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 1 Outline

- Information Structure Partitioning
- Question test for IS
- IS Realization Means
- IS Semantics
- Meaning Differences due to IS
- IS and Discourse Dynamics
- Course Outline

Motivation

Motivation

(1) Sign in London underground: *Dogs must be carried.*

(2) Sign in a synagogue: *Hats must be worn.*

How are they (intended to be) read and understood?

(1′) Dogs must be CARRIED.
If you have a dog, you must carry it.

(2′) HATS must be worn.
You must wear a hat.

(capitals denote intonation center)

Motivation

(3) Sign in London underground: *Dogs must be carried.*

Can be read in (at least) two different ways: (Halliday, 1970)

(4) a. Dogs must be CARRIED.
b. DOGS must be carried.

and there is a difference in meaning:

(3′) a. If you have a dog, you must carry it.
b. What you must do is carry a dog. (i.e., not allowed to enter without)

- The same or similar meanings can be realized in various ways.
- Different languages may use different ways.

Motivation

(5) German:

a. Hunde mussen GETRAGEN werden.
Dogs must carried be

b. Es mussen HUNDE getragen werden.
It_{explet} must_{3pl} dogs_{nom} carried_{part} be_{inf}.

(6) Czech:

a. Psi se musí NÉST.
Dogs_{nom} refl must_{3pl} carry_{inf}

b. Musí se nést PES.
Must_{3sg} refl carry_{inf} dog_{nom}

Motivation

Czech newspaper 1990: (Hajičová)

(7) Dobrá zpráva je, že Češi udělali revoluci.
Good news is that Czechs made revolution.
The good news is that the Czechs made a revolution.

Špatná zpráva je, že revoluci udělali Češi.
Bad news is that revolution made Czechs.

The bad news is that the/a revolution was made by the Czechs.
(or: . . . it was the Czechs who made the/a revolution)

Motivation

Dialog with an intelligent-home application: (Kruijff-Korbayová et al., 2003)

(8) U: What devices are there in the house?
S: There is a STOVE in the KITCHEN, a RADIO in the kitchen and a radio in the BATHROOM.
U: What is the status of the radios?
S: The radio in the KITCHEN is ON. The radio in the BATHROOM is OFF.
U: Which devices are on?
S: The radio in the KITCHEN is on. The STOVE in the kitchen is on.

- The same (default) realization would not be appropriate in all cases.
- Wrong realization maybe be disturbing or misleading.
- The realization of system output needs to be controlled according to context.

Information Structure Partitioning

What is Information Structure?

- IS comprises the *utterance-internal structural and semantic properties reflecting the relation of an utterance to the discourse context*, in terms of the discourse status of its contents, the actual and attributed attentional states of the discourse participants, and the participants' prior and changing attitudes (knowledge, beliefs, intentions, expectations, etc.) (Kruijff-Korbayová and Steedman, 2003)

- IS is represented as a partitioning of utterance meaning w.r.t. how parts of an utterance *depend on* and *affect* the context

- IS is reflected in/by the surface realization of the utterance

Two Dimensions of Information Structure

- A partitioning of utterance meaning into what the speaker means to address and what the speaker means to say about it:

  Theme the part which relates it to the purpose of the discourse and anchors the content to the context (i.e., what speaker and hearer are attending to); "*point of departure*"

  Rheme the part which advances the discourse, i.e., adds or modifies some information "*about the Theme*"

- A partitioning of utterance meaning according to which parts contribute to distinguishing the actual content from alternatives in the discourse context:

  Background the non-discriminating part(s), same across alternatives

  Focus the discriminating part(s), different from alternatives

Approaches to IS

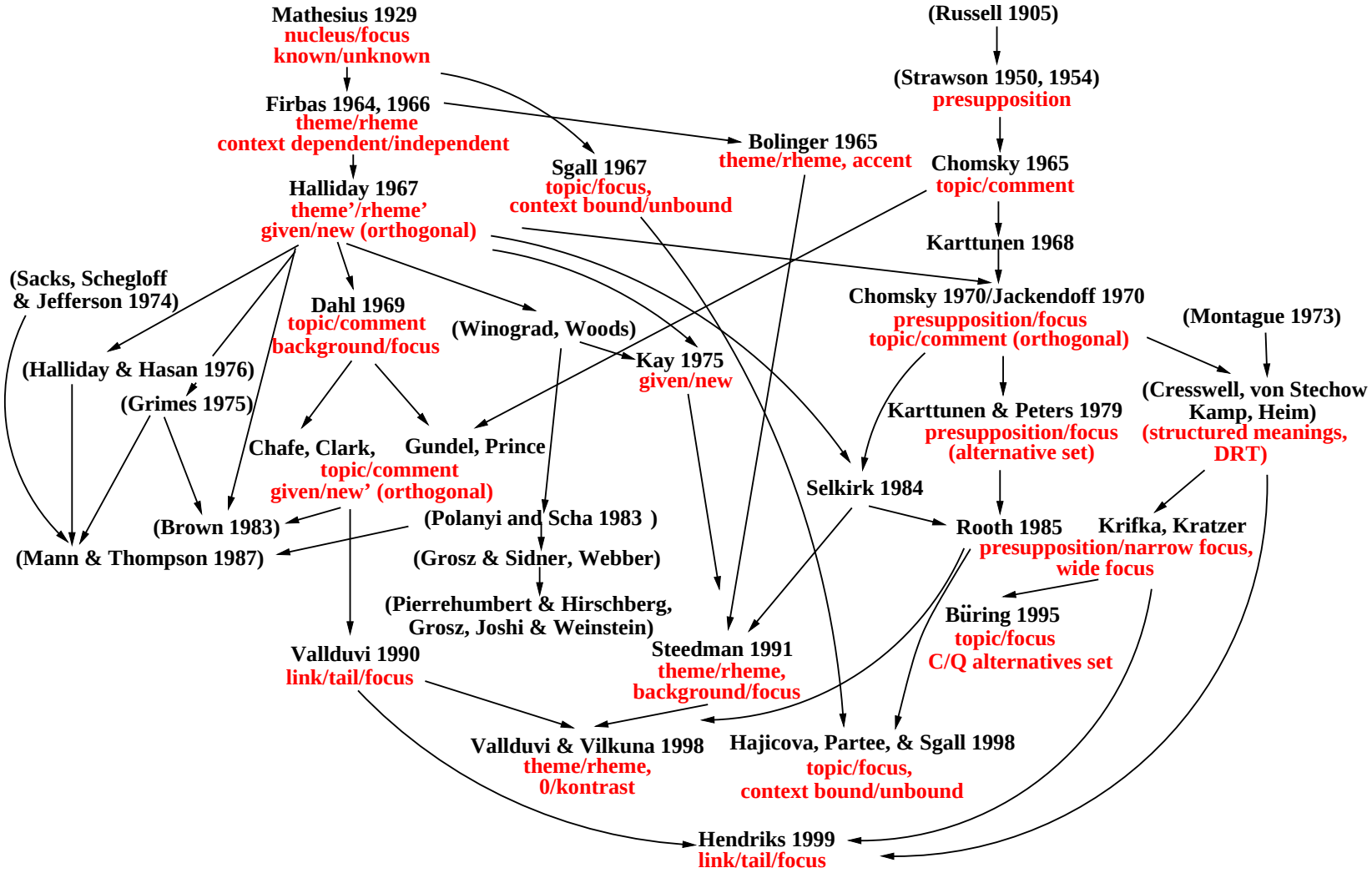
Theories of IS differ in

- how they define the partitioning more precisely

- which of the two dimensions they concentrate on
  when they consider both, how they combine them

  – "embedded": Sgall, Hajičová et al. (CB/NB deeper within Topic and Focus; contrastive Topic, Focus proper); Valdduví (Link/Tail in Ground); Steedman (Background-Focus within Theme and Rheme)

  – "orthogonal": Halliday (Thematic Structure vs. Information Stucture); Chomsky, Jackendoff . . . (Topic-Comment vs. Presupposition-Focus)


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

Question Test for IS

- Operational test of the appropriateness of a particular IS w.r.t. given context
  - Question represents the context
  - Theme reflects the question, and Rheme is what answers the question

(9) Q. What does John do? A. $\underbrace{\text{John}}_{Theme}$ $\underbrace{\text{writes NOVELS}}_{Rheme}$.

(10) Q. Who writes NOVELS? A. $\underbrace{\text{JOHN}}_{Rheme}$ $\underbrace{\text{writes novels}}_{Theme}$.

- Exchanging (9.Q) and (10.Q) yields incoherent Q-A pairs.

- Do not confuse with Q-A pairs in a natural dialogue!

Question Test: Focus Projection

- Phonological focus: word(s) carrying pitch accent

  (11) John flew from London to PARIS.

- Semantic focus: *narrow* vs. *broad* projection of phonological focus

  (12) Where did John fly (to) from London?
  John flew from London to PARIS. (narrow)

  (13) What flight did John make?
  John flew from London to PARIS. (broad 1)

  (14) What did John do?
  John flew from London to PARIS. (broad 2)

  (15) What happened?
  John flew from London to PARIS. (broad 3)

(16) From which place did John fly to Paris?
John flew from LONDON to Paris. (narrow)

(17) Who flew from London to Paris?
JOHN flew from London to Paris. (narrow)

(18) What happenned to Nixon?
Nixon DIED. (narrow)

(19) Who died?
NIXON died. (narrow)

(20) What happenned?
a. NIXON died. (broad)
b. Nixon DIED. (?)

IS Realization Means

IS Realization

- There are various means of IS realization.
- Various means can be used also in combination, and they interact.
- Different languages employ and combine the means differently, depending on their typological characteristics.
- The means (strategies):
  - Intonation: placement and type of pitch accents and boundary tones
  - Word order: ordering of constituents within a clause, ordering of clauses
  - Syntactic structure: e.g., fronting ("topicalization"), lef/right dislocation, there-insertion, it-cleft, wh-cleft, dative-shift, passivization, etc.
  - Morphological marking; e.g., particle 'wa' in Japanese
  - Ellipsis (deletion)
- Marked vs. unmarked (default, "out-of-the-blue")

IS Realization Means: Intonation

(Steedman, 2000) for English; similarly (Uhmann, 1991; Fery, 1993) for German:

- 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 on particular words
  - Focus: marked by pitch accent
  - Background: unmarked by pitch accent

IS Realization Means: Intonation

Intonation is (said to be) the predominant means of IS realization in English. Example from (Kruijff-Korbayová et al., 2003)

(21) U: What devices are there in the house?

There is a STOVE in the KITCHEN, a RADIO in the kitchen
H* H* LH% H* LH%
and a radio in the BATHROOM.
H* LL%

U: What is the status of the kitchen devices?

S: The STOVE in the kitchen is ON.
L+H* L% H*LL%
The RADIO in the kitchen is OFF.
L+H* L% H*LL%

IS Realization Means: Word Order

- "Default" progression is from "old" to "new" information:
  - Theme before Rheme
  - Background before Focus (at least within Rheme)

- Different ordering typically motivated in/by discourse context, e.g., Rheme before Theme (*subjective ordering* in (Firbas, 1971; Firbas, 1992))

- But: Modulo syntactic constraints!
  - Typological characteristic of a language as SVO, VSO, OVS, etc.
  - Focus-position before verb (e.g., Hungarian, Turkish)
  - Verb-secondness, clitic-placement, heavy-constituent shift, adjectives before head-noun, etc.

IS Realization Means: Word Order

WO is (said to be) the predominant IS realization means in Czech ("free WO").

(22) What happened? $\underbrace{\text{Češi udělali revoluci.}}_{Rheme}$ (The Czechs made a revolution)

(23) Who made a revolution?

$\underbrace{\text{Revoluci udělali}}_{Theme}$ $\underbrace{\text{Češi.}}_{Rheme}$

$\underbrace{\text{Češi}}_{Rheme}$ $\underbrace{\text{udělali revoluci.}}_{Theme}$

(24) What about the Czechs?

$\underbrace{\text{Češi}}_{Theme}$ $\underbrace{\text{udělali revoluci.}}_{Rheme}$

(25) What about the Czechs and revolution?

$\underbrace{\text{Češi revoluci}}_{Theme}$ $\underbrace{\text{udělali.}}_{Rheme}$

$\underbrace{\text{Revoluci Češi}}_{Theme}$ $\underbrace{\text{udělali.}}_{Rheme}$

IS Realization Means: Word Order

- WO freedom is a matter of degree.

- Even in languages with "fixed" WO, there may be some freedom, e.g.:

(26) German: "free" WO in middle field (*G. Mittelfeld*)

Jan hat Anna gestern	gesehen.	Jan hat gestern	Anna gesehen.
Jan has Anna yesterday	seen.	Jan has yesterday	Anna seen.

(27) English: some freedom in order of modifiers

a. John flew from LONDON to Paris.

b. John flew to Paris from LONDON.

IS Realization Means: Syntax

Syntactic constructions that allow one to change order:

- fronting (so-called topicalization)
- left dislocation
- right dislocation
- cleft
- pseudo-cleft
- passivization
- dative shift
- there-insertion

Differences across languages!
Differences in contextual appropriateness. (Prince, 1978)

IS Realization Means: Syntax

(28) Comics, John hates.

(29) Comics, John hates them.

(30) John hates them, comics.

(31) It is John who hates comics.
It is comics John hates.

(32) Who hates comics is John. / John is (the one) who hates comics.
What John hates are comics. / Comics are what John hates.

(33) Comics are hated by John.

(34) John gave Mary a book.
John gave a book to Mary.

(35) There is a troll in the garden.

IS Realization: Example

(36) I know John writes novels. But what does Bill write?

a. BILL writes **POETRY**.	# **BILL** writes POETRY.
b. **POETRY** is written by BILL.	# POETRY is written by **BILL**.
c. It is **POETRY** BILL writes.	# It is POETRY **BILL** writes.
d. What BILL writes is **POETRY**.	# What **BILL** writes is POETRY.
e. **POETRY**, BILL writes.	# POETRY, **BILL** writes.
f. BILL, he writes **POETRY**.	# POETRY, **BILL** writes it.
g. He writes **POETRY**, BILL.	# **BILL** writes it, POETRY.

IS Realization: Example

(37) I know John writes novels. But who writes poetry?

a. **BILL** writes POETRY.	# BILL writes **POETRY**.
b. POETRY is written by **BILL**.	# **POETRY** is written by BILL.
c. It is **BILL** who writes POETRY.	# It is BILL who writes **POETRY**.
d. Who writes POETRY is **BILL**.	# Who writes **POETRY** is BILL.
e. It is POETRY what **BILL** writes.	
f. What **BILL** writes is POETRY.	
g. POETRY, **BILL** writes.	# **POETRY**, BILL writes.
h. POETRY, **BILL** writes it.	# BILL, he writes **POETRY**.
i. **BILL** writes it, POETRY.	# He writes **POETRY**, BILL.

Intermezzo

"Focussing jokes"

(38) Why do you rob banks?
Because that's where the money is!

(39) Why do firemen wear red suspenders?
To keep their pants up.

(40) Why do we buy clothes?
Because we can't get them for free.

(41) Why do we dress girls in pink and boys in blue?
Because they can't dress themselves.

"Focussing jokes": Explanation

Surprise effect due to discrepancy between the answers and the IS of question: what is focused, whether focus narrow/broad

(38′) Why do you rob BANKS?

(39′) Why do firemen wear RED suspenders?
Why do firemen wear red SUSPENDERS?

(40′) Why do we buy CLOTHES?
Why do we BUY clothes?

(41′) Why do we dress GIRLS IN PINK and BOYS IN BLUE?
Why do WE dress girls in pink and boys in blue?

IS Semantics

IS-Sensitive Semantic Interpretation

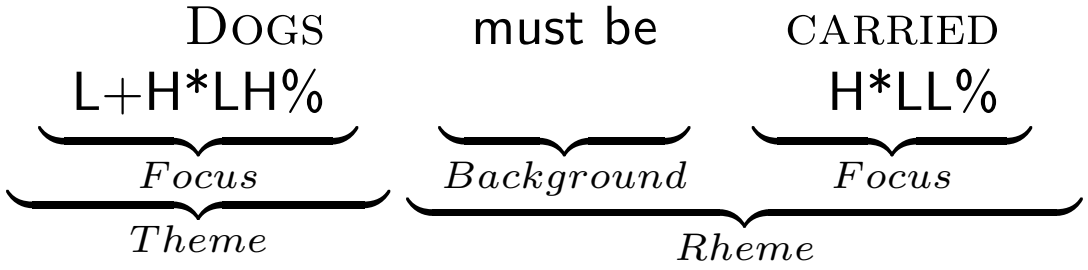
- von Stechow, Krifka: semantics of focus using structured meanings (von Stechow, 1990; Jackendoff, 1990; Krifka, 1992; Krifka, 1993)
- Hamblin: semantics of questions in terms of answer-alternative set (Hamblin, 1973)
- Rooth: semantics of focus in terms of focus-alternative set (Rooth, 1992)
- Büring: semantics of focus-marked topic in terms of question-alternative set (Büring, 1997; Büring, 1999)
- Steedman: semantics of two-dimensional IS-partitioning in terms of Rheme-alternative set and Theme-alternative set (Steedman, 2000)

IS-Sensitive Semantic Interpretation

- Semantics of IS in terms of selecting one member from a *presupposed set of alternatives* (Steedman, 2000)
  - 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.
- The systematic recognition of the alternative sets, and their maintenance as a discourse progresses are open research issues.

IS Semantics: Examples

(42) I know how to transport babies in the metro. But what about DOGS?



θ(42): $\lambda Q.\ Q\ (\star dog')$

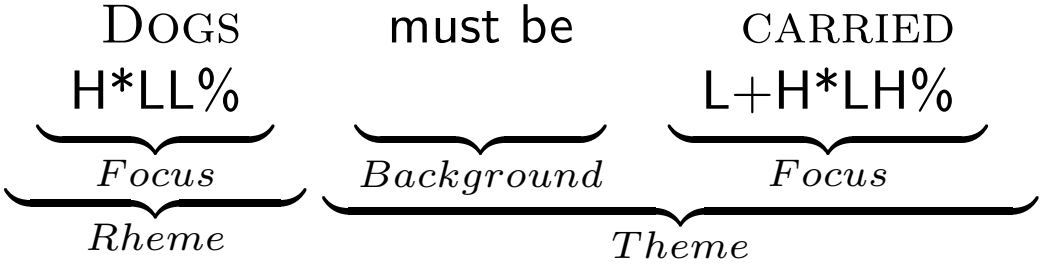
ρ(42): $\lambda x. \star carry'(hearer', x)$

ρ-AS(42): $\{carry'(hearer', dog'),$
$walk\_on\_lead(hearer', dog'), load\_on\_buggy(hearer', dog')\}$

θ-AS(42): $\{\exists Q.Q(dog'), \exists P.P(baby')\}$

IS Semantics: Examples

(43) I know what must be worn in the metro. But what must be CARRIED?



θ(43): $\lambda x. \star carry'(hearer', x)$

ρ(43): $\lambda Q.\ Q\ (\star dog')$

ρ-AS(43): $\{carry'(hearer', dog'), carry'(hearer', baby')\}$

θ-AS(43): $\{\exists x.carry'(hearer', x), \exists x.wear(hearer', x)\}$

Meaning Differences

IS: Meaning Differences

(44) Dogs must be carried. (Halliday, 1970)

a. What about dogs? $\underbrace{\text{Dogs}}_{Theme}$ $\underbrace{\text{must be CARRIED.}}_{Rheme}$

i.e., If there is a dog, it must be carried.

Presupposed alternative set: $\{\exists P.(P(e) \wedge patient(e, dog))\}$

b. $\underbrace{\text{DOGS must be carried.}}_{Rheme}$

Presupposed alternative set: $\{\exists P.P(e)\}$

IS: Meaning Differences

(45) Smoke in the hallway. (Hajičová, 1993)

a. Where should one smoke? $\underbrace{\text{Smoke}}_{Theme}$ $\underbrace{\text{in the HALLLWAY.}}_{Rheme}$

i.e., If you (want to) smoke, do it in the hallway.
Presupposed alternative set: $\{\exists x.(smoke(e) \wedge location(e, x))\}$

b. What should one do in the hallway? $\underbrace{\text{SMOKE}}_{Rheme}$ $\underbrace{\text{in the hallway.}}_{Theme}$

i.e., If you are in the hallway, smoke.
Presupposed alternative set: $\{\exists P.(P(e) \wedge location(e, hallway))\}$

IS: Meaning Differences

(46) Staff behind counter. (Hajičová, 1993)

a. Where should staff be? $\underbrace{\text{Staff}}_{Theme}$ $\underbrace{\text{BEHIND COUNTER.}}_{Rheme}$

i.e. Where staff should be is (only) behind the counter.
Presupposed alternative set: $\{\exists x.location(staff, x)\}$

b. Who should be behind the counter? $\underbrace{\text{STAFF}}_{Rheme}$ $\underbrace{\text{behind counter.}}_{Theme}$

i.e., Who should be behind the counter is (only) staff.
Presupposed alternative set: $\{\exists y.location(y, behind\_counter)\}$

IS: Meaning Differences

(47) a. What language is (mostly) spoken on the Shetlands?

$\underbrace{\text{On the SHETLANDS one speaks}}_{Theme}$ $\underbrace{\text{ENGLISH.}}_{Rheme}$

Presupposed alternative sets:

ρ-AS: $\{\exists x.(speak(e) \wedge location(e, shetlands) \wedge language(e, x))\}$

θ-AS: $\{\exists z.\exists x.(speak(e) \wedge location(e, z) \wedge language(e, x))\}$

b. Where is English (mostly) spoken?

$\underbrace{\text{One speaks ENGLISH}}_{Theme}$ $\underbrace{\text{on the SHETLANDS.}}_{Rheme}$

Presupposed alternative sets:

ρ-AS: $\{\exists y.(speak(e) \wedge language(e, english) \wedge location(e, y))\}$

θ-AS: $\{\exists z.\exists y.(speak(e) \wedge language(e, z) \wedge location(e, y))\}$

IS: Meaning Differences

(48) Officers always escorted ballerinas. (Partee et al., 1998)

a. Whom did officers always escort?

$\underbrace{\text{Officers escorted}}_{Theme}$ always $\underbrace{\text{BALLERINAS.}}_{Rheme}$

$\{\exists x.(escort(e) \wedge actor(e, officer) \wedge patient(e, x))\}$

b. What did officers always do?

$\underbrace{\text{Officers}}_{Theme}$ always $\underbrace{\text{escorted BALLERINAS.}}_{Rheme}$

$\{\exists P.(P(e) \wedge actor(e, officer))\}$

c. Who always escorted ballerinas?

$\underbrace{\text{OFFICERS}}_{Rheme}$ always $\underbrace{\text{escorted ballerinas.}}_{Theme}$

$\{\exists y.(escort(e) \wedge actor(e, y) \wedge patient(e, ballerina))\}$

IS: Meaning Differences

Czech newspaper 1990: (Hajičová)

(49) Dobrá zpráva je, že Češi udělali revoluci.
Špatná zpráva je, že revoluci udělali Češi.

The good news is that the Czechs made a revolution;
the bad news is that a revolution was made by the Czechs.
(. . . the bad news is that the CZECHS made a revolution.)

$\{\exists x.(make(e) \wedge actor(e, czechs) \wedge patient(e, x))\}$
$\{\exists y.(make(e) \wedge actor(e, y) \wedge patient(e, revolution))\}$

(50) Problém není v tom, že Janouch koupil gamma nůž,
ale že gamma nůž koupil Janouch.

The problem is not that Janouch bought a gamma-knife,
but that the gamma-knife was bought by Janouch.
(. . . but that JANOUCH bought the gamma knife.)

$\{\exists x.(buy(e) \wedge actor(e, janouch) \wedge patient(e, x))\}$
$\{\exists y.(buy(e) \wedge actor(e, y) \wedge patient(e, gknife))\}$

IS and Discourse Dynamics

IS and Discourse Dynamics

- IS and the File-Change Metaphor (Vallduví, 1992)
  Theme : file-card address
  Rheme : information to add/modify on a card
- IS-Sensitive Context Update (Krifka, 1993; Kruijff-Korbayová, 1998; Steedman, 2000)

Theme update phase : $c_1[\theta(\psi)]c_2$
verify Theme presuppositions $\mathsf{AS}_\theta(\psi)$ and $\mathsf{AS}_\rho(\psi)$; restrict $\mathsf{AS}_\theta(\psi)$
Rheme update phase : $c_2[\rho(\psi)]c_3$
restrict $\mathsf{AS}_\rho(\psi)$

IS-Sensitive Context Updating: "otherwise"

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

If you do something other with a dog than carry it you might get hurt.

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

If you carry something other then a dog than you might get hurt.

- In "otherwise β", "otherwise" is an anaphor that maps its antecedent α to a context complementary w.r.t. α, in which β is then interpreted (Webber et al., 1999)

- The IS of α makes additional conditions available for "otherwise β" to pick anaphorically (Kruijff-Korbayová and Webber, 2001)

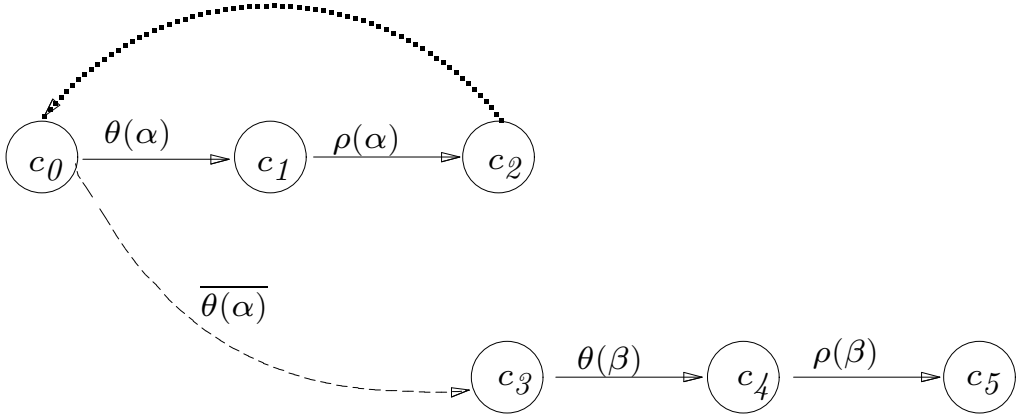
Full Theme-complement condition

"otherwise" can appeal to a condition derived from the Theme of the antecedent

(53) α: At a red light, STOP. β: Otherwise you can go straight on.

H*LL%

$\underbrace{\text{At a red light,}}_{Theme}$ $\underbrace{\text{STOP.}}_{Rheme}$



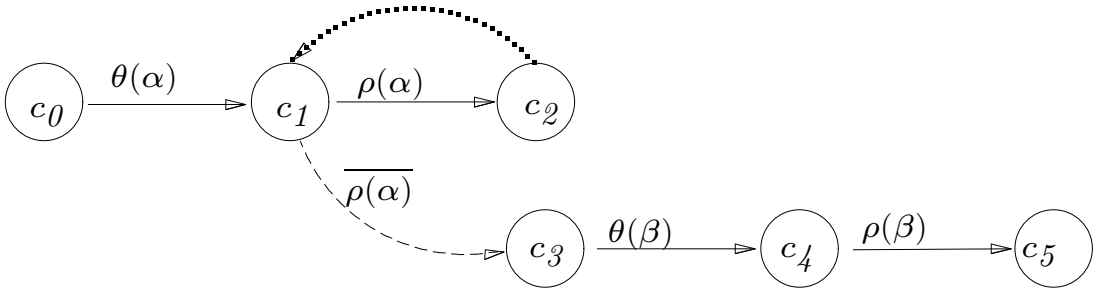
If the light is not red, go straight on.

Full Rheme-complement condition

"otherwise" can appeal to a condition derived from the Rheme of the antecedent

(54) α: At a red light, $\underbrace{\underline{\text{STOP}}}_{\text{H*LL\%}}$. β: Otherwise you will get a ticket.

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



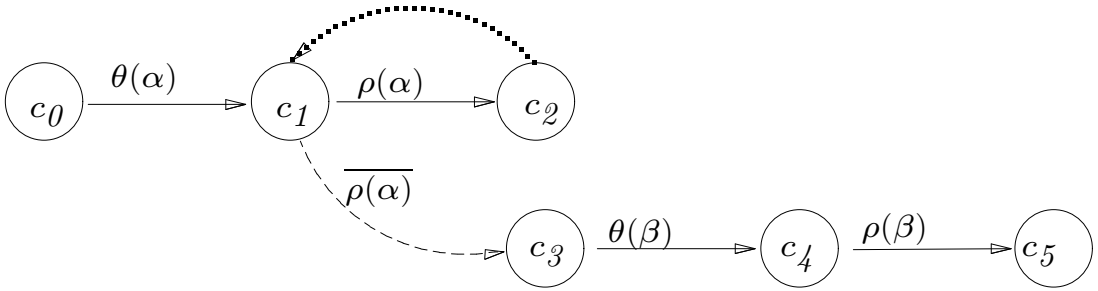
If the light is red and <u>you do not stop</u>, you will get a ticket.

Updating c_1 with "otherwise β"

(55) α: Stop at a red LIGHT. β: Otherwise you might get rear-ended.

H* LL%

$Theme$ $Rheme$



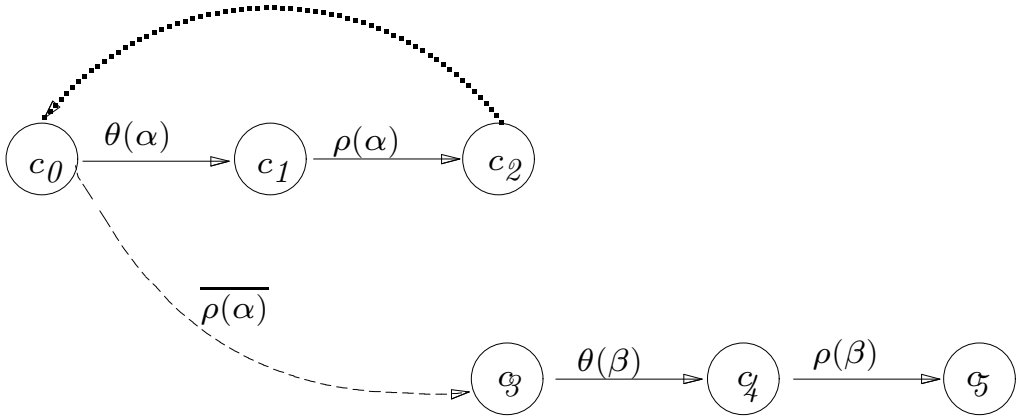
If you stop and the light is not red, you might get rear-ended.

Updating c_0 with "otherwise β"

(56) α: Stop at a red LIGHT. β: Otherwise you can go straight on.

H* LL%

$\underbrace{\text{Stop}}_{Theme}$ $\underbrace{\text{at a red LIGHT.}}_{Rheme}$



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

Summary

- Webber et al.'s analysis holds, with IS-sensitive analysis adding new possibilities:
- Antecedent's IS makes additional conditions available for "otherwise" to pick anaphorically:
  - *full Theme-complement* or *partial Theme-complement* condition
  - *full Rheme-complement* or *partial Rheme-complement* condition
- β can be asserted with respect to (at least) the following contexts:
  - context c_0 (prior to antecedent)
  - context c_1 (restricted by antecedent's Theme$_{is}$)
  - context c_ϕ (restricted by antecedent's "if"-clause)

Some Open Issues

- What (alternative) conditions could a speaker have in mind and what features of language give evidence for them?
- What do we learn about the relationships between IS and discourse structure if we analyse "otherwise" itself as a contrastive Theme$_{is}$, marking a contrast w.r.t. a preceding theme or rheme?
- Do postposed subordinate clauses in complex sentences have their own IS?
- How can claims about IS be tested?

IS in Computational Applications

Computational Modeling of IS in Applications

- analysis in question/text understanding, MT or TTS
  - word order: (Hoffman, 1995); (Styś and Zemke, 1995)
  - intonation: (Prevost, 1995)
  - anaphora resolution (Hajičová et al., 1990; Hajičová et al., 1992)

- production in NLG, MT, TTS or dialog systems
  - word order: (Hoffman, 1995; Hoffman, 1996); (Kruijff-Korbayová et al., 2002)
  - intonation: (Prevost, 1995); (Kruijff-Korbayová et al., 2003); (Moore et al., 2004)
  - referring expression generation in text (Hajičová et al., 1990); in dialogue (Krahmer and Theune, 2002)
  - embodied agents' gesture (Pelachaud et al., 1998; Cassell et al., 2000), gaze and turn-taking (Cassell-etal:1999)

Course Plan

Course Plan

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

Day 2 The Praguian Topic-Focus Articulation. Givenness. Familiarity Status. Salience. IS-sensitive Salience Modeling in Analysis and Generation.

Day 3 Steedman: Two Dimensions of IS. Alternative-set Semantics of IS. IS and Intonation. IS and Turn-Taking, Gesture and Gaze in Multimodal Dialog.

Day 4 Vallduví: Information Packaging & File-Change. Halliday: Thematic- vs. Information-Structure. Daneš: Thematic Sequences. IS in WO Generation.

Day 5 Wrapping Up and Looking Out. Aligning IS-Approaches. Empirical Studies. Testing Theories. Corpus Annotation Issues.

References

Daniel Büring. 1997. *The Meaning of Topic and Focus: The 59th Street Bridge Accent*. Routledge, London.

Daniel Büring. 1999. Topic. In Peter Bosch and Rob van der Sandt, editors, *Focus: Linguistic, Cognitive and Computational Principles*, Natural Language Processing, pages 142–165. Cambridge University Press, Cambridge.

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

Caroline Fery. 1993. *German Intonational Patterns*. Tuebingen:Niemeyer.

Jan Firbas. 1971. On the concept of communicative dynamism in the theory of functional sentence perspective. *Brno Studies of English*, (7).

Jan Firbas. 1992. *Functional Sentence Perspective in Written and Spoken Communication*. Studies in English Language. Cambridge University Press, Cambridge.

Eva Hajičová, Petr Kuboň, and Vladislav Kuboň. 1990. Hierarchy of salience and discourse analysis and production. pages 144–148.

Eva Hajičová, Vladislav Kuboň, and Petr Kuboň. 1992. Stock of shared knowledge - a tool for solving pronominal anaphora. pages 127–133.

Eva Hajičová. 1993. *Issues of sentence structure and discourse patterns*, volume 2 of *Theoretical and computational linguistics*. Charles University, Prague, Czech Republic.

Michael A.K. Halliday. 1970. *A Course in Spoken English: Intonation*. Oxford Uniersity Press, Oxford.

C. Hamblin. 1973. Questions in Montague english. *Foundations of Language*, pages 41–53.

Beryl Hoffman. 1995. Integrating free word order, syntax, and information structure. In *Proceedings of the 7th Conference of the European Chapter of the Association for Computational Linguistics, Dublin*, pages 245–252, San Francisco, CA. Morgan Kaufmann.

Beryl Hoffman. 1996. Translating into free word order languages. In *Proceedings of the International Conference on Computational Linguistics (COLING-96), Copenhagen*, pages 556–561, Copenhagen, Denmark. Center for Sprogteknologi.

Ray Jackendoff. 1990. *Semantic Structures*. MIT Press, Cambridge.

Emiel Krahmer and Mariet Theune. 2002. Efficient context-sensitive generation of referring expressions. In van Deemter and Kibble (van Deemter and Kibble, 2002), pages 223–264.

Manfred Krifka. 1992. A compositional semantics for multiple focus constructions. In J. Jacobs, editor, *Informationsstruktur und Grammatik*, Linguistische Berichte, Sonderheft 4, pages 17–53.

Manfred Krifka. 1993. Focus and presupposition in dynamic semantics. *Theoretical Linguistics*, 19:269–300.

Ivana Kruijff-Korbayová and Mark Steedman. 2003. Discourse and information structure. *Journal of Logic, Language and Information: Special Issue on Discourse and Information Structure*, 12(3):249–259.

Ivana Kruijff-Korbayová and Bonnie L. Webber. 2001. Information structure and the semantics of "otherwise". In Ivana Kruijff-Korbayová and Mark Steedman, editors, *Information Structure, Discourse Structure and Discourse Semantics, ESSLLI2001 Workshop Proceedings*, pages 61–78, Helsinki, Finland, August 20-24. European Summer School in Logic, Language and Information (ESSLLI), The University of Helsinki. `http://www.coli.uni-sb.de/~korbay/esslli01-wsh/Proceedings/24-Kruijff-Webber.ps.gz`.

Ivana Kruijff-Korbayová, John Bateman, and Geert-Jan M. Kruijff. 2002. Generation of contextually appropriate word order. In van Deemter and Kibble (van Deemter and Kibble, 2002), pages 193–222. http://www.coli.uni-sb.de/publikationen/softcopies/Kruijff-Korbayova:1999:GCA.pdf http://www.coli.uni-sb.de/publikationen/softcopies/Kruijff-Korbayova:1999:GCA.ps.

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.

Ivana Kruijff-Korbayová. 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.

Johanna Moore, Mary Ellen Foster, Oliver Lemon, and Michael White. 2004. Generating tailored, comparative descriptions in spoken dialogue. In *Proceedings of the Florida Artificial Intelligence Research Symposium (FLAIRS)*.

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

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. 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.

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

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

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

Malgorzata E. Styś and Stefan S. Zemke. 1995. Incorporating discourse aspects in english-polish mt: Towards robust implementation. In *Proceedings of the RANLP Conference*, Velingrad, Bulgaria.

Susanne Uhmann. 1991. *Fokusphonologie*. Tuebingen:Niemeyer.

Enric Vallduví. 1992. *The Informational Component*. Garland, New York.

Kees van Deemter and Rodger Kibble, editors. 2002. volume 143 of *Lecture Notes*. CSLI.

Arnim von Stechow. 1990. Categorial grammar and linguistic theory. *Studies in Language*, 14:433–478.

Bonnie Webber, Alistair Knott, Matthew Stone, and Aravind 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*, pages 41–48, College Park MD. url.