



Einführung in Pragmatik und Texttheorie

Information Structure and Common Ground

Ivana Kruijff-Korbayová

korbay@coli.uni-sb.de

<http://www.coli.uni-sb.de/cl/courses/PTT/>

Summer Semester 2004



Information Structure and Common Ground

Lecture Plan:

- Can speakers choose any information structure partitioning?
- Common ground and cognitive status
- Cognitive status and linguistic form
- Taxonomies of cognitive status
- So, can speakers choose any information structure partitioning?

Basic reading: Prince 1981



What is 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.
- Information Structure concerns a division of the expressed propositional content into:
 - that part which the speaker presents as reflecting (relying on) the (assumed) context**
 - that part which the speaker presents as affecting the context.**



Various notions of “givennes”

Prince (1981): A number of authors evoke some notion of “givennes” (e.g., old-new, known-new, presupposition-focus, given-new, etc.) in discussions of sentence-level phenomena, such as

- gapping
- dative (shift)
- pronominalization
- left/right dislocation
- (choice of) sentence subjects
- topicalization
- discourse structure

But, she observes they do not mean the same:
predictability/recoverability; saliency; knowledge



Taxonomies of Cognitive Status of Discourse Entities

- Chafe 1974, 1976 (see Lambrecht 1994, Ch. 3)
- Prince 1981
- Gundel et al. 1980, 1988, 1989, 1993 etc.
- Hajičová et al. 1982, 1993, etc.
- Centering theory: Grosz et al. 1983 ... 1995, etc.



Chafe's Taxonomy

- “knowing something and thinking something are different mental states”
- “not only being aware, but having easy access to a mental representation is important for interpretation”
- cognitive states of concepts in hearer's consciousness at utterance time:
 - active
 - semi-active (accessible)
 - * textual: deactivation
 - * inferential: from cognitive schema
 - * situational: presence in external world
 - inactive
- correlations between cognitive states and verbalization
 - active: lack of pitch accent, pronominal coding
 - inactive: accentuation, full lexical coding



Common ground

Hypothesis:

Since the Theme of an utterance is its point of departure, it should be something the speaker and the hearer “share”, therefore:

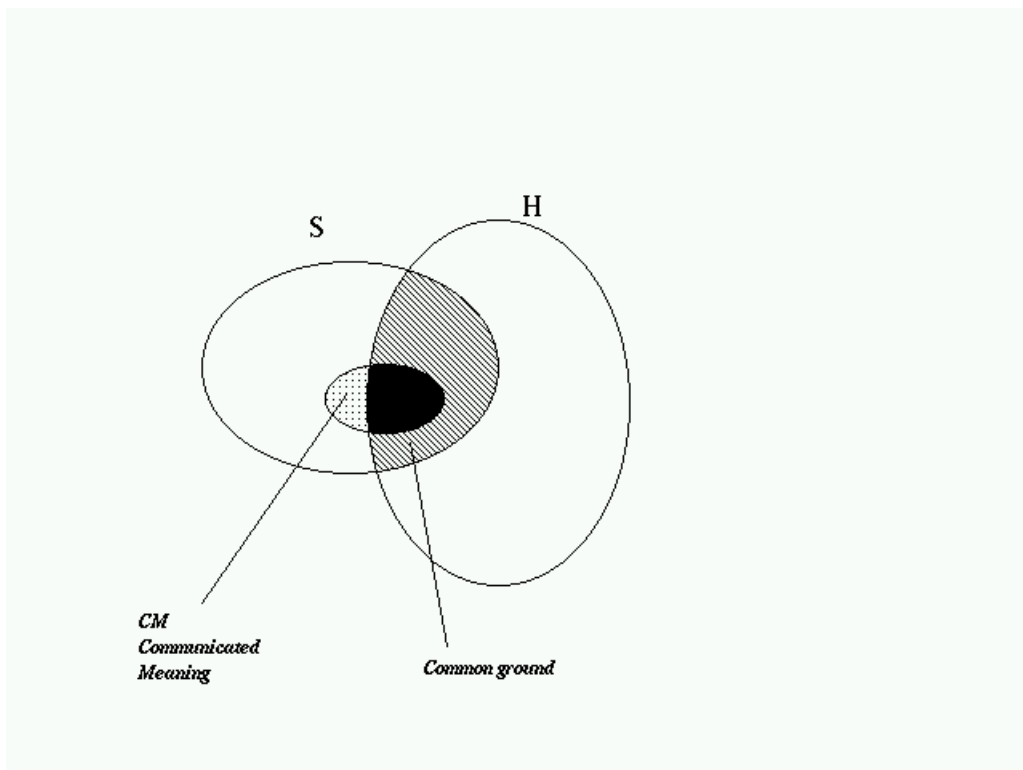
A cooperative speaker chooses as thematic entities in the **common ground**.

Questions:

- When exactly can an item be considered part of the common ground?
- Is the distinction between in vs. outside the common ground sufficiently finegrained to explain acceptable uses of various linguistic expressions?
- If not, what is the appropriate taxonomy of cognitive status?
- And, what is the taxonomy of available linguistic forms?
- What is the relation between the two?



Communicated Meaning & Common Ground





Prince's taxonomy of assumed familiarity

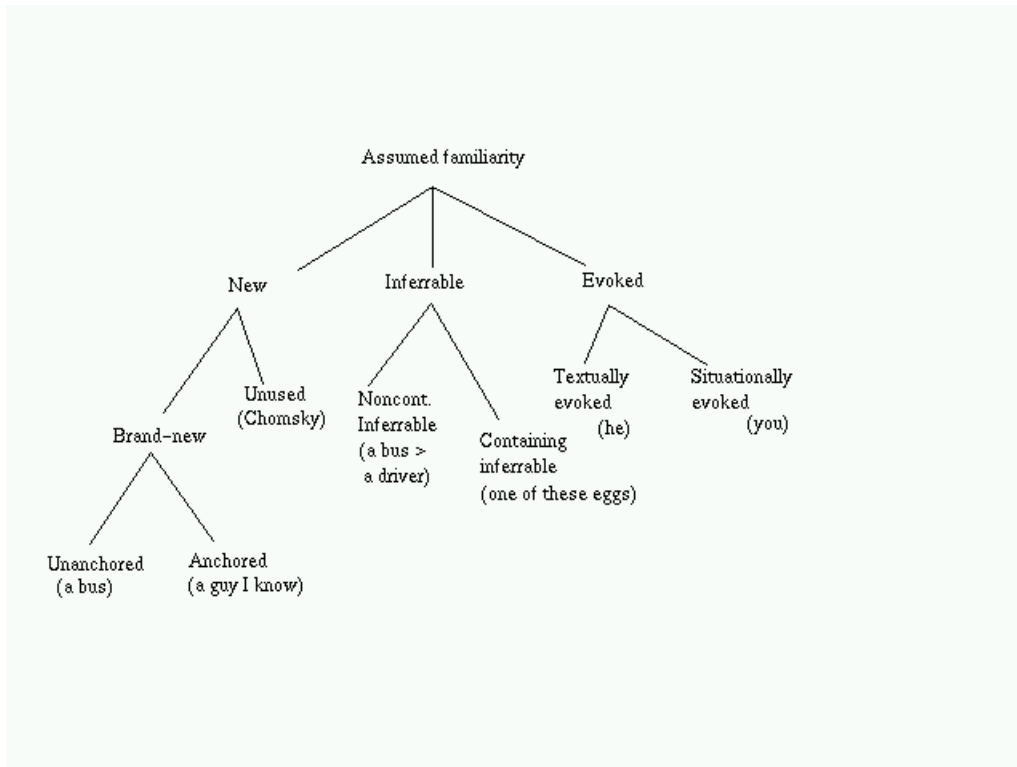
- *brand new*: create a new discourse referent for a previously unknown entity
- *unused*: create a new discourse referent for a known entity
- *inferable*: create a new discourse referent for an inferable entity
- *evoked* (textually or situationally): access an available discourse referent

Familiarity scale:

$$E/E^S > U > I > I^C > BN^A > BN$$



Prince's Taxonomy





Prince's Taxonomy: Examples

Brand new *I bought a dress.*

(Hearer creates a new entity in DM)

Unused *Chomsky is famous.*

(Hearer moves entity to DM.)

Inferrable *I went to the postoffice and the cleck sold me a stamp.*

(Hearer infers entity from an entity in DM.)

Containing Inferrable *One of these eggs is rotten.*

(Hearer infers entity from an entity denoted by containing NP.)



Textually evoked *Sue went to see her grandma and the sweet lady baked a cake for her.*

(Entity is already in DM because speaker gave hearer an instruction to put it there.)

Situationally evoked *The whiteboard is dirty.*

(Entity is in the situational context.)



Linguistic Forms

There is a variety of linguistic forms that can be used to refer to or describe entities, e.g.,:

- definite NPs: *the N*
- demonstrative NPs: *this/that N*
- indefinite NPs: *a/some/one N, Ns*
- quantified NPs: *every/five N, . . .*
- personal pronouns: *it, she . . .*
- proper names: *John*



Prince's Taxonomy

Familiarity scale: $E/E^S > U > I > I^C > BN^A > BN$

This scale can give rise to implicatures based on Grice's Maxim of Quantity, i.e., the use of a weaker expression implicates that the stronger expression would not have been appropriate/possible.

- (1) a. She
- b. Ellen
- c. One of the people that work at Penn
- d. A person that works at Penn
- e. A person
- . . . bought a Toyota.



Prince's Taxonomy

In informal conversational discourse, a tendency has been observed (in English) to reserve subjects for entities with higher familiarity; in other words, constructions are used which enable to keep entities with low familiarity out of subject position

(2) “run-on”

- a. I had a little boy, black, about ten years old, he . . .
- b. There's some male beauty shops, they . . .

(3) “deletion of subject relative markers”

- (2) We got a lot of fancy Cadillac cars don't tip.
- (3) I had a great-great-great-grandfather or something fought that Revolution.
- (4) There was a piece of four-inch bone never mended.



Information Structure \approx Linguistic Form \approx Familiarity Status



Gundel's Givenness Hierarchy

	Cognitive Status	Ling. Form
1	in focus	<i>it</i>
2	activated	<i>that, this this N</i>
3	familiar	<i>that N</i>
4	uniquely identifiable	<i>the N</i>
5	referential	<i>an N, this N</i>
6	identifiable type	<i>an N</i>

Claims:

- $1 \supset 2 \supset 3 \supset 4 \supset 5 \supset 6$
- The cognitive status of an item is a necessary and sufficient condition for the use of the corresponding ling. form.



Gundel's Givenness Hierarchy: Examples

Identifiable type H knows the meaning of the type being used; she can access a representation of the type described by N.

I couldn't sleep last night. A rabbit kept me awake.

Referential S refers to specific entity. H does not know which.

I couldn't sleep last night. A rabbit kept me awake.

Uniquely identifiable H can identify the S's intended referent.

I couldn't sleep last night. The rabbit kept me awake.



Gundel's Givenness Hierarchy: Examples

Familiar H uniquely identifies the intended referent because she has a representation of it in memory.

I couldn't sleep last night. That rabbit in the garden kept me awake.

Activated H has a representation of the intended referent in short-term memory.

I couldn't sleep last night. That kept me awake.

"That" = e.g., the rabbit's gnawing on carrots occurring at utterance time.

In focus (center of attention) H has a representation of the intended referent in the center of attention in short-term memory.

I couldn't sleep last night. That rabbit in the garden kept me awake. It gnaws very loudly.



Predictions of Gundel's Givenness Hierarchy

- A particular ling. form is inappropriate if the required cognitive status is not met.
- A form corresponding to a weaker cognitive status than the referent actually has can be used (e.g., *the N* for an entity in center of attention).

Tested and mostly verified on naturally occurring discourse for Chinese, English, Japanese, Russian and Spanish (the hierarchy has been tailored to the specifics of each language, e.g., Russian has no articles).

What where the failures like?



Predictions of Gundel's Givenness Hierarchy

Prediction 1:

All but few occurrences meet the necessary condition.

When not, two things can happen:

- H fails to identify the intended referent (and asks for clarification) A. *Do these go in here?* B. *These?*
(The referent of “these” was not activated.)
- H accommodates to be able to associate the correct referent with the form A. *Barb has it. I suspect she was a cat in some previous life. Oh, did I tell you that they have a cat?*
(Neither Barb's family wasn't in center of attention.)



Predictions of Gundel's Givenness Hierarchy

Prediction 2:

Although it is confirmed that forms coding higher status can be used, the distribution varies across different statuses for the various forms.

- Demonstrative pronouns (2) are rarely used for referents in center of attention (1).
- Demonstrative NPs (2,3) are rarely used for familiar referents (3).
- No occurrences of indefinite NPs (6) for referents with higher status than referential (5).

What explains this?

Conversational implicatures!!!



Failures of Predictions Explained

Maxim of quantity (Grice 1975)

- Q1: Say enough.
- Q2: Do not say too much.

Scalar implicatures: based on an entailment scale

- Q1: Use of weaker form implicates that stronger form does not hold:
 1. Use of indefinite (6) implicates referent not uniquely identifiable (4).

Jon is meeting a woman tonight.

2. Use of demonstrative pronouns (2) implicates referent not in center of attention (1).



Going back from the kitchen is a little hallway and across from the kitchen is a big walk-through closet. On the other side of {that/it} is another hallway.
(kitchen is if center of att.; “that” does not refer to kitchen, “it” does)

- Q2: Use of weaker form implicates that stronger form holds: Use of definite description (4) implicates familiarity (3).



Accessibility vs. Identifiability

Accessibility/Activation “awareness”, “easy access” (cf. Chafe)

Identifiability hearer’s ability to pick out a particular referent (“file”) from among all those which can be designated with a part. ling. expression, and identify it as the one the speaker intends

- no one-to-one correspondence between (non)identifiability and (in)definiteness
- other dimensions: specific vs. non-specific indefinite NPs; generic NPs

Identifiability and activation/accessibility are independent but correlated:

- unidentifiable are outside the activation parameter (Prince’s brand new)
- identifiable can be inactive, accessible or active



Summary

Information Structure \approx Linguistic Form \approx Familiarity Status

Gundel's predictions concern the choice of ling. form for a referent of a given cognitive status, concentrating particularly on various types of determiners (definite/indefinite article, demonstrative determiners, etc.). But, as we have seen, the properties of NPs, e.g., definiteness, are not directly related to IS partitioning.

We need predictions about IS partitioning as such, or about IS-related aspects of ling. form, e.g., following up on Prince's early work; also: the approach to activation by Hajičová et al. and the Centering Theory approach (although not directly about IS).



Hajičová's Activation Model

- degrees of accessibility
- no taxonomy (classes), continuous scale from “highly accessible” to “faded away”
- assignment of activation degree calculated taking into account IS partitioning:
 - Theme-expression → maintain activation degree
 - Rheme-expression → raise to top
 - referents not re-accessed directly fade quicker than referents indirectly accessed
- also relates degrees of activation and the appropriateness of using particular types of referring expressions (but not as detailed as Prince or Gundel)



Centering Theory

(Grosz, Joshi and Weinstein 1995)

- attentional limitations of discourse participants: they can only attend to a small number of referents at the same time
- local vs. global model of attention centering
- the Centering Theory concerns the local modelling of attention
 - each utterance has one backward looking center and an ordered set of forward looking centers
 - types of center-transitions depending on whether backward looking center is maintained or changed
 - preference for center continuation, or smooth (=gradual) shift