



Einführung in Pragmatik und Diskurs

Information Structure

Ivana Kruijff-Korbayová

korbay@coli.uni-sb.de

<http://www.coli.uni-saarland.de/courses/pd/>

Summer Semester 2005

Information Structure

- Introduction
- Basic concepts of IS partitioning
- Question test for IS
- IS interpretation
- IS realization
- Where does IS come from?
- IS in practical applications
- More about Mary

Reading: [Steedman2000]; [Kruijff-Korbayová et al.2003]

Illustration

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

Interpretation:

- (4) *If you have a dog, you must carry it.*

Illustration

- (5) Sign in a synagog:
Hats must be worn.
- (6) Man muss einen Hut tragen.
- (7) Es müssen Hüte getragen werden.
- (8) Hüte müssen getragen werden.

Interpretation:

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

Illustration

- (10) Dogs must be CARRIED.
Hunde müssen GETRAGEN werden.
- (11) # Hüte müssen GETRAGEN werden.
- (12) # DOGS must be carried.
HUNDE müssen getragen werden.
- (13) HÜTE müssen getragen werden.

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

Illustration

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

(14) *Pes se musí NÉST.*

Dog refl must carry

Ein Hund muss getragen werden.

(15) *# Musí se nést PES.*

Must refl carry dog

Es muss ein Hund getragen werden.

⇒ intonation, word order and syntactic structure are tightly related aspects of linguistic form; a particular linguistic form can be appropriate or inappropriate in a given context

Information Structure

- Speakers organize their utterances in a way that reflects their model of the context and their communicative intentions (the intended context change)
Utterances both *reflect* and *affect* the context.
- IS concerns a division (partitioning) of an utterance meaning into:
 - 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)
 - the part which advances the discourse, i.e., adds or modifies some information
- IS is an inherent aspect of meaning —it is an important factor in establishing coherence with respect to the context in which a sentence is uttered.



Basic Concepts

Information Structure Partitioning

Two dimensions of utterance meaning partitioning:

- Theme-Rheme partitioning
- Background-Focus partitioning

IS Partitioning: Theme-Rheme

Theme-Rheme partitioning Partitioning of utterance meaning into what speaker means to address vs. what she wants to say about it. This reflects an *aboutness* relation, i.e., the Rheme is semantically predicated over the Theme

Terminology as used in (Steedman 2000)

Theme-Rheme Partitioning: Examples

- (16) Das Mädchen kauft eine Bluse.
Theme *Rheme*
- (17) Das Mädchen kauft eine Bluse.
Theme *Rheme*
- (18) Das Mädchen kauft eine Bluse.
Theme *Rheme* *Theme*
- (19) 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.

(20) Was macht das Mädchen?

Das Mädchen kauft eine Bluse.
Theme *Rheme*

(21) 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

(22) Was passiert? / Was ist neu?

Das Mädchen kauft eine Bluse.

Rheme

(23) Was macht das Mädchen?

Das Mädchen kauft eine Bluse.

Theme

Rheme

(24) Was kauft das Mädchen?

Das Mädchen kauft eine Bluse.

Theme

Rheme

(25) Was macht das Mädchen mit eine Bluse?

Das Mädchen kauft eine Bluse.

Theme

Rheme

Theme



Question Test for IS

An utterance can have a range of possible IS partitionings of an utterance:

(26) John flew from London to PARIS.

(27) a. What happened?

b. What did John do?

c. Where did John fly?

d. To which place did John fly from London?

(26) is good as an answer to (27a-d). It has several possible IS partitionings.

(28) is only a matching answer to (31).

(28) John flew to PARIS from London.

(29) From which place did John fly to Paris?



(30) is only a matching answer to (31).

(30) John flew from LONDON to Paris.

(31) From which place did John fly to Paris?

Neither of the above is a good answer to (32).

(32) Who flew from London to Paris?

(33) JOHN flew from London to Paris.

IS Partitioning: Background-Focus

(34) Was kauft das Mädchen?

Das Mädchen kauft eine seidene Bluse.

Theme *Rheme*

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

Das Mädchen kauft eine SEIDENE Bluse.

Theme *Rheme*

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

Das Mädchen kauft eine seidene BLUSE.

Theme *Rheme*

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

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

Das Mädchen kauft eine SEIDENE Bluse.

Theme *Focus*
Rheme

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

Das Mädchen kauft eine seidene BLUSE .

Theme *Focus*
Rheme

IS Partitioning: Background-Focus in Theme

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

Das DEUTSCHE Mädchen kauft eine Bluse .
Focus *Rheme*
Theme

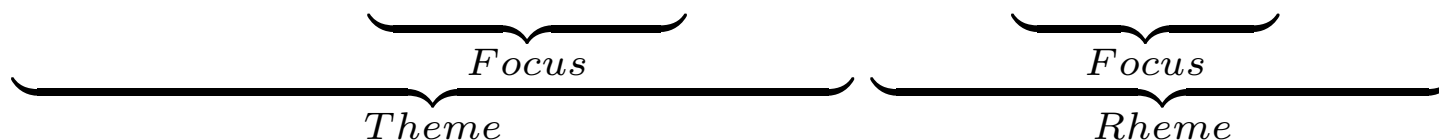
- (40) Es gibt eine deutsche Frau und ein deutsches Mädchen.
Was kauft das deutsche Mädchen?

Das deutsche MÄDCHEN kauft eine Bluse .
Focus *Rheme*
Theme

IS Partitioning: Background-Focus in Theme and Rheme

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

Das deutsche MÄDCHEN kauft eine SEIDENE Bluse .



IS Interpretation

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

IS Interpretation

(42) Das deutsche MÄDCHEN kauft eine SEIDENE Bluse .



Rheme-alternatives:

{ das deutsche Mädchen kauft eine seidene Bluse, das deutsche Mädchen kauft eine Bluse aus Baumwolle, das deutsche Mädchen kauft eine Bluse aus Nylon, etc. }

Thenme-alternatives:

{ das deutsche Mädchen kauft etwas, die deutsche Frau kauft etwas, der deutsche Man kauft etwas, der deutsche Junge kauft etwas, etc. }

IS interpretation Examples

Various IS-partitionings of (43)

- (43) The British amplifier has clean treble.
- (44)
 - i. Origin: British, American
 - ii. Component: amplifier, speaker
 - iii. Characteristic: treble, bass
 - iv. Property: clean, muddy

IS Interpretation Examples

Difference in Theme/Rheme partitioning:

- (45) Q. Which component has clean treble?
A. The British amplifier has clean treble.

- (46) Q. Which property does the British amplifier have?
A. The British amplifier has clean treble.

IS Interpretation Examples

Difference in Rheme-Focus:

- (47) Q. I know the American amplifier has muddy treble. But which component has clean treble?
A. The British amplifier has clean treble.
- (48) Q. I know the British speaker has muddy treble. But which component has clean treble?
A. The British amplifier has clean treble.

IS Interpretation Examples

Difference in Theme-Focus:

- (49) Q. I know the American amplifier has muddy treble. But which property does the British amplifier have?
A. The British amplifier has clean treble.
- (50) Q. I know the British speaker has muddy treble. But which property does the British amplifier have?
A. The British amplifier has clean treble.

IS Realization

- There are various means that can be used;
- The means can be used also in combination;
- Different languages employ and combine differently, depending on their typological characteristics;
- The means (strategies):
 - **intonation** (prosody); e.g., the predominant means in English
 - (word) **ordering**; e.g. the predominant means in Czech, also to some extent in German (“Mittelfeld”)
 - **morphological/grammatical marking**; e.g., particles ‘wa’ and ‘ga’ in Japanese
 - **syntactic constructions**, e.g. it-cleft, wh-cleft, passivization, etc.

IS Realization Means: Word Order

- Default order from Theme to Rheme
- Other orders typically combined with marking by other means (e.g., intonation)

(53) What does John hate?

John hates **COMICS**.

COMICS, John hates.

(Rheme before Theme)

(54) Wer liebt den Man?

Den Man liebt die **FRAU**.

Die **FRAU** liebt den Man.

(Rheme before Theme)

IS Realization Means: Syntax

Syntactic constructions such as cleft, pseudo-cleft, passivization, there-insertion

- (55) What does John hate?
It is COMICS that John hates.
- (56) Who hates comics?
It is JOHN who hates comics.
- (57) What does John hate? / What John hates are COMICS.
COMICS are what John hates.
- (58) Who hates comics?
Who hates comics is JOHN. / JOHN is (the one) who hates comics.
- (59) Who hates comics?
Comics are hated by JOHN.
- (60) What is in the garden?
There is a TROLL in the garden.

IS: Meaning Differences

- (61) a. Smoke OUTSIDE!
i.e., If you (want to) smoke, do it outside. or, where one smokes is outside.
- b. SMOKE outside!
i.e., If you are outside, smoke, or, what one does outside is smoke.
All-Rheme reading: There is smoke outside!

IS: Meaning Differences

- (62) a. STAFF behind counter.
i.e., Who should be behind the counter is (only) staff.
- b. Staff BEHIND COUNTER.
i.e., What staff should do is be (only) behind the counter.
or, where staff should be is (only) behind the counter.

IS: Meaning Differences

- (63) a. *On the Shetlands one speaks* ENGLISH.
i.e., English is the (only) language that is spoken on the Shetlands.
(correct)
- b. *One speaks English on the* SHETLANDS.
i.e., Shetlands is the (only) place where English is spoken.
(correct, but seems to give incomplete information)

Similarly with no change in structure, just intonation:

- (64) a. ENGLISH *is spoken on the Shetlands*.
b. *English is spoken on the* SHETLANDS.

In Czech, different word order has the same effect:

- (65) a. *Na Shetlandech se mluví Anglicky.*
On Shetlands refl speaks English
- b. *Anglicky se mluví na Shetlandech.*
English refl speaks on Shetlands

How about German? . . .

IS: Meaning Differences

Scope effects (=different truth conditions!):

- (66) a. *Everybody in this room knows two languages.*
i.e., Everybody in this room knows at least two, possibly different languages.
- b. *Two languages are known by everybody in this room.*
i.e., There are at least two languages such that everybody in this room knows at least these two languages.

IS: Meaning Differences

Scope effects (=different truth conditions!):

- (67) a. Officers always escorted BALLERINAS.
i.e., in all cases, what officers did was escort ballerinas
- b. OFFICERS always escorted ballerinas.
i.e., in all cases, whom ballerinas were escorted by were officers

“Focussing jokes”

- (68) Why do you rob banks?
Because that’s where the money is!
- (69) Why do firemen wear red suspenders?
To keep their pants up.
- (70) Why do we buy clothes?
Because we can’t get them for free.
- (71) 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

(68') Why do you rob BANKS?

(69') Why do firemen wear RED suspenders?

Why do firemen wear red SUSPENDERS?

(70') Why do we buy CLOTHES?

Why do we BUY clothes?

(71') Why do we dress GIRLS IN PINK and BOYS IN BLUE?

Why do WE dress girls in pink and boys in blue?

Computational Modeling of IS in Applications

- analysis in question/text understanding, MT or TTS
 - word order: [Hoffman1995]; [Styś and Zemke1995]
 - intonation: [Prevost1995]
 - anaphora resolution [Hajičová et al.1990, Hajičová et al.1992]
- production in NLG, MT, TTS or dialog systems
 - word order: [Hoffman1995, Hoffman1996]; [Kruijff-Korbayová et al.2002]
 - intonation: [Prevost1995]; [Kruijff-Korbayová et al.2003]; [Moore et al.2004]
 - referring expression generation in text [Hajičová et al.1990]; in dialogue [Krahmer and Theune2002]
 - embodied agents' gesture [Pelachaud et al.1998, Cassell et al.2000], gaze and turn-taking (Cassell-etal:1999)



References

- [Cassell et al.2000] Justine Cassell, Matthew Stone, and Hao Yan. 2000. Coordination and context-dependence in the generation of embodied conversation. In *Proceedings of the INLG Conference*.
- [Hajičová et al.1990] Eva Hajičová, Petr Kuboň, and Vladislav Kuboň. 1990. Hierarchy of salience and discourse analysis and production. pages 144–148.
- [Hajičová et al.1992] Eva Hajičová, Vladislav Kuboň, and Petr Kuboň. 1992. Stock of shared knowledge - a tool for solving pronominal anaphora. pages 127–133.
- [Hoffman1995] 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.
- [Hoffman1996] 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.
- [Krahmer and Theune2002] Emiel Krahmer and Mariet Theune. 2002. Efficient context-sensitive generation of referring expressions. In van Deemter and Kibble [van Deemter and Kibble2002], pages 223–264.
- [Kruijff-Korbayová et al.2002] 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 Kibble2002], 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>.

- [Kruijff-Korbayová et al.2003] 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.
- [Moore et al.2004] 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)*.
- [Pelachaud et al.1998] 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.
- [Prevost1995] 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.
- [Steedman2000] Mark Steedman. 2000. Information structure and the syntax-phonology interface. *Linguistic Inquiry*, 31(4):649–689.
- [Styś and Zemke1995] 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.



[van Deemter and Kibble2002] Kees van Deemter and Rodger Kibble, editors. 2002. volume 143 of *Lecture Notes*.
CSLI.