR 96-01, University of Pennsylvania, Philadelphia.

Scott Prevost. 1996. An information structural approach to spoken language generation. In *Proceedings of 34th Annual Meeting of the Association for Computational Linguistics, Santa Cruz*.

Ellen Prince. 1981. Toward a taxonomy of given-new information. In Peter Cole, editor, *Radical Pragmatics*, pages

223–256. Academic Press.

Ellen Prince. 1992. The ZPG letter: subjects, definiteness, and information status. In Sandra Thompson and William Mann, editors, *Discourse Description: diverse analyses of a fund raising text*, pages 295–325. John Benjamins.

Marc Schröder and Jürgen Trouvain. 2001. The german text-to-speech synthesis system MARY: A tool for research, development and teaching. In *The Proceedings of the 4th ISCA Workshop on Speech Synthesis, Blair Atholl, Scotland*.

Petr Sgall, Eva Hajičová, and Jarmila Panevová. 1986. *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects*. D. Reidel Publishing Company, Dordrecht, Boston, London.

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

Mark Steedman. 2000b. *The Syntactic Process*. M.I.T. Press, Cambridge, MA.

Bonnie Webber and Aravind Joshi. 1998. Anchoring a lexicalized tree-adjoining grammar for discourse. In *Coling/ACL Workshop on Discourse Relations and Discourse Markers*, pages 86–92, Montreal, Canada. url.

Bonnie Webber, Alistair Knott, Matthew Stone, and Aravind Joshi. 1999. Discourse relations: A structural and presuppositional account using lexicalised tag. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics, College Park MD, June*, pages 41–48, San Francisco, CA. Morgan Kaufmann.