



# Einführung in Pragmatik und Diskurs

## Informationsstruktur

Ivana Kruijff-Korbayová  
korbay@coli.uni-sb.de  
<http://www.coli.uni-saarland.de/courses/pd/>

Summer Semester 2006



## Information Structure

- Introduction
- Question test for IS
- IS realization
- IS meaning
- IS in practical applications

Reading: [Ste00]; [KKERK03]



## Illustration (1)

- (1) Sign in London underground:  
*Dogs must be carried.*
- Hunde müssen getragen werden.  
Einen Hund muss man tragen.

Interpretation:

*If you have a dog, you must carry it.*



## Illustration (2)

- (2) Sign in a synagog:  
*Hats must be worn.*
- (3) Man muss einen Hut tragen.  
(4) Es müssen Hüte getragen werden.  
(5) Hüte müssen getragen werden.

Interpretation:

- (6) *You must wear a hat.*



4

## Illustration

- (7) Dogs must be CARRIED.  
Hunde müssen GETRAGEN werden.
- (8) # DOGS must be carried.  
# HUNDE müssen getragen werden.
- (9) HÜTE müssen getragen werden.
- (10) # Hüte müssen GETRAGEN werden.

(Capitals denote intonation center = main stress in the sentence.)



5

## Illustration

Different languages use various means to realize the different meanings:  
e.g., Czech: word order

- (11) *Psi se musí NĚST.*  
Dog refl must carry  
Hunde müssen getragen werden.
- (12) # *Musí se nést PSI.*  
Must refl carry dog  
Es müssen Hunde getragen werden.



6

## Illustration (4)

Example from a dialogue system:

- (13) U: What is the status of the stove?  
Im welchen Zustand befindet sich die Heizung?  
S: The stove is switched ON.  
Die Heizung ist ANGESCHALTET.
- (14) U: Which device is switched on?  
Was ist angeschaltet?  
S: The STOVE is switched on.  
Die HEIZUNG ist angeschaltet.



7

## Observations

- Intonation, word order, syntactic structure and other aspects of linguistic form are tightly related
- Variation in these aspects of linguistic form is not arbitrary!
- These aspects of linguistic form reflect the relation of the utterance to the context
- Whether an utterance is appropriate or inappropriate in a given context depends not only on **what is said** but also **how**
- Recall: Grice's Maxim of Manner





## Theme-Rheme Partitioning: Examples

- (20) U: What is the status of the stove?  
Im welchen Zustand befindet sich die Heizung?

S: The stove is switched ON.  
Die Heizung ist ANGESCHALTET.

*Theme*
*Rheme*

- (21) U: Which device is switched on?  
Was ist angeschaltet?

S: The STOVE is switched on.  
Die HEIZUNG ist angeschaltet.

*Rheme*
*Theme*



## Theme-Rheme Partitioning: Examples

- (22) Das Mädchen kauft eine Bluse.
- Theme* *Rheme*
- (23) Das Mädchen kauft eine Bluse.
- Theme* *Rheme*
- (24) Das Mädchen kauft eine Bluse.
- Theme* *Rheme* *Theme*
- (25) Das Mädchen kauft eine Bluse.
- Rheme* *Theme*



## Question Test for IS

Question-answer pairs are commonly used to indicate or test the context in which a particular IS is appropriate: the question determines the Theme; what is asked for is the Rheme; the answer "fills" the Rheme.

- (26) Was macht das Mädchen?  
Das Mädchen kauft eine BLUSE.
- Theme* *Rheme*
- (27) Wer kauft eine Bluse?  
Das MÄDCHEN kauft eine Bluse.
- Rheme* *Theme*

Swapping the questions in the two examples results in incoherent Q-A pairs: the answers become infelicitous, because the IS partitioning then does not match the context set by the question.



## Question Test for IS

The linguistic form of an utterance may be compatible with several different IS partitionings (as reflected by different questions)

- (28) Das Mädchen kauft eine BLUSE.
- a. Was kauft das Mädchen?  
Das Mädchen kauft eine BLUSE.
- Theme* *Rheme*
- b. Was macht das Mädchen?  
Das Mädchen kauft eine BLUSE.
- Theme* *Rheme*
- c. Was passiert gerade? / Was ist neu?  
Das Mädchen kauft eine BLUSE.
- Rheme*

- (29) Das Mädchen kauft eine BLUSE.  
 # Wer kauft eine Bluse?  
 # Was macht das Mädchen mit eine Bluse?

- (30) Wer kauft eine Bluse?  
 Das MÄDCHEN kauft eine Bluse.  
*Rheme Theme*

- (31) Was macht das Mädchen mit eine Bluse?  
 Das Mädchen KAUFT eine Bluse.  
*Theme Rheme Theme*

- (34) is only a matching answer to (35).  
 (34) John flew to PARIS from London.  
 (35) Where did John fly from London?  
 (36) is only a matching answer to (37).  
 (36) John flew from LONDON to Paris.  
 (37) From which place did John fly to Paris?  
 Neither of the above is a good answer to (38).  
 (38) Who flew from London to Paris?  
 (39) JOHN flew from London to Paris.

## Question Test for IS

- (32) John flew from London to PARIS.  
 (33) a. What happened?  
 b. What did John do?  
 c. Where did John fly?  
 d. To which place did John fly from London?  
 (32) is good as an answer to (33a-d). It has several possible IS partitionings.

## Second Dimension of IS Partitioning: Background-Focus

**Background-Focus partitioning** reflects an abstract notion of contrast between alternatives available in the discourse context, against which the actual utterance is cast; B/F partitioning within Theme and/or Rheme reflects alternative Theme(s) and/or Rheme(s) in the context.

### IS Partitioning: Background-Focus in Rheme

- (40) Es gibt eine Bluse aus Baumwolle und eine aus Seide.  
Was kauft das Mädchen?

Das Mädchen kauft eine SEIDENE Bluse.

*Theme* (under "Das Mädchen kauft")

*Focus* (under "SEIDENE")

*Rheme* (under "eine SEIDENE Bluse")

- (41) Es gibt eine seidene Bluse und eine seidene Jacke.  
Was kauft das Mädchen?

Das Mädchen kauft eine seidene BLUSE .

*Theme* (under "Das Mädchen kauft")

*Focus* (under "BLUSE")

*Rheme* (under "eine seidene BLUSE")

### IS Partitioning: Background-Focus in Theme

- (42) Es gibt ein deutsches und ein amerikanisches Mädchen.  
Was kauft das deutsche Mädchen?

Das DEUTSCHE Mädchen kauft eine BLUSE .

*Focus* (under "DEUTSCHE")

*Rheme* (under "eine BLUSE")

*Theme* (under "Das DEUTSCHE Mädchen kauft")

- (43) Es gibt eine deutsche Frau und ein deutsches Mädchen.  
Was kauft das deutsche Mädchen?

Das deutsche MÄDCHEN kauft eine BLUSE .

*Focus* (under "MÄDCHEN")

*Rheme* (under "eine BLUSE")

*Theme* (under "Das deutsche MÄDCHEN kauft")

### IS Partitioning: Background-Focus in Theme and Rheme

- (44) Ich weiss, dass die deutsche Frau eine Bluse aus Baumwolle kauft.  
Aber was kauft das deutsche Mädchen?

Das deutsche MÄDCHEN kauft eine SEIDENE Bluse .

*Focus* (under "MÄDCHEN")

*Focus* (under "SEIDENE")

*Theme* (under "Das deutsche MÄDCHEN kauft")

*Rheme* (under "eine SEIDENE Bluse")

### IS Realization

- Various means that can be used (i.e., IS influences various aspects of linguistic form)
  - **intonation** (prosody); e.g., the predominant means in English
  - (word) **ordering**; e.g. the predominant means in Czech and other Slavic languages, also to some extent in German (particularly in the “Mittelfeld”)
  - **morphological/grammatical marking**; e.g., particles ‘wa’ and ‘ga’ in Japanese
  - **syntactic constructions**, e.g. it-cleft, wh-cleft, passivization, etc.
  - **ellipsis**
- The means can be used also in combination;
- Different languages employ and combine the means differently, depending on their typological characteristics;

## IS Realization Means: Intonation

- placement and type of accent and type of boundary tone
  - accent placed on Focus part
  - Theme-Focus and Rheme-Focus use different type of accent and boundary tone (tune): L+H\**LH*% is a prototypical Theme-tune, H\**LL*% is a prototypical Rheme-Tune
- placement of boundaries: boundary between Theme and Rheme

(45) Ich weiss, dass die deutsche Frau eine seidene Jacke anzieht.  
Aber was macht das deutsche Mädchen?

(46) Das deutsche MÄDCHEN kauft eine seidene BLUSE.

$\underbrace{\text{L+H**LH*\%}}_{\text{Focus}} \quad \underbrace{\text{H**LL*\%}}_{\text{Focus}}$

$\underbrace{\hspace{10em}}_{\text{Theme}} \quad \underbrace{\hspace{10em}}_{\text{Rheme}}$

## IS Realization Means: Word Order

“Normal” (default) order: Theme before Rheme

(47) Was kauft das Mädchen? Das Mädchen kauft eine BLUSE.

$\underbrace{\hspace{10em}}_{\text{Theme}} \quad \underbrace{\hspace{10em}}_{\text{Rheme}}$

(48) MP3: Was soll mit den Liedern gemacht werden?

U: Mit den Liedern soll eine Playlist erstellt werden.

(49) U: Bitte suche Titel von Madonna.

MP3: Einen Moment bitte. ... Von Madonna haben wir 1711 Treffer.

“Subjective ordering”: Rheme before Theme (is marked)

(50) Was kauft das Mädchen? Eine BLUSE kauft das Mädchen.

$\underbrace{\hspace{10em}}_{\text{Rheme}} \quad \underbrace{\hspace{10em}}_{\text{Theme}}$

## IS Realization Means: Syntax

Syntactic constructions can be used to mark IS:

Cleft: *It is Rheme (that/who) Theme*

1. What does John hate? It is COMICS John hates.
2. Who hates comics? It is JOHN who hates comics.

Pseudo-cleft: *Who/What Theme is/are Rheme*

1. What does John hate? What John hates are COMICS.
2. Who hates comics? Who hates comics is JOHN.

Passivization: allows opposite ordering than corresponding active sentence

- (51) Who hates comics?
- a. Comics are hated by JOHN.
  - b. JOHN hates comics.

There-insertion: gets Rheme-subject away from the beginning of the sentence

- (52) a. What is in the garden? There is a TROLL in the garden.  
b. Where is a troll? A troll is in the GARDEN.

(53) Es spielt die Tschechische Philharmonie. Es dirigiert Hilary Griffiths.

## IS Realization Means: Ellipsis

Theme can be left implicit when it is “known” (available, retrievable):

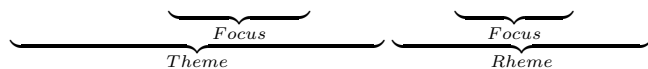
- (54) A: What does John hate?  
B: COMICS.
- (55) A: Who hates comics?  
B: JOHN.
- (56) MP3: Was soll mit den Liedern gemacht werden?  
U: Mit den Liedern soll eine Playlist erstellt werden.
- (57) U: Bitte suche Titel von Madonna.  
MP3: Einen Moment bitte. ... Von Madonna haben wir 1711 Treffer.
- (58) U: Wieviele Titel von Madonna gibt es?  
MP3: Einen Moment. ... 1711.

## The Meaning of IS

- Steedman defines the meaning of the IS partitioning in terms of selecting one member from a *presupposed set of alternatives*
  - The Theme-Rheme partitioning 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
  - The Focus-Background partitioning in 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 is available in the context.
- The systematic recognition of the alternative sets, and their maintenance as a discourse progresses are open research issues.

## IS Meaning

- (59) Das deutsche MÄDCHEN kauft eine SEIDENE Bluse .



Rheme-alternatives: *das deutsche Mädchen kauft x*  
 {das deutsche Mädchen kauft eine seidene Bluse, das deutsche Mädchen kauft eine Bluse aus Baumwolle, das deutsche Mädchen kauft eine Bluse aus Nylon, etc. }

Theme-alternatives: *y kauft etwas*  
 {das deutsche Mädchen kauft etwas, die deutsche Frau kauft etwas, der deutsche Man kauft etwas, der deutsche Junge kauft etwas, etc. }

## IS: Meaning Differences

- (60) a. Smoke OUTSIDE!  
b. SMOKE outside!
- (61) a. STAFF behind counter.  
b. Staff behind COUNTER.
- (62) a. *On the Shetlands one speaks* ENGLISH.  
b. *One speaks English on the* SHETLANDS.
- (63) a. *Several people in this room know three languages.*  
b. *Three languages are known by several people in this room.*
- (64) a. Officers always escorted BALLERINAS.  
b. OFFICERS always escorted ballerinas.

## “Focussing jokes”

- (65) Why do you rob banks?  
Because that's where the money is!
- (66) Why do firemen wear red suspenders?  
To keep their pants up.
- (67) Why do we buy clothes?  
Because we can't get them for free.
- (68) Why do we dress girls in pink and boys in blue?  
Because they can't dress themselves.

## IS in Multimodal Interaction

- Appropriate and synchronized speech, intonation, facial expressions and hand gestures [PCB<sup>+</sup>98]
- Integrating turn-taking and IS provides better explanation for gaze behavior [CTP99]
- Generation of either speech, gesture or combination of both as a function of IS status and surprise value of a discourse entity [CSY00]
- Various researches have observed that distance, posture shifts and other body movements seem to accompany changes in the topic or social relationship

## Computational Modeling of IS in Applications

- word order:
  - anaphora resolution/generation [HKK90, HKK92]
  - database-querying dialogue [Hof95]
  - MT: En. → Turkish [Hof96]; En. → Polish [SZ95]
  - text generation: Cz., En. [KKBK02]
- intonation in dialogue: En. [Pre95], Ger., En. [KKERK03]; En. [MFLW04]
- referring expression generation in text [HKK90] and in dialogue [KT02]
- non-linguistic aspects in modeling embodied conversational agents: gestures [PCB<sup>+</sup>98, CSY00], gaze and turn-taking [CTP99]

## IS and Gestures

Gilbert and George [PCB<sup>+</sup>98]; Greta ECA (<http://www.iut.univ-paris8.fr/greta/>)

- Some facial expressions are automatically generated according to intonation (e.g., raising eyebrows with accented material), based on corpus studies
- 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

- IS-partitioning: Theme-Rheme, Focus-Background
- IS realization: intonation, word order, syntactic constructions, ...
- IS is an essential aspect of meaning at the interface between utterance and discourse
- IS is important for accurate NL processing
- formal accounts are emerging, some embodied into practical systems
- many questions concerning IS partitioning and its realization in different languages still open
- Further research topics: [KKS03]
  - further systematization of terminologies
  - formalization and computational modeling&testing
  - empirical and corpus-based studies
  - cross-linguistic investigations and multilingual applications

## References

- [CSY00] Justine Cassell, Matthew Stone, and Hao Yan. Coordination and context-dependence in the generation of embodied conversation. In *Proceedings of the INLG Conference*, 2000.
- [CTP99] Justine Cassell, Obed. E. Torres, and Scott Prevost. Turn taking vs. discourse structure: How best to model multimodal conversation. In Y. Wilks, editor, *Machine Conversations*. Kluwer, 1999.
- [HKK90] Eva Hajičová, Petr Kuboň, and Vladislav Kuboň. Hierarchy of salience and discourse analysis and production. pages 144–148, 1990.
- [HKK92] Eva Hajičová, Vladislav Kuboň, and Petr Kuboň. Stock of shared knowledge - a tool for solving pronominal anaphora. pages 127–133, 1992.
- [Hof95] Beryl Hoffman. 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, 1995. Morgan Kaufmann.
- [Hof96] Beryl Hoffman. Translating into free word order languages. In *Proceedings of the International Conference on Computational Linguistics (COLING-96), Copenhagen*, pages 556–561, Copenhagen, Denmark, 1996. Center for Sprogteknologi.
- [KKBK02] Ivana Kruijff-Korbayová, John Bateman, and Geert-Jan M. Kruijff. Generation of contextually appropriate word order. In van Deemter and Kibble [vDK02], pages 193–222.

- [KKERK03] Ivana Kruijff-Korbayová, Stina Ericsson, Kepa Joseba Rodríguez, and Elena Karagjosova. 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, 2003.
- [KKS03] Ivana Kruijff-Korbayová and Mark Steedman. Discourse and information structure. *Journal of Logic, Language and Information: Special Issue on Discourse and Information Structure*, 12(3):249–259, 2003.
- [KT02] Emiel Krahmmer and Mariet Theune. Efficient context-sensitive generation of referring expressions. In van Deemter and Kibble [vDK02], pages 223–264.
- [MFLW04] Johanna Moore, Mary Ellen Foster, Oliver Lemon, and Michael White. Generating tailored, comparative descriptions in spoken dialogue. In *Proceedings of the Florida Artificial Intelligence Research Symposium (FLAIRS)*, 2004.
- [PCB<sup>+</sup>98] Catherine Pelachaud, Justine Cassel, Norman Badler, Mark Steedman, Scott Prevost, and Mathew Stone. 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, 1998.
- [Pre95] Scott Prevost. *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, 1995.
- [Ste00] Mark Steedman. Information structure and the syntax-phonology interface. *Linguistic Inquiry*, 31(4):649–689, 2000.

- [SZ95] Malgorzata E. Styś and Stefan S. Zemke. Incorporating discourse aspects in english-polish mt: Towards robust implementation. In *Proceedings of the RANLP Conference*, Velingrad, Bulgaria, 1995.
- [vDK02] Kees van Deemter and Rodger Kibble, editors. volume 143 of *Lecture Notes*. CSLI, 2002.