Lecture 4

IS and Information Status

Motivation

Recall:

- **Speakers organize their utterances** in a way that reflects their model of the context and their communicative intentions (the intended context change).
  - Theme (Topic): presented as reflecting (relying on) the (assumed) context
  - Rheme (Focus): presented as affecting the context, e.g., adding to it or modifying it

Questions:

- Can a cooperative speaker at a given point in discourse use any IS partitioning?
- What items can she present/realize as (part of) Theme/Given/CB?
- What items if any must she present/realize as (part of) Rheme/New/NB?
- What information status do items in context (common ground) have?

Outline

- Familiarity status according to Prince (Prince, 1981; Prince, 1992)
- Cognitive states according to Chafe/Lambrecht (Chafe, 1976; Lambrecht, 1994)
- Gundel's givenness hierarchy (Gundel, 1974; Gundel et al., 1993)
- Praguian model of salience based on TFA, e.g. (Hajičová, 1993; Hajičová et al., 1995a)
- Centering theory (Grosz et al., 1995)

Points to consider:

- taxonomy (classification) vs. hierarchy (partial ordering) of statuses
- statuses vs. the dynamics (i.e., how referents acquire and change status)
- incrementality (utterance-by-utterance vs. continuous)
- relation between statuses and linguistic form
- relation between statuses and IS

Prince's taxonomy of assumed familiarity

Information Status and Linguistic Form

(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.

Three notions of "givenness"

(Prince, 1981)

Givennes$_P$: Predictability/Recoverability:
S presumes H can predict or could have predicted that a particular linguistic item will or would occur in a particular position in a sentence.

Givennes$_S$: Saliency:
S assumes that H has or could appropriately have some particular entity in consciousness at the time of hearing the utterance.

Givennes$_K$: Knowledge:
A assumes that H "knows", assumes or can infer a particular entity (not necessarily thinking about it)

These notions are different, though not mutually independent. $G_P \subset G_S \subset G_K$

Prince's taxonomy of assumed familiarity

(Prince, 1981; Prince, 1992)

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

	Discourse-new	Discourse-old
Hearer-new	brand new	d.n.a.
Hearer-old	unused (inferrable)	evoked
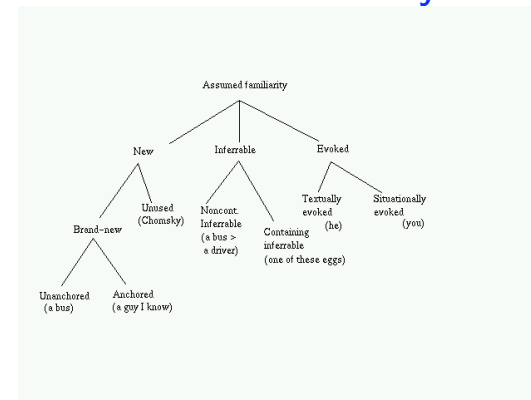
Prince's Taxonomy

Prince's Taxonomy: Examples

Brand new *I bought <u>a dress.</u>*   (Hearer creates a new entity in DM)
Unused <u>*Chomsky*</u> *is famous.*   (Hearer moves entity to DM.)

Inferrable *I went to the postoffice and <u>the clerck</u> sold me <u>a stamp.</u>*
    (Hearer infers entity from an entity in DM.)
Containing Inferrable <u>*One of these eggs*</u> *is rotten.*
    (Hearer infers entity from an entity denoted by containing NP.)

Textually evoked *Sue went to her grandma and <u>the sweet lady</u> baked a cake.*
    (Entity already introduced into in DM.)
Situationally evoked <u>*The board*</u> *is dirty.*
    (Entity already in DM because it is in the situational context.)

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.

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

The scale also seems to hold for NPs representing anchors within BN^A entities.

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

(109) "run-on"
    a. I had **a little boy, black, about ten years old**, **he** . . .
    b. There's **some male beauty shops**, **they** . . .
(110) "deletion of subject relative markers"
    a. We got **a lot of fancy Cadillac cars** ∅ don't tip.
    b. I had **a great-great-great-grandfather or something** ∅ fought that Revolution.
    c. There was **a peice of four-inch bone** ∅ never mended.

Prince's Taxonomy

(Prince, 1978) corpus-based findings:

- The presupposed part of a wh-cleft represents information that the speaker can assume the hearer is thinking about (hearer-old)

- In one variety of it-cleft, the presupposed part represents information the speaker assumes the hearer knows or can deduce but is not thinking about (hearer-new)

- In another variety of it-cleft, the presupposed part represents information the speaker takes to be a known fact, though not known to the hearer (hearer-new)

cf. also (Ward, 1988; Birner and Ward, 1998) for findings concerning preposing, postposing and argument reversal in English

Chafe's taxonomy of cognitive states

Chafe's Taxonomy

(Chafe, 1976), cf. also (Lambrecht, 1994)

- "knowing something and thinking something are different mental states"
- 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
- interested in correlations between cognitive states and verbalization:
  - active: lack of pitch accent, pronominal coding
  - inactive: accentuation, full lexical coding

Accessibility vs. Identifiability

(Lambrecht, 1994)

Accessibility/Activation "awareness", "easy access" (cf. Chafe)
Identifiablity 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

Gundel's Givenness Hierarchy

Gundel's Givenness Hierarchy

(Gundel et al., 1993) combine activation and identifiability in one hierarchy

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

Predictions: 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. <u>A rabbit</u> kept me awake.

Referential S refers to specific entity. H does not know which.
I couldn't sleep last night. <u>This rabbit in the garden</u> kept me awake.

Uniquely identifiable H can identify the S's intended referent.
I couldn't sleep last night. <u>The rabbit</u> 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. <u>That rabbit in the garden</u> kept me awake.

Activated H has a representation of the intended referent in short-term memory.
I couldn't sleep last night. <u>That</u> kept me awake.
"That" = e.g., the rabbit's gnawing on carrots ocurring 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. <u>It</u> 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).

Analysis of the mispredicted cases.

Predictions of Gundel's Givenness Hierarchy

Prediction 1:

All but few occurrences meet the necessary condition.

When not, two things are happening:

- 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?*
  (Barb's family was not in focus.)

Predictions of Gundel's Givenness Hierarchy

Prediction 2:

Although it is confirmed that forms coding weaker 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?

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. Idefinite NP (6) implicates referent not uniquely identifiable (4).
     Jon is meeting a woman tonight.
  2. Demonstrative pron. (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 in focus; "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).

The Praguian Model of Salience

The Praguian Model of Salience

The *stock of knowledge assumed by the speaker to be shared by the hearer* (SSK) comprises "not only knowledge in literal sense, but also a wide range of psychological phenomena including beliefs and other attitudes" (Sgall et al., 1986, p.55–56), (Hajičová, 1993, p.70)

- Dynamically changing stock of shared knowledge (SSK):

  - *Set of senses of linguistic expressions* ($\approx$ discourse referents)
  - *Activation*: SSK is partially ordered to reflect the degrees of salience of its elements (i.e., immediate accessibility in the hearer's memory)
  - *Dynamic character*: the set and the ordering are relativized to the time-points of utterances

- Activation of entities in SSK w.r.t. IS (Topic-Focus Articulation, TFA)

The Praguian Model of Salience

(Hajičová et al., 1990; Hajičová, 1993)

- Degrees of activation (salience):

  - continuous scale from "highly salient" to "faded away"
  - assignment rules take into account TFA partitioning and ling. form:

1. Reference by an expression in Focus $\rightarrow$ assign top activation
2. Reference by a noun phrase in Topic $\rightarrow$ assign high activation (not top)
3. Reference by a pronoun in Topic $\rightarrow$ maintain current activation
4. Not re-accessed referent $\rightarrow$ decrease activation
   Referents that remained highly activated (mentioned in Topic) fade away slower than referents only accessed in Focus
5. Entity *associated* with an activated referent $\rightarrow$ assign a lowered activation

(111)   i.   John went to his favorite music store to buy a piano.
$ssk(s_1, j) = 1; \ ssk(s_1, s) = 0; \ ssk(s_1, p) = 0$
     ii.   He had frequented the store for many years.
$ssk(s_2, j) = 1; \ ssk(s_2, s) = 1; \ ssk(s_2, y) = 0; \ ssk(s_2, p) = 2$
     iii.   He was excited that he could finally buy a piano.
$ssk(s_3, j) = 1; \ ssk(s_3, s) = 2; \ ssk(s_3, y) = 2; \ ssk(s_3, p) = 0$
     iv.   He arrived just as the store was closing.
$ssk(s_4, j) = 1; \ ssk(s_4, s) = 1; \ ssk(s_4, y) = 4; \ ssk(s_4, p) = 2$

(112)   i.   John went to his favorite music store to buy a piano.
$ssk(s_1, j) = 1; \ ssk(s_1, s) = 0; \ ssk(s_1, p) = 0$
     ii.   It was a store John had frequented for many years.
$ssk(s_2, j) = 1; \ ssk(s_2, s) = 1; \ ssk(s_2, p) = 2; \ ssk(s_2, y) = 0$
     iii.   He was excited that he could finally buy a piano.
$ssk(s_3, j) = 1; \ ssk(s_3, s) = 2; \ ssk(s_3, p) = 0; \ ssk(s_3, y) = 2$
     iv.   It was closing just as John arrived.
       it=piano

The Praguian Model of Salience

Predictions Based on Activation:

- A cooperative speaker chooses as CB items (Topic) what is already at/near the top of the activation scale; in contrast, NB items (Focus) can but don't have to be already activated.

- Appropriate use of linguistic forms according to activation.

- *Discourse topic*: an entity at/near top of activation scale (through a stretch of discourse)

- *Discourse segment*: a stretch of discourse with relatively "constant" constellation at/near the top of the activation scale

Salience and Reference Resolution

(Hajičová et al., 1990) propose the following rules for reference resolution:

1. null-subject → preceding subject
2. relative pronoun → head of preceding noun phrase
3. weak pronoun → preceding topic
4. strong or demonstrative pronoun → preceding focus
5. apply salience in SSK to decide between competitors after (1-4) using salience
6. if no item activated enough or difference too small, no resolution

Reported precision of 80-85% just with rules (1-4),
SSK is in addition, but no evaluation results presented.

Determination of TFA?

Salience and Reference Resolution

- (Hajičová et al., 1995a): refinements for different pronominal forms and recursive CB/NB

  - group 1: CB dependents of main verb (in T)
  - group 2: NB dependents of main verb (in F)
  - group 3 and 5: CB dependents deeper below (in T/F)
  - group 4 and 6: NB dependents deeper below (in T/F)

- (Hajičová et al., 1992): algorithm based on the above, also taking simple syntactically encoded associative links, repeated reference and distance into account; set of weighted factors; looking for optimal solution.

Determination of CB/NB?
No evaluation results presented.

Salience and Reference Production

(Hajičová et al., 1990) propose the following rules for deciding between reference by pronominal or nominal expression:

1. if intended referent X not salient above a MIN threshhold, then use an NP
   else,
2. if X has no (recent) competitor, then use a weak pronoun
   else
3. if neither X nor competitor(s) have MAX salience, then use an NP
   else
4. if X was and remains subject, then use a weak pronoun
   else if X becoming subject, then use a strong pronoun to refer to non-subject
   else use an NP

No evaluation results presented.

Salience and Reference Production: Example

d_1: black chihuahua; d_2: white chihuahua
$ssk(s_0, d_1) = MIN, ssk(s_0, d_2) = MIN$

(113) i. The white chihuahua was angry.　　　　　(Rule 1)
   　　　　$ssk(s_1, d_2) = 1$

   ii. It viciously attacked the black chihuahua.　　(Rule 2)
   　　　　$ssk(s_2, d_1) = 0; \; ssk(s_2, d_2) = 1$

   iii. { It / <u>This one</u> / The chihuahua } barked loudly.　　(Rule 4)

Salience in GRE

(Krahmer and Theune, 2002) propose a modification of the *Incremental Algorithm* for generating referring expressions (Dale and Reiter 1995)

- the IA determines what properties to include in a referring expression to uniquely determine an object X w.r.t. to a context set (set of objects in discourse domain from which intended referent X needs to be distinguished)

- use salience to determine the context set: set of salient objects

- salience weighting according to the Praguian model (Hajičová, 1993) and Centering Theory (Grosz et al., 1995) compared

- combined approach: continuous values, subject-preservation preference

Salience in GRE: Example

d_1: black chihuahua; d_2: white chihuahua

(114)   i.   The white chihuahua was angry.
$$ssk(s_1, d_2) = 1$$
$$C_f(s_1) = \{d_2\}$$

    ii.   It viciously attacked the black chihuahua.
$$ssk(s_2, d_1) = 0;\ ssk(s_2, d_2) = 1$$
$$C_f(s_2) = \{d_2, d_1\}, C_b(s_2) = d_2$$

    iii.   { It / This one / The chihuahua / The black dog } barked loudly.

Centering Theory

Centering Theory

(Grosz et al., 1995)

- Attempt to
  - account for attentional limitations of discourse participants: they can only attend to a small number of referents at the same time
  - reduce inference load in the process of discourse interpretation (i.e., more likely candidates for coreference considered first)

- There are local and global aspects of attention centering (e.g., based on overall task structure or communicative goals)

- CT is a computational model of local centering of attention

Centering Theory

Compare:

(115) a. John went to his favorite music store to buy a piano.
      b. He had frequented the store for many years.
      c. He was excited that he could finally buy a piano.
      d. He arrived just as the store was closing.

(116) a. John went to his favorite music store to buy a piano.
      b. It was a store John had frequented for many years.
      c. He was excited that he could finally buy a piano.
      d. It was closing just as John arrived.

Centering Theory

- each utterance has one backward looking center C_b and an ordered set of forward looking centers C_f

- proposed C_f ordering $Subj < Obj < Other$
  various other proposals: syntactic function, theta-role, surface order, familiarity status (Walker et al., 1998)

- the most highly ranked item on C_f is the C_p, i.e., the preferred C_b for the next utterance

- types of center-transitions depending on whether backward looking center is maintained or changed: continuation, retaining, shift

---

Centering Theory

Rule 1 If any item is pronominalized, then the C_b is pronominalized
Rule 2 preference for sequences of center transitions: continue > retain > shift

Centering transition types:

	$C_b(U_i) = C_b(U_{i-1})$	$C_b(U_i) \neq C_b(U_{i-1})$
$C_b(U_i) = C_p(U_i)$	Continue	Smooth Shift
$C_b(U_i) \neq C_p(U_i)$	Retain	Rough Shift

$C_b(U_k)$ — backward looking center of utterance U_k
$C_f(U_k)$ — (partially) ordered list of forward looking centers of utterance U_k
$C_p(U_k)$ — highest ranked item on $C_f(U_k)$, the preferred (next) center

---

Centering theory: Examples

(117) a. John has had trouble arranging his vacation.
        $C_f = John > trouble > vacation$
      b. He cannot find anyone to take over his responsibilities.
        $C_b = John,\ C_f = John > responsibilities,\ Cont(U_a, U_b)$
      c. He called up Mike yesterday to work out a plan.
        $C_b = John,\ C_f = John > Mike > yesterday, plan,\ Cont(U_b, U_c)$

(118) d. Mike has annoyed him a lot recently.
        $C_b = John,\ C_f = Mike > John,\ Shift(U_c, U_d)$
      e. He called John at 5 A.M. on Friday last week.
        $C_b = Mike,\ C_f = Mike > John > 5AM, Fri, lw,\ Cont(U_d, U_e)$

(119) d'. He has been pretty annoyed with Mike recently.
        $C_b = John,\ C_f = John > Mike,\ Cont(U_c, U_{d'})$
      e'. He got a call from him at 5 A.M. on Friday last week.
        $C_b = John,\ C_f = John > Mike > 5AM, Fri, lw,\ Cont(U_{d'}, U_{e'})$

(120) a. John went to his favorite music store to buy a piano.
$C_f = John > store > piano$
b. He had frequented the store for many years.
$C_b = john, \ C_f = John > store > years, \ Cont(U_a, U_b)$
c. He was excited that he could finally buy a piano.
$C_b = John, \ C_f = John > piano, \ Cont(U_b, U_c)$
d. He arrived just as the store was closing.
$C_b = John, \ C_f = John > store, \ Cont(U_c, U_d)$

(121) a. John went to his favorite music store to buy a piano.
$C_f = John > store > piano$
b. It was a store John had frequented for many years.
$C_b = store, \ C_f = store > John > years, \ Shift(U_a, U_b)$
c. He was excited that he could finally buy a piano.
$C_b = John, \ C_f = John > piano, \ Shift(U_b, U_c)$
d. It was closing just as John arrived.
$C_b = piano, \ C_f = piano > John, \ Shift(U_c, U_d)$

Centering Theory

- Comparison with SSK (Kruijff-Korbayová and Hajičová, 1997): backward-looking center $\approx$ least communicatively dynamic Topic-item; but: not only pronominalized, continuous salience values, fading away; differences in ordering

- other comparisons and integrations, e.g., (Krahmer and Theune, 2002; Styś and Zemke, 1995)

- There exist various implementations, e.g., for NLG and for anaphora resolution.

- Corpus annotation: GNOME project (Poesio, 2004)

Summary

- various taxonomies of information status exist

- proposed to explain certain distributions of linguistic forms, e.g., pronominalization, NP forms, marked syntactic constructions, prosody

- information status needs to be considered orthogonal to information structure:
  - there is a tendency for entities referred to in Theme to be given/familiar/contextually bound to some degree,
  - particular information status is neither sufficient nor necessary for an entity to be referred to in Theme

  cf. also (Reinhart, 1981) for a discussion