Modeling Information Structure
in Discourse and Dialogue Processing

Ivana Kruijff-Korbayová

korbay@coli.uni-sb.de

http://www.coli.uni-saarland.de/~korbay/

Nordic Graduate School,
Helsinki, April 3-8 2005

Lecture 3 Outline

- Steedman's two dimensions of IS
- Alternative-set based semantics of IS
- IS and intonation (in English)
- Intonation assignment in answers to questions
- Assignment in monologue generation
- Intonation assignment for TTS
- Intonation assignment in the FLIGHTS dialogue system (!!!)
- Gestures, turn-taking and eye-gaze in multimodal interaction

Steedman's Two Dimensions of IS

Steedman's IS Partitioning

(Steedman, 2000b; Steedman, 2000a) distinguishes two dimensions of IS within a sentence:

Theme-Rheme partitioning reflects an *aboutness* relation, i.e., the Rheme is semantically predicated over the Theme. This dimension connects the utterance to the rest of the discourse.

Background-Focus partitioning within Theme and Rheme reflects an abstract notion of "kontrast" between alternatives available in the discourse context, against which the Theme and Rheme of the actual utterance are cast. Words whose interpretation contributes to distinguishing Theme/Rheme from alternatives belong to Focus, other words belong to Background.

The Semantics of IS

- Semantics of IS in terms of selecting one member from a *presupposed set of alternatives* (Steedman, 2000a), following (Rooth, 1992; **?**)
  - Theme presupposes a *Rheme-alternative set*, i.e., a set of alternative propositions that could possibly answer the corresponding question in the given context; Rheme then restricts the Rheme-alternative set to a singleton
  - Theme also presupposes a *Theme-alternative set*, i.e. a set of alternative questions; Focus within Theme then restricts the Theme-alternative set to a singleton
- These are pragmatic presuppositions that the relevant alternative set(s) be available in the context. They can get bound or accommodated.
- The systematic recognition of the alternative sets, and their maintenance as a discourse progresses are open research issues.

The Semantics of IS

(1)   Marcel   PROVED   COMPLETENESS.

$$\underbrace{\underbrace{Background}\ \underbrace{Focus}}_{Theme}\ \underbrace{\underbrace{Focus}}_{Rheme}$$
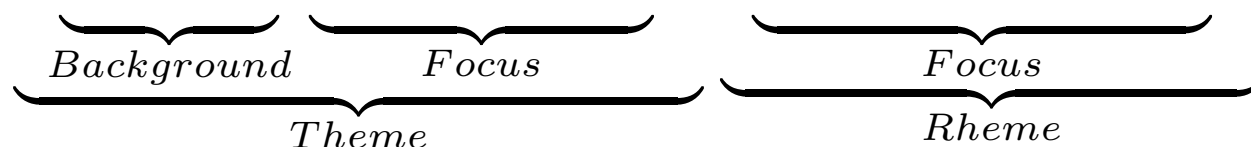
(2) $prove'\ completeness'\ marcel'$

(3) $\exists x.\ \star\, prove'\ x\ marcel'$

(4) $\{\ prove'completeness'marcel',\ prove'decidability'marcel',$
      $prove'soundness'marcel'\ \}$

(5) $\{\ \exists x.\ prove'\ x\ marcel',\ \exists x.\ predict'\ x\ marcel'\ \}$

(6)   I know which result Marcel predicted. But what did Marcel prove?

The Semantics of IS

(7)   MARCEL   proved   COMPLETENESS.

$$\underbrace{\underbrace{\text{Focus}}\_{\text{Theme}}}\quad \underbrace{\underbrace{\text{Background}}\quad\underbrace{\text{Focus}}}\_{\text{Rheme}}$$

(8) $prove'\ completeness'\ marcel'$

(9) $\exists P.\ P \star marcel'$

(10) $\{\ prove'completeness'marcel',\ prove'decidability'marcel',$
       $predict'soundness'marcel'\ \}$

(11) $\{\ \exists P.Pmarcel',\ \exists P.Psusan'\ \}$

(12)   I know Susan proved soundness. But what did Marcel do?

The Semantics of IS

(13)   MARCEL   proved   COMPLETENESS.

$$\underbrace{\underbrace{\text{Focus}}}_{Rheme} \quad \underbrace{\underbrace{Background} \quad \underbrace{Focus}}_{Theme}$$

(14) $prove'\ completeness'\ marcel'$

(15) $\exists x.\ prove'completeness'x$

(16) $\{\ prove'completeness'marcel',\ prove'completeness'susan'\ \}$

(17) $\{\ \exists x.prove'\ completeness'\ x,\ \exists x.prove'\ soundness'\ x,$
$\exists x.prove'\ decidability'\ x\ \}$

(18)   I know Susan proved soundness. But who proved completeness?

IS and Intonation

(Steedman, 2000b; Steedman, 2000a) proposes a compositional account of the semantics of tones for English, cast in CCG

- Theme/Rheme partitioning determines overall intonation pattern

  - Theme and Rheme as one intonation phrase each (boundary between)
  - Theme-accents: L+H*, L*+H   (prototypical Theme-tune: L+H*LH%)
  - Rheme-accents: H*, L*, H*+L, H+L* (prototypical Rheme-tune: H*LL%)

- Background/Focus partitioning Determines placement of pitch accents

  - Focus: (words) marked by pitch accent
 - Background: (words) without pitch accent

(ToBI intonation notation (Beckman and Hirschberg, 1999).)

IS and Intonation

(19) Marcel PROVED   COMPLETENESS.
              L+H* LH%   H*   LL%

$\underbrace{\qquad}_{Background}$ $\underbrace{\qquad}_{Focus}$ $\underbrace{\qquad}_{Focus}$

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

(20) MARCEL   proved   COMPLETENESS.
      L+H*LH%   H*LL%

$\underbrace{\qquad}_{Focus}$ $\underbrace{\qquad}_{Background}$ $\underbrace{\qquad}_{Focus}$

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

(21) MARCEL   proved   COMPLETENESS.
      H*L   L+H*LL%

$\underbrace{\qquad}_{Focus}$ $\underbrace{\qquad}_{Background}$ $\underbrace{\qquad}_{Focus}$

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

More Examples

(22)   I know that Marcel likes the man who wrote the musical.
       But who does he ADMIRE?

Marcel	ADMIRES	the woman who	DIRECTED	the musical.
	L+H* LH%		H*	LL%
$\underbrace{}_{Background}$	$\underbrace{}_{Focus}$	$\underbrace{}_{Background}$	$\underbrace{}_{Focus}$	$\underbrace{}_{Background}$

Background — *Focus* — *Theme* | *Background* — *Focus* — *Background* — *Rheme*
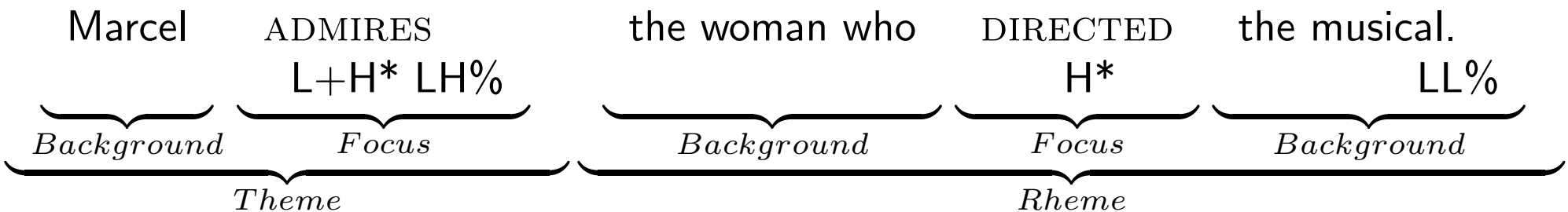
- the Background/Focus partitioning of this Rheme is supported just in case all individuals considered have something to do with the musical, and the property of directing it uniquely identifies one such individual (Prevost and Steedman, 1994; Prevost, 1995; Steedman, 2000b)
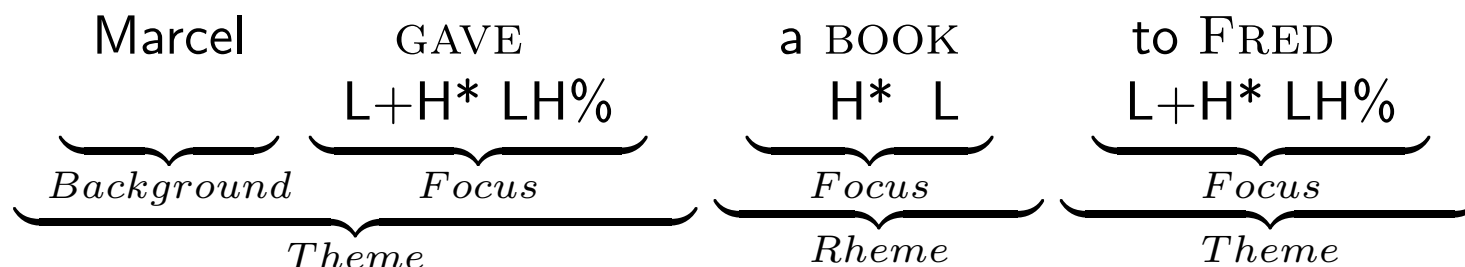
More Examples

Marcel	ADMIRES	the woman who	DIRECTED	the musical.
	L+H* LH%		H*	LL%

$\underbrace{Background}$ $\underbrace{Focus}$ $\underbrace{Background}$ $\underbrace{Focus}$ $\underbrace{Background}$

$\underbrace{\qquad\qquad Theme \qquad\qquad}$ $\underbrace{\qquad\qquad Rheme \qquad\qquad}$
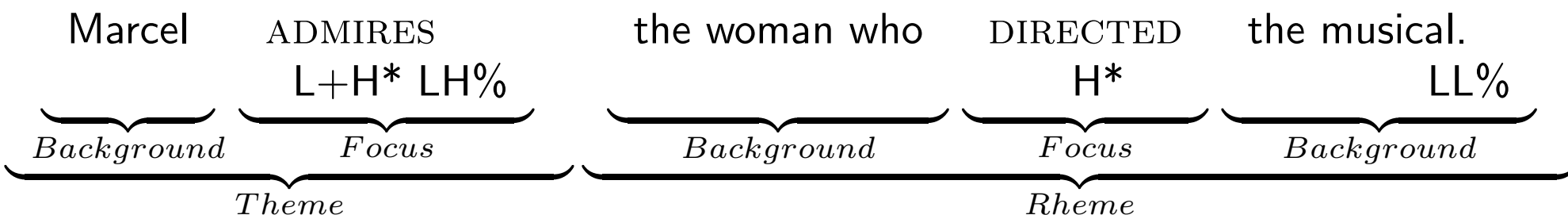
(23) $admire'\ woman_1'\ marcel'$

(24) $\exists x.\ \star\,admire'\ x\ marcel'$

(25) { $admire'woman_1'marcel'$, $admire'woman_2'marcel'$, $admire'man_1'marcel'$ }

(26) { $\exists x.\ admires'\ x\ marcel'$, $\exists x.\ likes'\ x\ marcel'$ }

More Examples

(27) I know what Marcel SOLD to Harry. But what did he GIVE to FRED?

$$\underbrace{\underbrace{\text{Marcel}}_{Background} \quad \underbrace{\underset{\text{L+H* LH\%}}{\text{GAVE}}}_{Focus}}_{Theme} \quad \underbrace{\underbrace{\underset{\text{H* L}}{\text{a BOOK}}}_{Focus}}_{Rheme} \quad \underbrace{\underbrace{\underset{\text{L+H* LH\%}}{\text{to FRED}}}_{Focus}}_{Theme}$$

(28) $give'\ fred'\ book'\ marcel'$

(29) $\exists x.\ \star\, give'\ \star\, fred'\ x\ marcel'$

(30) $\{\ give'\ fred'\ book'\ marcel',\ give'\ fred'\ record'\ marcel',$
       $give'\ fred'\ biscuit'\ marcel'\}$

(31) $\{\ \exists x.\ give'\ fred'\ x\ marcel',\ \exists x.\ sell'\ fred\ x\ marcel',$
       $\exists x.\ give'\ harry\ x\ marcel',\ \exists x.\ sell'\ harry\ x\ marcel'\ \}$

IS and Intonation

further examples, cf. (Steedman, 2000a)[p. 662 and on], incl. all-theme, ownership by hearer vs. speaker

Intonation Assignment in various Applications

IS-Based Assignment of Intonation in Answers to Questions

- IS used to control intonation synthesized spoken output

- (Prevost and Steedman, 1993) IS of question determines the IS of the answer

- Theme/Rheme determination:
  - rheme of the question determines the theme of the answer

- Focus determination:
  - terms focused in question are focused in asnwer
  - term instantiating question variable is also focused
  - for more complex rhemes, only *new* elements are focused

Example

(32) *I know that widgets contain cogs,*

    but what parts do WODGETS *include?*
       L+H* LH%   H*   LL%

   prop: $s : \lambda x[part(x)\&include(\star wodgets, x)]$
   theme: $s : \lambda x[part(x)\&include(\star wodgets, x)]$ /
               $(s : include(\star wodgets, x)/np : x)$
   rheme: $s : include(\star wodgets, x)/np : x$

     prop: $s : include(\star wodgets, \star sprockets)$
(33) theme: $s : include(\star wodgets, x)/np : x$
     rheme: $np : \star sprockets$

   WODGETS *include* SPROCKETS.
    L+H*   LH%   H* LL%

IS-Based Assignment of Intonation in Answers to Questions

- (Prevost and Steedman, 1994): IS determination in answer not from question alone but also from discourse model (database)

- Theme/Rheme determination:

  - rheme of the question determines the theme of the answer

- Focus determination:

  - terms focused in question's rheme are focused in answer's theme
  - rheme-focus in answer determined from alternative sets in the database

Constructing Rheme-Alternative Sets

(Prevost and Steedman, 1994):

Given database D, object x and a set of properties P that uniquely describe x:

1. construct a set of objects, A, (and their referring properties) which can be considered alternatives to x w.r.t. D

2. restrict A by properties of objects mentioned in theme $\rightarrow A'$

3. mark as contrastive those properties of x in P that exclude some alternatives from A'

Example . . .

IS-Based Assignment of Intonation in Text Generation

(Prevost, 1996)

- Content and text planning: determine a sequence of propositions about an object and the rhetorical relations, segment each proposition into theme/rheme

  - discourse model contains previous themes and rhemes (ISstore)
  - to determine theme, search for most recent match, prefer theme-continuation
  - determine rheme as complement of theme

- Sentence planning: determination of realization, focus assignment

  - each new (not mentioned) property or discourse entity get focus
  - contrasting elements get focus

Assigning Contrastive Focus

Step 1: Determine contrasting propositions

- containing 2 contrasting pairs of entities or 1 pair of contrasting entities and contrasting functors
- discourse entities are contrasting when they are alternatives w.r.t. `isa` in DB

Step 2: Contrastive focus algorithm
Given: object x, properties P, alternatives A:

1. restrict A to objects mentioned in discourse $\rightarrow A'$

2. for each property p in P, include p in set of contrasting propertied P_c iff p excludes some object from A'

IS-Based Intonation in TTS

- Previous work: acenting affected by "givenness" (Hirschberg 1990), (Hirschberg, 1993)

- (Hiyakumoto et al. 1997):

  - combine first mention and contrastiveness as reasons for accenting
  - use of WordNet in givenness and contrast determination: to identify sets of synonyms and contrasting words for open-class words (nouns, verbs, adjectives, adverbs)
  - determine theme/rheme in propositional constituents by heuristics applied to pre- and post-verbal and verb-complex material, and considering presence of focus within it

Summary

- IS and intonation (only); formalization in CCG

- Theme/Rheme determined by (a) question, (b) linking in text, (c) heuristics

- Focus is determined from (a) question, (b) discourse newness, (c) contrast w.r.t. alternatives in discourse model

- Discourse model contains entities, propositions, and themes and rhemes

- Theme/Rheme partitioning at proposition level, i.e., recursive in complex utterances

- IS assignment in questions differs from Hoffman

IS in Multimodal Interaction

- Appropriate and synchronized speech, intonation, facial expressions and hand gestures (Pelachaud et al., 1998)

- Integrating turn-taking and IS provides better explanation for gaze behavior (Cassell et al., 1999)

- Generation of either speech, gesture or combination of both as a function of IS status and surprise value of a discourse entity (Cassell et al., 2000)

- Various researches have observed that distance, posture shifts and other body movements seem to accompany changes in the topic or social relationship (cf. Cassell et al. 2001)

IS and Gestures

Gilbert and George (Pelachaud et al., 1998)

- Some facial expressions are automatically generated according to intonation (cf. also the COMIC project (**?**))
- Head nods and look-toward listener punctuate accented and emphasized items
- Iconic and metaphoric gestures (i.e., representing something) are generated for
  - rhematic verbal elements (roughly, information not yet spoken about)
  - hearer new references
  provided that the semantic content can receive such a gesture (e.g., spatial)
- Beat gestures are generated
  - otherwise
  - to accompany discourse new definite references
- Duration of intonation phrases is used to time gestures

Summary and Conclusions

- Contextually apropriate system output needed in a variety of applications: dialog, monolog; written, spoken, multimodal; etc.

- Contextually appropriate realization requires account of IS to motivate realization choices by various means in a uniform way

- Modeling of the interplay of linguistic IS-realization choices in practical systems so far largely not done, i.e., concentration on either intonation or word order

- Multimodal systems try to combine linguistic and non-linguistic signals, based on empirically observed correlations

References

Mary Beckman and Julia Hirschberg. 1999. The tobi annotation conventions. Ms. Ohio State University.

Justine Cassell, Obed. E. Torres, and Scott Prevost. 1999. Turn taking vs. discourse structure: How best to model multimodal conversation. In Y. Wilks, editor, *Machine Conversations*. Kluwer.

Justine Cassell, Matthew Stone, and Hao Yan. 2000. Coordination and context-dependence in the generation of embodied conversation. In *Proceedings of the INLG Conference*.

Jonathan Ginzburg. 1996. Interrogatives: Questions, facts and dialogue. In Shalom Lappin, editor, *The Handbook of Contemporary Semantic Theory*, chapter Chapter 15, pages 385–422. Blackwell Publishers.

M. Grice, S. Baumann, and R. Benzmüller. to appear. German intonation in autosegmental phonology. In S.-A. Jun, editor, *Prosodic typology*. OUP.

Julia Hirschberg. 1993. Pitch accent in context: Predicting intonational prominence from text. *Artificial Intelligence*, (63):305–340.

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.

Catherine Pelachaud, Justine Cassel, Norman Badler, Mark Steedman, Scott Prevost, and Mathew Stone. 1998. Synthesizing cooperative conversation. In Harry Bunt and R. J. Beun, editors, *Multimodal Human-Computer*

Communication - Systems, Techniques and Experiments, Springer Notes in Artificial Intelligence 1374, pages 68–88. Springer Verlag.

Scott Prevost and Mark Steedman. 1993. Generating contextually appropriate intonation. In *Proceedings of the Sixth Conference of the European Chapter of the Association for Computational Linguistics, Utrecht*, pages 332–340, San Francisco, CA. Morgan Kaufmann.

Scott Prevost and Mark Steedman. 1994. Specifying intonation from context for speech synthesis. *Speech Communication*, 15:139–153.

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*.

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

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*.

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.