Information Structure According to Hendriks & Dekker; Vallduví

Elizabeth Allyn Smith

LING795K, OSU, Spring 2005

Hendriks and Related Approaches

- Background for Understanding Hendriks

  - File Change Semantics (Heim 1983)
  - Discourse Representation Theory (Kamp 1984)
  - Information Packaging (Vallduví 1995?)

- Hendriks' Approaches...
  - ...with Dekker (1996)
  - ...Cross-linguistically (1999)

File Change Semantics

- Developed in order to return to a previous notion (i.e. Jespersen 1949) of the difference between definite and indefinite noun phrases:

  'A definite is used to refer to something that is already *familiar* at the current stage of the conversation. An indefinite is used to introduce a *new* referent' (Partee & Portner 2002: 223)

- Heim accomplishes this via the file card metaphor where each individual in the discourse has its own card. Then definites/indefinites differ according to their file-keeping instruction:
  - Indefinites: Start a new card
  - Definites: Update an old card

Discourse Representation Theory

- Developed in order to account for anaphora and so-called Donkey Sentences such as:

  Every farmer who owns a donkey beats it

  Problem: 'a donkey' should require an existential quantifier (i.e. there exists a donkey), but instead has a universal interpretation: for all donkeys owned by some farmer, they are all beaten by him.

- While these were the empirical issues, this work also marked an attempt to integrate truth-theoretic and model-theoretic semantics, i.e. 'a sentence S or discourse D with representation m is true in a model M iff M is compatible with m; and compatibility...can be defined as the existence of a proper embedding of m into M, [which is]...a map from the universe of m into that of M which...preserves all the properties and relations which m specifies of the elements in its domain' (189)

DRT Continued

- The Discourse Representations themselves (represented by boxes) are formed in response to discourse according to certain rules. These rules operate on syntactic structures, but not all are syntactic in nature (those dealing with anaphor, for example, are considered to be semantic) (194)

- An example: If Pedro owns a donkey, he beats it

Vallduví's Information Packaging

- Assumes the tripartite distinction we've seen before:

topic	comment	comment
link	tail	focus
ground	ground	focus

- Information Structure purpose:
  (loosely quoted from H&D, 29)
  - Link: directs hearer to a location in her info. structure
  - Tail: signals a particular mode of info. update
  - Focus: encodes information to be added

Hendriks' Interpretation

- Though Vallduví doesn't state it as such, his proposal identifies four information structures in English:
  1. focus
  2. focus-tail
  3. link-focus
  4. link-focus-tail

- These correspond (as stated in H&D) to the following functions:
  1. update-add (I)
  2. update-replace (I, record(fc))
  3. goto (fc) (update-add (I))
  4. goto (fc) (update-replace (I, record (fc)))

(where 'I' here represents their I_s, which is the focus information)

Additional Commentary

- One one hand, delineating the various information structures is nice in that it corresponds to what we know is true, i.e. that people often respond with single-word answers rather than complete sentences, which often seems to be assumed in accounts of focus (although many theories do specifically discuss cases where the entire sentence is focused, as Detmar pointed out)

- On the other hand, how are these four types flexible enough to allow later sentences, such as:

  (4) I hate to spoil the fun, but [$_L$John][$_T$is][$_F$not][$_T$dead]

  Is this considered 2 tails, or one split tail? (NB: In class, Detmar raised the question of whether this example contained a typo)

Intonational Comparison of Steedman and Henriks et al.

- Though H&D aren't explicit about intonation in the 1996 paper, they are assuming Vallduví's analysis (labeled below as Hendriks et al.)
- **Steedman**

  - Focus (his *rheme)* can be one of several tunes, but usually H* L or H* LL%

  - Ground (his *theme*) is associated with a L+H* LH% tune

- **Hendriks et al.**
  - Focus is always marked with a H*
  - Link (crucially not the entire Ground) is always marked with a L+H*
  - Tail (the other part of Ground) is always deaccented

How the theory works

- Example (H&D's (1) and (2)):

  (1) [_L_ The **president**][_F_ hates the Delft CHINA SET]

  (2) [_F_ He always uses DISPOSABLES]

- The link tells us to find the president's file card: #125
- The focus tells us to add information to the card. This information is 'hates object #136 (i.e. Delft China set)
- The next sentence has no go-to link, so it is interpreted as adding some information to the same file (#125). This time it adds 'uses object #147 (disposables)'

Problems (according to Hendriks)

- In the following example, pretend that Speaker A believes that Kasparov beat Karpov in Linares (thus, it is recorded on his file card)

A: Karpov was beaten by Kasparov, so he may also be beaten by Timman

B: I hate to spoil the fun, but [_L_ **Karpov**][_T_ was][_F_ NOT][_T_ beaten by Kasparov]

- The update-replace theory detailed here would lead A to try to find a 'beaten-by(kasparov)(karpov)' entry on the Karpov file card, but what he actually needs to do is to correct the record on the Kasparov file card (page 35)
- Adriane noted that the active/passive forms of each verb on the file cards is an additional practice that is not justified in these articles.

Difference between DRSs & FCs

- DRSs lack a marked discourse referent corresponding to the 'current locus of update,' which they add a marker for indicating (thus they are supposed to be equally able to account for the 'data')

- While the authors are not arguing that the file card metaphor is not possible, they are merely stating that Vallduví does not give sufficient evidence to prefer that metaphor to the discourse representation notion.

Hendriks (1999)

- Now that we've seen H&D's critique of Vallduví, what do they propose instead?

- Hendriks (1999) argues against Steedman's CCG proposal because it cannot account for focus cross-linguistically

- Following Vallduví, he contrasts languages like English (where pitch accents are the primary conveyor of focus) with those like Catalan or Turkish, where syntax plays a larger role and intonation is secondary.

- He likewise argues against the Multiset CCG proposal of Hoffman (1995) for erring on the other side: it accounts only for languages such as Catalan and not those similar to English.

Initial Assumptions

- Noting that languages use both strategies (intonation and word order) to varying degrees, Hendriks wants to deal with both at the same level of the grammar

- His point of departure, then, is a type of logical grammar based on the non-associative Lambek calculus, which means that the rules in (26)-(29) but not (30) apply; really, he uses a "doubling" of this calculus, as we will see.

- Signs are the basic units of the formalism, and they are represented like this: intonational term <| type |> informational term

- He dismisses the notion that focus-sensitive operators (like *only* in English) take scope over non-focused parts of the sentence so that he can make all-focus sentences the basic case from which he derives others

Core of Argument/Miscellaneous Points

- From p. 141: "Instead of focusing operators, then, the present treatment of information packaging employs 'defocusing' operators, heither-order functors that license the presence of links and tails. In...Catalan, such defocusing operators are realized as clitics and agreement morphology, whereas in English they are abstract items that merely have intonational repercussions"

- Note: The notation used in this theory looks like Steedman's CCG but is not actually the same. In particular, it is not always true (as it is in Steedman's work) that the argument comes to the right of the slash.

- My question from last time (echoed by Detmar) was 'how does focus projection work?' (or 'does it exist in this framework?', 'what needs to be said instead?,' etc.) is (not) answered on page 147 in footnote 8: "Discussion of this phenomenon will have to be resumed on another occassion" (sic)

More Miscellany

- Link and Tail are taken to be non-logical constants, but their formal interpretation is not given in the article. Hendriks provides a reference to some of his earlier work where he apparently discusses this (148)

- In order to account for the English data, the unary modals of Moortgat must be added (thus the system is multimodal rather than combinatory categorial grammar) (152)

- The essence of the proposal for English is that the vp undergoes a transformation into something that requires arguments with specific pitch accents (i.e. 'I take a L+H* subject np to my left')

- This is accomplished with the abstract defocusing operator, epsilon. It is important to see that this yields many different epsilons; one for each possible prosodic subcategorization for the verb.

Bibliography

Heim, I. (1983) File Change Semantics and the Familiarity Theory of Definiteness. In Bauerle, Schwarze, & von Stechow (eds), *Meaning, Use and the Interpretation of Language*. Walter de Gruyter, 303-323.

Hendriks, H. (1999) The Logic of Tune: A Proof-Theoretic Analysis of Intonation. In Lecomte, Lamarche, and Perrier (eds), *Logical aspects of computation linguistics: Second International Conference, LACL '97*. Berlin: Springer.

Hendriks, H. & Dekker, P. (1996) Links without Location. In P. Dekker & M. Stokhof (eds), *Proceedings of the Tenth Amsterdam Colloquium*. Amsterdam: ILLC, 339-358.

Hoffman, B. (1995) *Integrating 'Free' Word Order Syntax and Information Structure*. ms.

Kamp, H. (1984) A Theory of Truth and Semantic Representation. In Groenendijk, Janssen, & Stokhof (eds), *Formal Methods in the Study of Language*. Amsterdam: Mathematische Centrum.

Pierrehumbert and Hirschberg (1990). The Meaning of Intonational Countours in the Interpretation of Discourse. In Cohen, Morgan & Pollack (eds), *Intentions in Communication*. Cambridge: MIT Press

Vallduví, E (1995?) *Information Packaging: A Survey*. Report prepared for the European Research Project 'Word order, prosody, and information structure.'