A Computational Approach to Topic and Focus in a Production Model

Vincenza PIGNATARO
Fak. f. Linguistik u.Literaturwiss.
Universität Bielefeld
Postfach 8640
D—4800 Bielefeld 1

Abstract: The aim of the presented research is the development of a linguistic model of the functional concepts topic and focus that can be used in natural language processing systems. The paper deals with two points of investigation: the first point concerns the identification of the topic and focus of an utterance. Within the frame of the linguistic discussion on such concepts topic and focus will be considered as semantic, pragmatic and intonational rather then as syntactic phenomena. An operational definition of topic and focus is obtained on the basis of basic semantic-pragmatic categories which are defined in relation to a specified context. The second point concerns the integration of the topic and focus identification rules in a system for natural language generation. The aim of the application is the validation of the developed topic and focus model with respect to some aspect of the generation process like thematic progression and accent mapping. Moreover the identification of topic and focus can be used to make prediction about the thematic progression and the accent mapping in the blocks world texts. For the prediction of indefinite pronouns like "one" and of definite articles within a noun phrase it is necessary to recur to the semantic-pragmatic categories.

1. Topic and Focus identification

The importance of contextual factors for the whole comunication process and for the subprocesses running in parallel like the distribution of information and the marking of topic and focus is the reason for attempting a definition of the "context" for this restricted domain. This new approach is designed to replace the traditional simple question criterion used in determining the topic and focus of single sentences and make a genuine semantic-pragmatic definition of topic, salient topic and focus possible. It extends certain ideas of the Prague School /cf. Hajičová,Vrbová 1982, Hajičová,Sgall 1885, Sgall et al. 1986/.

The present generation model is developed as a simulation of the generation of simplified German texts taken from blocks world experiments, in which a speaker has to advise the hearer on how to build a pyramid, a bridge and a façade. The underlying discourse model consists of the action sequence TAKE, PUT and PROVE, which was found to be constant in the produced conversations. The number of blocks involved in the TAKE action determines the number of the following PUT actions.

1.2 Context Definition

First a language L_1 is defined, which allows a description of the world of the experiments using statements like green(b_1), on(b_1,b_2), UHB(b_1) meaning that b_1 is an element of the unordered set of the hearer's blocks. But for the communication process knowledge and experience of the partecipants are as important as the tangible things around them. Therefore a language L_2 is defined, which allows speaking about the assumptions of the speaker about the hearer and his world. In fact in the experiment the speaker and the hearer do not see each other; they mainly rely on assumptions about their mutual knowledge /Schiffers 1972/. Statements of L_2 are for ex. *amk*(a_1) where *amk*is a property symbol of L_2 and a_1 is the statement on(b_1,b_2). *amk*(on(b_1,b_2)) means that it is assumed to be ***mutual knowledge*** that the block b_1 is on the block b_2. Other examples of statements of L_2 are *id*(b_1) meaning that the block b_1 has been ***identified*** in a TAKE action, *ueq*(TAKE) meaning that TAKE can be ***interpreted unequivocally*** and *apl*(l_1) meaning that l_1 is assumed to be a ***potential position*** for the moved blocks.

For practical reasons we consider the context C to be a pair of sets of statements in the language L_2: $<CO_1, CO_2>$. CO_1 contains statements about the world of the speaker and those assumptions about mutual knowledge which remain unchanged during an experiment. CO_2 contains the speaker's assumptions about the hearer's blocks, their actual and their potential positions.

1.2 Operationalisation of the categories assignment

The units of analysis are semantic representation of utterances from the blocks world texts. To every element of the semantic representation some semantic categories will be operationally assigned. In this session a formal definition of the units of analysis and of the operational rules is given. Every semantic representation of an illocutionary plan is an ordered set IP = $<x_1,...x_n>$, where x_1 is the verb and the remaining elements x_2 to x_n correspond to the elements of the case frame of the verb x_1. For every element x of IP there is an individual constant in the language L_2 referred to as x*. The assignment of semantic-pragmatic categories to the elements of IP is a function, which maps every pair (x,C), where x $\in$ IP and C is the context, onto the semanticäpragmatic categories of x, representing the status of x with respect to C.

The contextual labels are given (*g*), chosen (*ch*), mentioned (*m*), mentioned in the previous sentence (*mp*) and their negation $\neg$*ch*, $\neg$*m*, $\neg$*mp*. $\neg$*g* does not occur. These symbols build the alphabet A={*ch,m,mp,* $\neg$*ch*, $\neg$*m*, $\neg$*mp*}.

The operationalisation criteria are:

(1) If x $\in$ IP, then:

(i) if there is a property y of L_1 such that y(x*) $\in CO_1 \cup CO_2$ and for every other object x': y(x'*) $\notin CO_1 \cup CO_2$,then g(x). This criterion applies e.g. in case there is only one x* for which the property *hearer*(x*) holds.

(ii) If *amk*(x*) and *ueq*(x*) $\in CO_1 \cup CO_2$, then *g*(x). This criterion applies e.g. for x=TAKE being element of the action sequence <TAKE,PUT,PROVE>, which is considered to be ***assumed mutual knowledge*** and for the hearer ***unequivocally*** interpretable.

(2) If $x \in$ IP, UHB(x^*) $\in CO_2$, then $ch(x)$ and $\neg m(x)$. This criterion applies e.g. to the elements of the unordered set of the hearer's blocks.

(3) If $x \in$ IP, id(x^*) $\in CO_2$, then :
   (i) if x is the first object in a sequence of PUT actions, then $ch(x)$, $m(x)$, $mp(x)$.

   (ii) If x is neither the first nor the last object in a sequence of PUT actions, then $ch(x)$, $m(x)$ and $\neg mp(x)$.

   (iii) If x is the last object of the sequence of PUT actions, then $\neg ch(x)$, $m(x)$ and $\neg mp(x)$.

   (iv) If x is the only object of the single PUT-action, then $\neg ch(x)$, $m(x)$, and $\neg mp(x)$.

(4) If $x \in$ IP and apl(x^*) $\in CO_2$, then $ch(x)$ and $\neg m(x)$. This criterion applies e.g. when the speaker assumes that there is a position on(b_1), among others, that can be potentially occupied by the block being moved.

The labels *ch*, $\neg$*ch* mirror the step of the problem solving while the labels *m*, *mp*, $\neg$*m*, $\neg$*mp* directly refer to the dynamics of the utterance production.

4. Definition of Topic (t), Salient Topic (st) and Focus (f)

topic rules:
(5) If $g(x) \in$ IP then t(x).
(6) If $(\neg ch,\ m,\ mp(x)) \in$ IP then t(x).

salient topic rules:
(7) If $(ch,\ m,\ mp(x)) \in$ IP then st(x).
(8) If $(ch,\ m, \neg mp(x)) \in$ IP then st(x).
(9) If $(\neg ch,\ m,\ \neg mp(x)) \in$ IP then st(x).
The rules (7) and (8) can be replaced by the equivalent rule
(7*) If $(ch, m(x)) \in$ IP then st(x).

focus rules:
(10) If (ch, $\neg$m(x)) $\in$ IP then f(x).

1.4 Examples

For lack of space I will not give a detailed specification of the context. In order to give an idea about the relation between the single arguments of the representation of the illocutionary plans and their contextual status example 1 will be presented in the following order: rule number, assigned category and contextua' information.

Example 1
The arguments of the illocutionary plan <TAKE, ADRESSEE,OBJECT> yield the following labels:

(1 ii) ---> *g*(TAKE)
amk(TAKE*) $\in CO_1$, *ueq*(TAKE*) $\in CO_1$
(1 i) ---> g(ADRESSEE)
hearer(ADRESSEE*) $\in CO_1$
(2) ---> ((*ch*, $\neg$*m*)(OBJECT))
UHB(OBJECT*) $\in CO_2$
Therefore the new IP' is <*g*(TAKE), *g*(ADRESSEE), ((*ch*, $\neg$*m*)(OBJECT))> and the application of rule (5) to the first and second argument and of rule (10) to the third argument of IP' gives
IP'' : <t(TAKE), t(ADRESSEE),f(OBJECT)>.

The surface structure of the illocutionary plan IP would be: "du nimmst einen **roten** Klotz" meaning "(you) take a red block". Bold print within the examples designate possible occurrences of accents and underlining highlights the words responsible for the cohesion of the surface form.

Example 2
Application of the rule to
IP = <PUT,ADRESSEE,OBJECT, GOAL> gives:
(1 ii) ---> *g*(PUT),
(1 i) ---> *g*(ADRESSEE),
(3 iv) ---> (($\neg$*ch*,*m*,*mp*)(OBJECT)),
(4) ---> ((*ch*, $\neg$*m*)(GOAL)).
In this case the new illocutionary plan IP' is: <*g*(PUT),*g*(ADRESSEE),(($\neg$*ch*,*m*,*mp*)(OBJECT)), ((*ch*, $\neg$*m*)(GOAL)>.
The application of rule (5) to the first and second argument, of rule (6) to the third and of rule (10) to the fourth argument of IP' gives
IP'':<t(PUT),t(ADRESSEE),t(OBJECT)),f(GOAL)>.
The surface structure would be "du stellst ihn auf den tisch" meaning "put it on the table".

Example 3
In order to illustrate the application of the salient topic rule we assume that the following utterance is made as a consequence of an illocutionary TAKE plan: "du nimmst einen roten und einen **blauen** Klotz", meaning "take a red and a blue block". Two illocutionary PUT plans would follow:
IP_1 = <PUT,ADRESSEE,$OBJECT_1$,GOAL>,
IP_2 = <PUT,ADRESSEE,$OBJECT_2$,GOAL)>.
For the first , second and fourth argument of the set IP_1 and IP_2 the same conditions as in the above PUT examples hold. For the third argument the following rules apply:
(3 i) ----> ((*ch*,*m*,*mp*)($OBJECT_1$)),
(3 iii) ----> (($\neg$*ch*,*m*, $\neg$*mp*)($OBJECT_2$)).
The new illocutionary plans are therefore:
IP1'= <*g*(PUT),*g*(ADRESSEE),((*ch*,*m*,*mp*)($OBJECT_1$)), ((*ch*, $\neg$*m*)(GOAL)>,
IP2'= <*g*(PUT),*g*(ADRESSEE), (*ch*,*m*, $\neg$*mp*($OBJECT_2$)),((*ch*, $\neg$*m*(GOAL)>.
The application of rule (5) to the first and second arguments, of rule (7) to the third argument in IP_1',of rule (9) to the third argument of IP_2' and of rule (10) to the fourth arguments of IP_1' and IP_2' yields:
IP_1''= <t(PUT),t(ADRESSEE),st(OBJECT1)),f(GOAL)>,
IP_2''= <t(PUT),t(ADRESSEE),st(OBJECT2)),f(GOAL)>.
The surface structure would be : "du stellst den roten auf den **grünen** und den **blauen** auf den **roten**" meaning "put the red on the green and the blue on the red".
If in an illocutionary TAKE plan the third argument consists of a list of many objects, then for every object OBJ *ch*(OBJ) and $\neg$*m*(OBJ) holds. This can be abbreviated by the expression (*ch*, $\neg$*m*)* of the formal language over the alphabet A. For every third argument of an illocutionary PUT plan the following holds: for the first object (*ch*,*m*,*mp*), for the objects 2 to $n-1$ (*ch*,*m*, $\neg$*mp*) and for the last object ($\neg$*ch*,*m*, $\neg$*mp*). This can be abbreviated by the expression

$(ch,m,mp),(ch,m,\neg mp)^*,(\neg ch,m,\neg mp)$.

If "du nimmst zwei/drei grüne Klötze", meaning "take three red blocks", is uttered then a cohesive succeding utterance should be "du stellst einen ..., einen .... und einen / den letzten ....", meaning "put one.., one,... and one/the last one...". In case the taken blocks were "two reds and a blue" the succeeding answers must be: "du stellst einen roten,... einen roten,.. . und den - blauen,..." meaning " put one red..., one red... and the blue..."

2. Integration in a generation system

The control of the dynamics of the conversation through the labels mentioned (*m*), mentioned in the previous utterance (*mp*) and the marking of utterance elements topic (t) and focus (f) are only two of the various subprocesses that run parallel during the main production processes. In the automatic generation of natural language, spoken as well as written, the thematic progression of a sequence of utterances and their formal cohesion must also be taken into consideration. For the spoken language prosodic cohesion must be considered additionally. Our rules for the identification of topic (t), salient topic (st) and focus (f) guarantees the coherence of the thematic progression /Daneš 1970/ of two or more successive utterances of the action sequence. Two very simple rules for thematic progression with the respective number of the examples above are now given.

R1: The only focussed OBJECT of a TAKE-action becomes the topicalized OBJECT of the following PUTä action (ex.1,2).

R2: The two/three focussed OBJECTS of the TAKEaction become the OBJECT of the following two/three PUT actions and will be labelled salient topic (Ex.3).

Our topic, salient topic and focus identification rules also allows to make predictions about the distribution of accents. Indeed an accent will be assigned to the elements labelled salient topic (st) and focus (f); the topic elements (t) get no accents. In this phase of the work accents are assigned to all arguments of the proposition. The assignment of the accent to the adjective instead of the noun in phrases like "...den roten..." involves application of the same criteria inside lower level constituents. In order to generate cohesive surface structures it is also necessary to know when to use a definite article within noun phrases (the last one : der letzte) or an indefinite pronoun (one : ein). This choice depends on the pragmatic decision of taking one or more blocks and on the properties shared by the objects in question. Under the assumption that only the parallel processing semantic and pragmatic information allows the choice of appropriate lexical material. For this purpose we will extend our set of semantic categories to express if a certain objects is an underdetermined or a determined element of a set. /For an extended discussion see Pignataro 1987 and Pignataro (forthcoming)/.

3. The Generation Model

The generation model consists of four functions: F_1, F_2, F_3 and F_4. F_1 maps a illocutionary plan IP and the context C onto an illocutionary plan IP' with additional semanticä pragmatic categories. F_2 maps IP' onto IP'': i.e. semanticpragmatic categories onto topic, salient topic and focus. F_3 maps IP'' onto surface sentences. F_4 maps C and IP onto the changed context C'.

$$\begin{tikzcd} (IP,C) \arrow[r, dashed, "F1"] \arrow[d, dashed, "F4"] & (IP') \arrow[r, dashed, "F2"] & (IP'') \arrow[r, dashed, "F3"] & Surf.Str \\ (C') \end{tikzcd}$$

Literature

Daneš, F. (1970) Zur linguistischen Analyse der Textstruktur, in Folia Linguistica IV , pp.72−78

Hajicǒvá,E.,Vrbová (1982) "On the role of the Hierarchy of Activation in the process of Natural Language Understanding" in Horecky J.(ed) Coling '82 Prague.

Hajicǒvá,E.,Sgall,P.(1985) "Towards an automatic Identification of Topic and Focus"; Proceedings of the second Conference of the European Chapter of the Association for Computational Linguistics. Geneva pp.263−267.

Pignataro, V.(1987) Topik und Fokus in der Sprachproduktion; KoLiBri Nr. 8, Universität Bielefeld.

Pignataro, V. (forthcoming)

Schiffers,S.R.,(1972) Meaning;Oxford, at the Clarendon Press.

Sgall,P.,Hajicǒvá,E.,Panevǒvá,J.(1986) The Meaning of the sentence and its Semantic and Pragmatic Aspects; Dordrecht, Reidel P.C.