Back to Part III

5. Annotation of information structure

In data which do not differ much from the established examples from the theoretical or experimental (e.g. laboratory phonology) literature, focus can be reliably identified as the answer to an explicit question, as the element which "associates" with a focus-sensitive particle (only, also, even, again etc.), or as an item that occurs in a parallelism. However, these simple identification methods quickly loose their grip when it comes to the annotation of natural language data such as news or spontaneous speech.

5.1 Focus, the Question under Discussion, and not-at-issue content

Focus (as a semantic, formal pragmatic notion) is characterized as the material selecting between the alternatives introduced by the explicit or implicit Question Under Discussion (QUD). Yet, in addition to the material necessary to select between the alternatives, there can be additional material that formally is part of the focus but is attributing additional information, not selecting between alternatives. Consider, for example, the bold material in the following example:

(5.1.1)

a. Wer war 1985 Bundeskanzler der BRD?

b. 1985 war Helmut Kohl Bundeskanzler, **der in seinen 16 Jahren an der Spitze der Regierung einerseits für die Wiedervereinigung aber andererseits auch für eine Politik der schwarzen Kassen stand.**

It seems that in order to make the notion of focus sufficiently precise, the analysis and annotation of focus should differentiate between material selecting between alternatives and additional information provided on already selected alternatives. The missing distinction could potentially be informed by the difference between "main structure" and "side structure" of an answer introduced by Klein & Stutterheim (2002, sec. 3.1). This work stems from an independent strand of research on "Quaestio" (Klein & von Stutterheim 1987), a concept apparently closely related to the notion of Questions under Discussion – a potentially fruitful connection which does not yet seem to have been explored.

The previous example also raises the question about the domain in which focus is

assigned. There are two subquestions here. First, we need to consider whether the QUD defining the focus can change during one sentence. In (5.1.1b), after Helmut Kohl is uttered, the implicit QUD may be analyzed as changing to What can you tell me about Helmut Kohl? But under what exact conditions does this happen? It seems that this question can only be answered if we answer the question about the domain for which a QUD is established and the relevance of sentences in this context: Second, what is the status of a sentence for information structure in general? For example, if we split the sentence (5.1.1b) into two main clauses, as in (5.1.2), does the information structure analysis change – given that each separate sentence traditionally is assumed to have a separate focus?

(5.1.2) 1985 war Helmut Kohl Bundeskanzler. In seinen 16 Jahren an der Spitze der Regierung stand er einerseits für die Wiedervereinigung aber andererseits auch für eine Politik der schwarzen Kassen.

The general question here, which needs to be answered for reliable focus annotation, could be formulated as: What is the connection between a sentence as a syntactic notion and text unit and the formal pragmatic domain containing a focus as an answer to a question under discussion?

== 5.2 Relating information structure (focus/background) and information status (given/new) === There seems to be a consensus that the focus can consist of new and given material, and the background of given material. But can there be new material in the background, or is all new material part of the focus? This does seem to be the case. Consider the question answer pair below, in which the answer (5.5.1.b) contains new information that is not related to the focus of the question.

(5.2.1) a. Was war die größte Leistung von Helmut Kohl?

> b. Die größte Leistung der Birne, **die mich entsetzlicherweise die ganze Kindheit und Jugend als Kanzler begleitet hat, war wohl die deutscheWiedervereinigung.**

Alternative questions and implicit Questions Under Discussion:

Should material explicitly mentioned in an alternative question be analyzed as given in the answer? It would be possible to analyze the answer (5.2.2b) to the alternative question (5.2.2a) as answering the implicit QUD in (5.2.3) instead.

(5.2.2) a. Would you like tea or coffee?

> b. I'd like tea.

(5.2.3) What would you like to drink?

Positing such an implicit question under discussion for alternative questions may also receive support from the fact that one can answer an alternative question such as (5.2.2.a) with answer that is not one of the explicitly given alternatives, as in (5.2.4).

(5.2.4) I'd like a hot chocolate, if possible.

Analyzing the focus to an alternative questions with the help of an implicit question under discussion in this way would be in line with annotation schemes, such as that of the Prague Dependency Treebank Hajičová et al. (2000) assuming that the focus cannot consist entirely of Contextually Bound (i.e., Given) material.

5.3 Information structure annotation of non-compositional expressions

How does one annotate the information structure of (the parts of) non-compositional expression? Some non-compositional expressions seem to behave just like expressions with a regular, compositional semantic contribution. For example, the support verb construction die Tränen kommen (the tears come) in the example (5.3.1b) derived from the CREG corpus (Ott et al. 2012), behaves just like the simple compositional sentence (5.3.1c) with the predicate weinen (weep).

(5.3.1) a. Was war mit Wuschel los?

    b. Ihm [kamen die Tränen, weil die Stones Platte kaputt ging]_F.

    c. Er [weinte, weil die Stones Platte kaputt ging]_F .

But consider the status of *Augen* in the following example (taken from Zoo im ersten Stock von Peter Paul Hilbert, Ravensburger Taschenbücher 1964 Herold Verlag Stuttgart, S. 86):

(5.3.2) a. Obwohl unsere neue Wohnung in der Straße des 7. Juli wunderschön war, gefiel mir das Leben in der Stadt nicht besonders. Gewiß, wir wohnten jetzt in einem richtigen Steinhaus an einer gepflasterten Straße, so wie Mum und Josephine es sich immer gewünscht hatten.

    b. Aber in meinen Augen war das alles nicht mit Sacramento zu vergleichen.

At the prosodic level, clearly the pitch accent in the PP in meinen Augen can only be on the possessive meinen, not on Augen. But why is Augen deaccented? It is definitely not previously mentioned. In meinem Augen means something like in my view, but even

this non-literal meaning is not explicitly introduced.

5.4 Echo questions

Focus is generally formalized as selecting an alternative in the denotation. However, there are cases where the denotation is clear for the speaker and the hearer and instead the alternatives arise at the level of the language itself, not its denotation:

(5.4.1) a. Schau mal, das ist ein Triptychon.

      b. Das ist ein **was**?

Relatedly, what is the focus of echo questions in general?

(5.4.2) a. Also, Du suchst den Weg zur Bonnemannstrasse.

      b. Ich suche den Weg **zu welcher Strasse**?

5.5 Inter-annotator agreement for focus

Inter-annotator agreement: Calculating IAA presupposes a set of pre-determined markables in order to de ne the total number of choices annotators need to make. This point is often not described in detail in publications on focus annotation, so it is unclear what the IAA gures really mean. Independently, it would be helpful to get input on what agreement measures people have been using.

5.6 Aboutness Topic

Cook and Bildhauer (2011, to appear) identify two areas where they could see that difficulties with the annotation of topics arose. The first area involves utterances containing two suitable candidates for topic status within one single utterance. Often one of the candidates is structurally more prominent (e.g. in initial position) but is less prototypically topical with respect to other features (e.g. animacy, grammatical function, definiteness etc.). The other possible topic expression can be said to "compete" with the structurally more prominent one for topic status. In many of these cases, the annotators felt it was rather arbitrary to decide on a clear-cut aboutness topic for the utterance. The decision as to which expression should be considered topical in such cases has perhaps to do with the notion of perspectivization. The other problematic area concerned the decision whether an utterance had no aboutness topic (i.e. was a thetic utterance) or whether there was a topic. The annotators came across utterances which did not really appear to be about a particular referent (i.e. which seemed to lack an aboutness topic) but

which did not fit the standard definition of thetic utterances either. In the discussion section of the second study, Cook & Bildhauer (to appear) suggest that maybe it is an idealization to claim that all utterances (except thetics) necessarily encode a topic-comment articulation and that topicality is perhaps instead to be viewed as a gradient notion with some utterances encoding the topic-comment articulation prototypically while other utterances do so to a lesser degree. The authors also raise the possibility that overlapping a level of topic annotation with a level of annotation of discourse relations or rhetorical relations may prove helpful in future work. They suggest that it is plausible that some discourse relations may encode topic-comment articulations whereas others perhaps do not, or do so only to a lesser degree. Given the view of the topic-comment articulation as a particular kind of speech act (cf. Jacobs (2001), Krifka (2008)), this seems to be an interesting area for future developments.

References

RAIS/DocumentPartIV (zuletzt geändert am 2013-04-18 15:55:34 durch ArndtRiester)

- IMS home
- wiki-support(at)ims
- Legal Notice