ON THE SYNTAX-DISCOURSE INTERFACE IN HUNGARIAN

Anna Gazdik[1,2]   and   András Komlósy[2]
Université Paris 7   Research Institute for Linguistics
LLF[1]   Hungarian Academy of Sciences[2]

Abstract

In this paper, we investigate the syntax-discourse interface in Hungarian and propose the first steps of its formalization in LFG. Concerning the syntax, we propose that the c-structure is flat in Hungarian, and the preverbal domain in it is governed by the information structure. After examining the distribution of elements in the topic field and the Prominent Preverbal Position (PPP), we point out that they can be filled differently in *neutral* and *non-neutral* sentences, i.e. depending on the discourse. The PPP in non-neutral sentences can be occupied by an element/constituent referred to as the *hocus*, which cannot easily be accommodated in the i-structure assumed in the mainstream LFG literature. This is why a new architecture is put forth, building on the common features and not the discourse functions themselves, that can account for the Hungarian data more adequately. The syntax-discourse interface is then exposed via the mappings between the c- and i-structures.

1   Introduction

In configurational languages, like English, syntactic structure and the linear order of constituents are determined by syntactic functions, like *subject* or *object*, which constitute functions between constituents and the whole sentence (*"the subject/object of the sentence"*). In discourse-configurational languages (É. Kiss, 1995), syntactic structure and the positions of the elements reflect discourse structure, i.e. the role that the sentence plays in the discourse. Discourse functions (*topic, focus,* etc.) are thus not functions between a constituent and the sentence, but between a constituent and the discourse structure.

Many syntactic analyses, especially in derivational frameworks, account for the discourse-relatedness of syntactic structures by positing special functional projections (TopP, FocP) that host a particular discourse function (see for instance Rizzi (1997), É. Kiss (2002)). However, such analyses run into some serious problems. Firstly, positing separate functional projections for every discourse function has little explanatory adequacy. Secondly, as opposed to lexical projections (NP, PP, AP, etc.), discourse functional projections do not encode categories, but discourse-semantic information integrated into the syntax, without a clear formal account of the discourse/syntax interface. Thirdly, as we will show, discourse functions cannot exclusively be assigned to designated syntactic positions, and *vice versa*, a particular syntactic position can host more than one discourse function, even in discourse-configurational languages.

In this paper, we deal with Hungarian, more specifically with the preverbal part/left periphery of the sentence. Schematically, the Hungarian sentence can be divided into two fields: the *topic* and the *comment*, and the comment can be further divided into four subfields: the *pre-comment*, the *prominent preverbal posi-*

tion (henceforth PPP), the *finite verb*, and the *postverbal part* (based on Kálmán (2001)). This is illustrated in Figure (1):

Topic field   Comment

$$\overbrace{Quantifiers \quad PPP \quad \underbrace{V}}^{} \quad Postverbal \quad part}$$
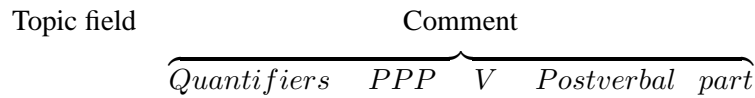
Figure 1: The schematic representation of the Hungarian sentence

Although the names (topic, comment, prominent preverbal position, etc.) are of semantic/pragmatic in nature, there are also syntactic (distributional) and prosodic arguments for this division of the Hungarian sentence into these fields and sub-fields. The topic field and the comment can be divided prosodically: the elements in the topic field exhibit a rising intonation pattern, constituting a preparatory section to the first obligatory stress, which falls on the first (content) word of the comment, starting a falling intonation pattern. This main stress can be followed by other main stresses in a sequence of downsteps (in a level-prosody sentence). The verb can lose its stress, cliticizing onto the element in the prominent preverbal position (verbal modifier, negative word, etc.). The first obligatory stress can also be followed by a so called *eradicating stress* (a sharp falling pitch accent) on the element in the prominent preverbal position, which can only be followed by another eradicating stress (on the right periphery), otherwise the rest of the sentence is deaccented, or all other main stresses are reduced. An eradicating stress can also fall onto an element in the topic field, which is often followed by another one in the PPP. The distribution of elements in these positions and fields is further characterized in the next section, concentrating mainly on the topic field and on the PPP.

2 The Data

2.1 Distribution in the topic field

The topic field is usually reserved for definite or specific indefinite noun phrases and referential (time and place) adverbials (individualizable elements), whose order is free in the topic field (É. Kiss, 2005). However, the rightmost position of certain sentence adverbials, like *tegnap* (yesterday), *idén* (this year), indicates the right frontier of the topic field itself as well. These adverbials are interpreted as sentence adverbials in the topic field (1), but as referring only to the immediately following constituent in the comment (2).[1]

---

[1] Main stress is indicated with ('), eradicating stress with (") and a rising-falling intonation pattern with (/) throughout this paper.

(1)   A ’vonaton   ’tegnap sok   ’gyerek ’utazott.
        the train.SUPERESS yesterday a lot of child   travel.PST
        Yesterday, there were a lot of children travelling on the train.

(2)   A vonaton   ”tegnap utazott sok gyerek.
        the train.SUPERESS yestarday a lot of child travel.PST
        It was yesterday that a lot of children were travelling on the train.

Example (1) conveys information about an event, answering a question like *What happened?*, whereas (2) contrasts *tegnap* (yesterday) with another day (or other days), when supposedly only a few children were travelling on the train.

The set of elements that can appear in the topic field with a special rising-falling intonation pattern is larger: infinitives, adjectives, bare nouns, quantifiers, verbal modifiers, and adverbs (other than the ones mentioned above) can appear there if pronounced that way. This is the case of *Mari* in the following example:

(3)   a.   Q: -Mit hoztak   a vendégek a   bulira?
                 what bring.PST the guests   the party.SUBL
                 What did the guests bring to the party?

        b.   A: -/**Mari** ”csokitortát   hozott.
                 Mari   chocolate cake bring.PST
                 As for Mary, she brought a chocolate cake.

2.2   Distribution in the PPP

The prominent preverbal position (PPP), which is between the precomment and the finite verb, can also be occupied by a wide range of elements. Some of them appear in the PPP in level-prosody sentences and can receive an eradicating stress *in situ*. However, they must follow the verb if there is another element that carries the eradicating stress. The explanation is that only one of them can precede the verb, when there is more than one potential element that can occupy the PPP in a sentence. The others occupy postverbal positions (except for some questions in which there is also a focused constituent).

2.2.1   Verbal Modifiers (VM)

The verbal modifiers we mention here are *verbal particles, bare nominal complements* and *secondary predicates*. **Verbal particles** can have an adverbial or a lexicalized aspectual meaning. When there is no other potential element appearing in the PPP, they precede the verb, otherwise they follow it[2]:

---

[2]In what follows, verbal particles will be referred to as verbal modifiers (VM). Verbal particles are written as one word with the noun when they precede it, but as two words when they follow it.

(4)   'János '**ki**olvasta   a   'könyvet.
          János VM.read.PST the book.ACC
          John finished the book.

(5)   'János "egy hét   alatt olvasta **ki** a   könyvet.
          János one   week under read.PST out the book.ACC
          John finished the book in one week.

About a classification and analysis of verbal particles, see Surányi (2009) and Laczkó and Rákosi (2011).

Another type of verbal modifier is **bare nominal complements** (see also de Swart and Farkas (2003)), illustrated by the following example:

(6)   János '**levelet**   ír.
          János letter.ACC writes
          John is letter-writing.

Secondary predicates co-occur with some (other) argument of the verb, about which they state something. They often express a goal or a result (7), and appear in the immediately preverbal position:

(7)   János '**pirosra** festette   a 'kerítést.
          János red.SUBL paint.PST the fence.ACC
          John has painted the fence red.

2.2.2   The Hocus

The hocus (introduced by Kálmán (1985a,b); Kálmán et al. (1986), and also referred to in Kálmán (2001)) is a noun phrase, a negative adverb or a monotone decreasing quantifier (as opposed to monotone increasing ones that appear in the precomment) expressing some participant or circumstance in the event denoted by the predicate. Such elements/phrases can bear main stress and appear in the immediately preverbal position, when the event denoted by the verb is not particularly newsworthy, or it is a regular event, apart from the circumstance or participant denoted by the hocus, which expresses something unusual or unexpected. In these cases the main proposition of the sentence is the identification of this participant or circumstance.

(8)   János 'tegnap   '**vonattal** 'utazott   'haza. (NP)
          János yesterday by train   travel.PST home
          Yesterday John took the train to go home.

(9)   'Ma a **feleségem**   'vitte   az 'óvodába   a 'gyerekeket.
           today the wife.POSS.1SG take.PST the kindergarten.ILL the children.ACC
           (NP)

           Today my wife took the children to the kindergarten.

(10)   '**Kevesen** 'jöttek   el   a   'bulira.   (monotone decreasing
           few   come.PST VM the party.SUBL quantifier)
           Only a few people came to the party.

(11)   'János '**ritkán** 'megy el   'kirándulni. (negative adverb)
           János   seldom goes   VM to hike
           John seldom goes hiking.

Example (8) implies that John usually does not take the train, according to (9) it is usually not the wife, but someone else that takes the children to the kindergarten, in (10) more people were expected to come to the party, and in (13) John goes hiking less often that it would be expected.

     In identificational sentences, the subject appears as the hocus, preceding the verb (copula):

(12)   '**János** volt az 'igazgató.
           János   was the director.
           John was the director.

(13)   A   'nyomozó '**a sógorom**   volt.
           the inspector   the brother-in-law.POSS.1SG was
           My brother-in-law was the inspector.

     In the mainstream linguistic literature on Hungarian, sentences containing a hocus are not discussed, and they are not clearly distinguished from narrow-focus sentences. This is a problem, since *hocus* and *focus* are clearly different (see below).

2.2.3   The Focus

The focused constituent differs from the above mentioned elements/phrases in that it bears sharp falling pitch accent, functioning as *eradicating stress*, referring to the fact that no main stress (only another eradicating stress) can follow it in the rest of the sentence. In Hungarian, the main function of focus is contrast, i.e. it identifies the entities about which the predicate holds and restricts the validity of the predicate to only these entities by excluding the other members of the relevant

set. Sentences with focus cannot be uttered *out of the blue*. In most cases, they are answers to questions (14), reactions or corrections (15) (capitals indicate the constituent carrying a pitch accent):

(14)   Answer:

   a.   Q: -Ki   hívta   meg Marit   a   bulira?
            who invite.PST VM Mari.ACC the party.SUBL
         Who invited Mary to the party?

   b.   A: -"**Zoli** hívta   meg (Marit   a   bulira).
            ZOLI invite.PST VM (Mari.ACC the party.SUBL)
         It was ZOLI who invited her (to the party).

(15)   Correction:

   a.   -'Mari 'tegnap   'kiolvasta   a   'Háború és békét.
         Mari   yesterday VM.read.PST the War   and Peace.ACC
         Mary finished yesterday *War and Peace*.

   b.   -Nem, a   "**Bűn és bűnhődést**   olvasta   ki.
         no,   the Crime and Punishment.ACC read.PST VM
         No, she finished *Crime and Punishment* yesterday.

We should note here that not all foci appear in the PPP. If the focus is a universal quantifier (16), it cannot occupy the immediately preverbal position. Similarly, in the presence of a focus in the PPP, a second one must appear on the right periphery of the sentence (17). Thus the syntactic position cannot help in the identification of all foci in Hungarian.

(16)   /A csillagok háborúját "**mindenki** megnézte.
        Star wars.ACC   everyone   VMsaw
        Star Wars was seen by everyone.

(17)   A "**lányok** nyerték meg tegnap   a "**kajakversenyt**), a "**fiúk** pedig
        the girls   won   VM yesterday the kayak contest,   the boys and
        a "**kenuversenyt**.
        the canoe contest
        It was the girls who won the kayak contest yesterday, and the boys who won the canoe contest.

The following examples (based on Kálmán (2001)) illustrate the difference between the hocus and the focus:

(18)   'Ezen   a héten   a 'Mecsekben raboltak   ki
        this.SUPERESS the week.SUPERESS the Mecsek.INESS rob.PST.3PL VM
        egy 'pénzszállító   autót.
        a   money transport car.ACC
        This week it was in the Mecsek (mountains) that a money transport vehicle was robbed.

(19)   'Ezen   a héten   a MECSEKBEN raboltak   ki
        this.SUPERESS the week.SUPERESS the Mecsek.INESS rob.PST.3PL VM
        egy pénzszállító   autót.
        a   money transport car.ACC
        This week it was in the Mecsek (mountains) that a money transport vehicle was robbed.

A Mecsekben is hocus in (18) and focus in (19). The difference between the two sentences can be illustrated with the different contexts. In the first case, robbing a money transport car counts as a usual event. The hocus identifies the place where the event happened this week. The location counts as non-canonical, unusual and surprising at the same time, either because this happens less often in mountains, or because the Mecsek is not known for such crimes. In (19), robbing a money transport car is not necessarily a usual event. The focus identifies the place where it happened, contrasting it to other locations, where it could have potentially happened, or correcting a previously proposed other location. In this latter case, the sentence does not form a true prosodic minimal pair with (18), but would have the following form:

(20)   (Nem!) A "Mecsekben rabolták   ki a pénzszállító   autót
        (No!)   the Mecsek.INESS rob.PST.3PL VM the money transport car.ACC
        (és nem a   Bakonyban).
        (and not the Bakony.INESS)
        No! It was in the Mecsek that the money transport vehicle was robbed (and not in the Bakony)!

(20) is about a single event (indicated by the definite article in front of the noun *pénzszállító autó* and the definite conjugation of the verb). It identifies the location of the event, by contrasting it to another location.

    We can thus conclude that both the hocus and the focus are identificational elements, appearing in different discourse contexts. The focus is prosodically distinguished, carrying a pitch accent (followed by the deaccenting or reduced stress of the post-focal material), whereas the hocus is not more prominent prosodically than the other lexical elements of the sentence (except for the verb which cliticizes onto it). In addition, the focused constituent presupposes that the proposition can-

not be true simultaneously with another, in which the focused element is changed to an alternative to its denotation (the robbery cannot take place at two locations at the same time). To illustrate this, consider the possible continuations of (18) and (19):

(21) a. 'Ezen   a héten   a 'Mecsekben raboltak
          this.SUPERESS the week.SUPERESS the Mecsek.INESS rob.PST.3PL
          ki egy 'pénzszállító   autót.
          VM a   money transport car.ACC
          This week it was in the Mecsek (mountains) that a money transport vehicle was robbed.

      b. Nem, nem csak ott. A "Bakonyban is kiraboltak
          no   not only there the Bakony.INESS too VM.rob.PST.3PL
          egyet.
          one.ACC
          No, not only there. One was robbed in the Bakony too.

(22) a. Ezen   a héten "a Mecsekben   raboltak   ki egy
          this.SUPERESS the week the Mecsek.INESS rob.PST.3PL VM a
          pénzszállító   autót.
          money transport car
          This week it was in the Mecsek that a money transport vehicle was robbed.

      b. #Nem, nem csak ott. A "Bakonyban is kiraboltak egyet.

2.2.4 Question words

Finally, question words typically appear immediately in front of the finite verb as well. In the presence of a question word, not only verbal modifiers (23), secondary predicates, etc. but elements of the precomment (24) also occupy postverbal positions:

(23) **Kit**   hívott   meg János a bulira?
      who.ACC invite.PST VM János the party.SUBL
      Who did John invite to the party?

(24) **Kire**   szavazott mindenki?
      who.SUBL vote.PST everybody
      Who did everybody vote for?

Question words are often argued to constitute a subclass of focus, based on similarities in prosody, syntactic position, semantics and, in some languages, morphology. Despite the apparent similarities, it would be too hasty a generalization to collapse question words into foci in Hungarian. Although foci and question words seem to share the same syntactic position and prosody, some important differences suggest that they belong to different types of objects. For instance, while strictly only one preverbal focus is permitted in Hungarian (if there is a second, it is obligatorily postverbal), two question words can appear preverbally, and they can even be preceded by a focused constituent or followed by a negated focus in some contexts. Furthermore, two question words can be coordinated, irrespective of their grammatical function, whereas this is not possible with two focused constituents. We will argue that their similarities can also be derived from the fact that they play parallel roles in the discourse they occur in: both question words and foci presuppose the rest of the sentence, and foci in the answers correspond to question words in the question.

The elements in the PPP in (4)-(13) are in complementary distribution with each other, i.e. a verb cannot appear simultaneously with a verbal modifier and a secondary predicate, for instance, even if one of those followed the verb. They can all receive an eradicating stress *in situ*, in the PPP. However, in the presence of the elements in (14)-(15), and (23)-(24) they have to follow the verb.

2.3 The role of discourse structure

Considering the diversity of elements that can occupy the topic field and the PPP, how could we identify what is in common in them? Concerning the PPP, as we have already seen in the case of secondary predicates, these elements contribute to the meaning of the sentence with a secondary/independent proposition that can sometimes modify the proposition formulated by the comment. According to É. Kiss (2006), not only verbal modifiers can be considered as resultative, locative or terminative secondary predicates, but structural focus can be reanalyzed as a *specificational predicate* (similarly to English cleft sentences) as well. Komlósy (1994) also showed that preverbal bare nominals function as predicates that predicate of an existentially bound variable incorporated into the verb. We will argue that apart from the common grammatical function (secondary predicates), the common properties of some of the elements in the PPP are related to the information structure and the discourse the sentence is uttered in.

To see this last point more clearly, an important remark is due here. Some of the above mentioned elements can never appear in the same sentence, since the discourse types they are part of are different. In Hungarian, based on formal, interpretational and discourse factors, two types of sentences can be distinguished: *neutral* ("all-focus")[3] and *non-neutral* "narrow-focus") sentences (see Kálmán (1985a,b)). Formally, non-neutral sentences contain an eradicating stress (25) in the PPP (and

---

[3]Neutral sentences can answer questions like *What happened?*. They cannot always be referred to as all-focus sentences, since they can contain topics:

possibly also in the topic field), whereas neutral sentences have level-prosody and can contain several main stresses (26):

(25)   'Tegnap ''**Marit** láttam   a városban, (nem ''Jánost).
           yesterday Mari.ACC see.PST.1SG the city.INESS (not János)
           Yesterday I saw MARY in the city, not JOHN.

(26)   = (1)

           'János 'tegnap   'vonattal   utazott   'haza.
           János   yesterday train.INSTR travel.PST home

           Yesterday John took the train to go home.

In (25), the focus (*MARI*) bears a sharp falling pitch accent, after which the rest of the sentence is deaccented (or bears reduced stress), except for the second focus (*JÁNOST*). In (26), the topics (*János*, *tegnap*) have a slightly rising tone, pointing forward to the comment, whereas the falling tone starts on the first constituent of the comment, the hocus (*vonattal*).

Concerning the topic field, the elements occurring there have different properties in narrow-focus and all-focus sentences. It is common in the two cases that they introduce subtopics/subquestions. In a neutral context, there is no topic in the sentence if the sentence continues the previous subtopic. However, when a sentence changes the subtopic, the element in the linearly first position indicates the topic shift. This is why this type of topic is often called *thematic shifter*. The other type of topic that we find in narrow-focus sentences (indicated prosodically with eradicating stress and a rising tone) is closely related to the contrastive property of these sentences and is called *contrastive topic* in the literature. The contrastive topic restricts the domain of the validity of the focused constituent to some element of a set, implying that to other elements of the relevant set the focused constituent does not hold. For instance in (3-b), repeated here as (27-b), the contrastive topic (Mari), indicates *the strategy* of decomposing the set of guests into its elements, the individual guests and associates, each of them with an answer (i.e. a focused constituent). This association means at the same time that as opposed to Mary, there is at least someone else who did not bring a chocolate cake.

(27)   a.   Q: -Mit hoztak   a vendégek a bulira?
                what bring.PST the guests   the party.SUBL
                What did the guests bring to the party?

(i)   'János 'találkozott 'Marival   a   'városban.
        János meet.PST   Mari.INSTR the city.INESS
        John met Mary in the city.

In this example *János* is the topic, and the sentence is not *all-focus*.

     b.   A: -/**Mari** "csokitortát   hozott.
            Mari   chocolate cake bring.PST
            As for Mary, she brought a chocolate cake.

The two types of sentences are used in different contexts. Neutral sentences, present mostly in narrative contexts, only convey information and continue the narrative, whereas non-neutral sentences are used for asking questions, answering questions, corrections and confirmations, disagreement, and highlighting parallels. If we analyze the discourse as the hierarchy of topics and subtopics (or, a question under discussion, subquestions and the possible answers; see for instance Büring (1997, 2003)), we see that both sentence types contain two prominent preverbal parts (the topic field and the PPP), and a set of (in the sense of Jacobs (1984)) prominent element types that can fill these positions. Elements in the topic field relate to the discourse in a way that they thematize it by selecting the subtopic/subquestion with respect to which the given sentence adds new information to the common ground. On the other hand, elements appearing in the PPP (or possibly in the precomment) constitute the most informative, prominent part of the sentence. In some cases, this can be new information, or the part that answers a question, or the unexpected or unusual part of the meaning (as we have seen in the case of the *hocus*). In order to formalize the above sketched syntax-discourse interface, first we take a look at the syntactic structure of Hungarian in the next section.

3   The syntactic structure of Hungarian

The LFG representation of the syntactic structure (c-structure) of Hungarian can differ from that of other theories, since in LFG, constituent structure corresponds to a flexible X-bar theory representation, in which no node, not even the head is obligatory, and exocentric constituents are permitted (there is no binary-branching constraint). The question is, what kind of c-structure should be associated with Hungarian. To our knowledge, there have been two proposals in the LFG literature for the c-structure in Hungarian, but they concentrated mostly on the problem of the preverbal position and the elements it can host: focus and question words.

    In the first analysis (Börjars et al., 1999), the immediately preverbal constituent is sister to the verb in an extended verbal projection, which is supposed to host also all the elements of the preverbal domain (topics and quantifiers). The discourse functions are associated with syntactic positions via functional annotations. This analysis does away with the set of functional projections (TopP, CTopP, DistP/QP), whose head position is usually empty, since they are only postulated for accommodating one type of element in their specifier position. FocP is an exception to this, since the verb is supposed to move into its head position, leaving behind the verbal modifier. However, according to Börjars et al. (1999), even a FocP is superfluous

in a theory in which no Foc feature is supposed to be assigned or checked. The authors assume OT-type constraints as well, which account for word-order and the immediately preverbal position of the focus. The second analysis to be mentioned here is that of Mycock (2006), who assumes that the focus and the question words are in Spec,VP, thus obligatorily sister to the verb.[4]

According to Dalrymple (2001), functional categories vary from language to language, and each of them has to be motivated for each language. According to this, the I head position can be occupied by a finite verb or an auxiliary, like the C position (in inversion contexts). Thus King (1993) assumes that in Russian, only non-finite verbs reside in the VP, finite verbs occupy the I position, the topic and the contrastive focus the Spec,IP and interrogative words the Spec,CP position. Dalrymple (2001) also mentions that positing a VP projection is motivated only if it contains only the verb and its complements (except for the subject) and these constituents can appear together at other parts of the sentence as well. On the other hand, if the subject can appear as sister to the V, the VP projection is unmotivated. Now, the syntactic structure of non-configurational languages is represented with the help of the non-configurational S node, which does not necessarily contain a CP or an IP projection. It is also possible that one part of the sentence is hierarchical and the other exhibits a free word order, flat structure, in which case the tree diagram contains both CP/IP and S nodes. Such languages are Warlpiri and Welsh.

These considerations about the VP undermine Mycock (2006)'s (and Laczkó and Rákosi (2011)'s) c-structure, since in Hungarian, the subject can be postverbal, appearing as sister to the verb, between the verb and the direct object:

(28) Marinak adta   oda János a könyvet.
      Mari.DAT give.PST VM János the book.ACC
      John gave the book to MARY.

In Hungarian, as we have seen, the preverbal and postverbal parts of the sentence differ, in that in the preverbal section, the position and the order of the elements depend on their role in the information structure. This can be directly represented in LFG via the functional annotations. The question is now, whether a hierarchical preverbal section is motivated even in the LFG framework. In the transformational frameworks, two factors motivated the hierarchical preverbal structure: the obligatory binary branching in the tree diagrams and the fact that the linear order of the elements determines their relative scope as well. As opposed to this, the postverbal part of the sentence exhibits free word-order (obeying, supposedly, certain phonological factors, such as heavy elements tend to follow lighter ones). In LFG, neither of these factors necessitate a hierarchical architecture, since the linear order of elements can in itself reflect the scopal relations, thus there is no reason for positing a hierarchical sentence structure in Hungarian. As was pointed

---

[4]Laczkó and Rákosi (2011) also assume a VP projection in Hungarian, in which the verbal modifiers occupy the specifier position.

out above, a VP projection is not motivated. In the c-structure, annotations indicate the grammatical and discourse functions. In addition, they can express other constraints as well, for instance, the obligatory presence of other elements, which is needed in the case of sentence-final focus (which can appear on condition if there is a preverbal focus in the sentence). With the annotations, thus, we can express and formalize the observation that the preverbal part of the Hungarian sentence is determined by the information structure. The flat, c-structure with preliminary annotations is shown in Figure (2).

```
                              S
          ┌──────┬──────┬─────┼────────┐
        XP*      XP      XP   V/AUX     XP*
      (↑GF)=↓  (↑GF)=↓  (↑GF)=↓  ↑=↓   (↑GF)=↓

      topics  universal quantifiers  PPP   verb   postverbal material
```
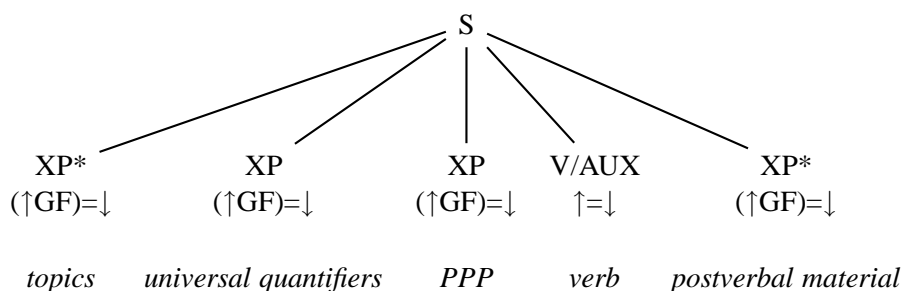
Figure 2: C-structure of Hungarian

The annotations in Figure (2) refer to the grammatical function represented in the f-structure. In order to formalize the interface of the syntactic structure with the information structure, we will consider the latter in the next section.

4   Information Structure

4.1   Information Structure in LFG

The information structure as a separate level of representation in LFG was proposed by Butt and King (1996) and King (1997). In earlier versions of the LFG framework, discourse functions were integrated in the functional structure and one syntactic unit was associated with two functions at the same time, for instance *topic* and *subject*). The proposal was motivated by the problem of focusing the predicate without its arguments and the observation that syntactic constituents do not correspond systematically to constituents of information structure (see King (1997)).

Butt and King (1996) proposed that the information structure consist of four sets, which are defined by the combination of two features: *new +/-* and *prominent +/-*. The *TOPIC* set contains elements that are prominent, but not new, the *(Information) FOCUS* set contains new and prominent elements, whereas old and not prominent elements belong to *BACKGROUND* and new but not prominent ones to *COMPLETIVE INFORMATION* (see Figure 3).

As a recent development, Dalrymple and Nikolaeva (2011) propose to encode semantic information in the information structure. In this representation, which

	Topic	Focus	Background Information	Completive Information
New	−	+	−	+
Prominent	+	+	−	−

Figure 3: I-Structure units (Butt and King, 1996)

we also follow in this paper, the values of the discourse function attributes are the meaning constructors of the elements that appear in a given set.

Although this classification simplifies the definition of discourse functions, for instance, foci are not always new, and prominence also has to be defined, we will propose an analysis which builds on this architecture, but it captures the facts presented above more adequately.

4.2 Problems and new proposal

The architecture of the information structure, as proposed by Butt and King (1996), King (1997) and Choi (1999), contains *topic* and *focus* as i-structure primitives. In this section, an alternative architecture is proposed, which is not fundamentally different from the one presented above, but it could be argued to capture the problematic facts better. The main problem concerns the fact that the set of elements with different discourse, semantic and prosodic properties is larger than the above architecture could accommodate without simplifying these properties. Let us now go through these elements, familiar from previous sections:

Thematic shifters can be defined as the element that links the sentence to the discourse by introducing a new subtopic of the discourse topic. It can be observed that in Hungarian (and, according to Vallduví (1992) in Catalan), a thematic shifter is present in the sentence only if it does not continue the previous subtopic. Such sentence topics are typical in narrative contexts. From this it follows that there are a number of sentence types which do not contain sentence topics, for instance the ones that continue the previous subtopic in narrative contexts, some questions, and answers introducing complex strategies.

The focus is the semantically/pragmatically prominent part of answers to questions, corrections, contrastive and parallel structures, which is usually formally highlighted as well (pitch accent, syntactic position).

Contrastive topics are similar to foci, in that they do not appear in *non-neutral* utterances. They usually co-occur with a focused constituent, which both Büring (2003) and Gyuris (2009) explain by claiming that contrastive topics carry the presupposition that there is a focus value (different from and not entailed by that of the sentence) associated with an alternative to the denotation of the contrastive topic. Contrastive topics can co-occur with sentence topics, in this case the sentence is linked both to a more general discourse topic and to a more restricted one:

(29) [$_T$ János] [$_{CT}$ a levest] $_F$ megette(, de a [$_{CT}$ húst] [$_F$
      John   the soup.ACC VM.eat.PST but the   meat.ACC
      nem]).
      not
      As for the soup, John did eat it (, but he did not eat the meat).

<div align="right">Gyuris (2002, p. 23, 15)</div>

In (29), the thematic shifter is *János*. The sentence contains a contrastive topic (*a levest*), which is implicitly or explicitly contrasted to *a húst*. In the two parallel clauses, the focus values are also different, since different contrastive topic values have to be mapped on different focus values (Gyuris, 2009). The different focus values are verum and falsum foci, respectively.

The hocus is an argument or adjunct appearing in the preverbal position in neutral (or all-focus) sentences in Hungarian. It lacks the pitch accent and the contrastive-exclusive reading of focused constituents in non-neutral sentences.

It follows from the facts presented above that the hocus cannot be analyzed as a subtype of focus, and thus it would be difficult to integrate it into Butt and King (1996)'s model of information structure.

Question words are often assumed to be a subtype of focus and analyzed as such at the level of information structure. However, as King (1993) remarks,

"[t]he discourse functions associated with questions are not fully understood. The term Q(uestion)-Foc(us) is used to indicate the role which corresponds to the focus in the answer to the question."

Mycock (2006) also distinguishes between *interrogative* and *non-interrogative* foci in order to account for Hungarian multiple questions in which two (or more) question words can precede the verb, unlike the case of multiple foci, where only one of them can appear in the PPP. We assume (see also Gazdik (2011, Chapter 7) for the details) that question words cannot be collapsed into focus, but their common properties can be captured if we assume that they have a parallel status in the discourse.

An alternative solution would be to emphasize the common properties of the different discourse functions and to build the i-structure architecture on them. Thus a set would include elements based on a common property, without claiming that these elements must be semantically and discourse-wise identical. The exact semantic and discourse properties would follow, as mentioned above, from the meaning constructors of the individual elements and the discourse structure the sentence appears in. In what follows, our proposed i-structure architecture is presented. It keeps some aspects of Choi (1997)'s features, but also deviates from it in others. First of all, we have seen that certain elements are semantically prominent and formally (syntactically or prosodically) highlighted. These elements are referred to as +PROMINENT, and others as −PROMINENT. Semantic prominence can be defined, based on Jacobs (1984), with respect to the illocutionary operator associated with the sentence. Prominent elements are the ones specially affected by the illocu-

tionary operator. These elements are different in reactive (focus, contrastive topic) and out of the blue sentences (thematic shifter, hocus, certain question words), but constitute the prominent set at i-structure. This distinction defines two sets in the i-structure. Furthermore, we saw that among prominent elements we find some that link the sentence to the discourse (by introducing a subtopic of the discourse topic or reshaping the discourse topic), and others which do not. The first set is called D-LINKED (see Pesetsky (1987); Comorovski (1996), etc.), and the second ¬D-LINKED. This way, we stay neutral with respect to the *new* status of focus, since the focus does not necessarily constitute new information (in the sense of introducing a new discourse referent). In the −PROMINENT set, we also find a D-LINKED and a ¬D-LINKED subset, the first corresponding to background, the second to completive information (these are kept from Butt and King (1996)). The proposed architecture hosts the above mentioned elements as shown in Figure (4).[5] Assuming this structure, the annotated c-structure of Hungarian would look like that in Figure (5).
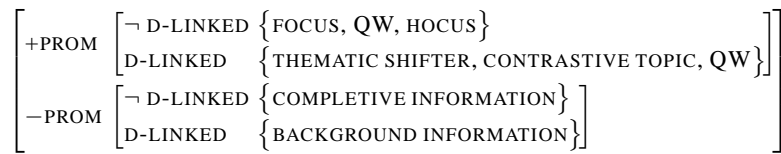
$$
\begin{bmatrix}
+\text{PROM} & \begin{bmatrix} \neg \text{D-LINKED} & \{\text{FOCUS, QW, HOCUS}\} \\ \text{D-LINKED} & \{\text{THEMATIC SHIFTER, CONTRASTIVE TOPIC, QW}\} \end{bmatrix} \\
-\text{PROM} & \begin{bmatrix} \neg \text{D-LINKED} & \{\text{COMPLETIVE INFORMATION}\} \\ \text{D-LINKED} & \{\text{BACKGROUND INFORMATION}\} \end{bmatrix}
\end{bmatrix}
$$

Figure 4: Proposed i-structure

---

[5]QW stands for *Question Word*, without specifying its discourse function as topic or focus. Arguably, some question words in multiple questions show similar properties with topics (Gazdik, 2010).
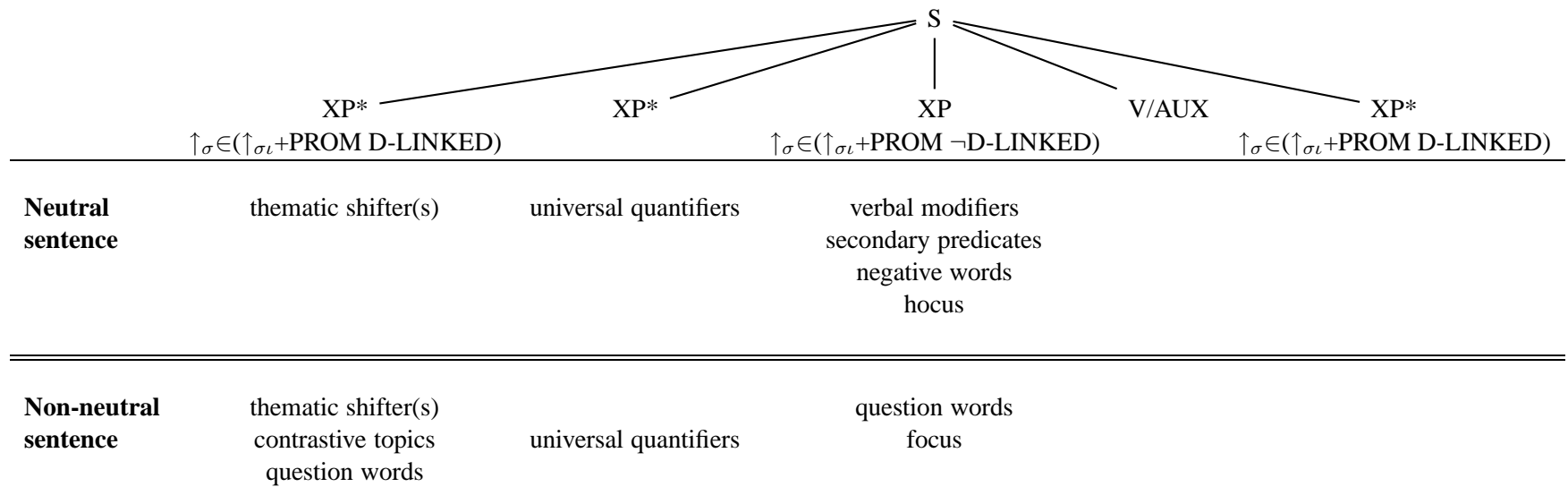
$$S$$

XP*	XP*	XP	V/AUX	XP*
$\uparrow_\sigma \in (\uparrow_{\sigma\iota}+\text{PROM D-LINKED})$		$\uparrow_\sigma \in (\uparrow_{\sigma\iota}+\text{PROM } \neg\text{D-LINKED})$		$\uparrow_\sigma \in (\uparrow_{\sigma\iota}+\text{PROM D-LINKED})$

	XP*	XP*	XP	V/AUX	XP*
Neutral sentence	thematic shifter(s)	universal quantifiers	verbal modifiers secondary predicates negative words hocus		
Non-neutral sentence	thematic shifter(s) contrastive topics question words	universal quantifiers	question words focus		

Figure 5: Annotated C-structure

5 Conclusion

In this paper, we presented the first steps of the possible formalization of the syntax-discourse interface in Hungarian in the framework of LFG. The discourse-configurational property of the language is reflected by the assumption that the preverbal part of the sentence is determined by the information structure via a direct mapping between this latter and the c-structure. We argued, furthermore, that the traditionally assumed i-structure, containing *topic* and *focus* as its basic sets cannot account for all the Hungarian data, mainly since it could not accommodate the *hocus* (the constituent in the PPP of neutral sentences), and it usually collapses question words into foci. The alternative architecture proposed is based on two features: $+/-$ PROMINENT and D-LINKED/¬D-LINKED, grouping together the elements with discourse functions that share some common properties. Further research should be conducted in order to specify how discourse structure could be included in the representations (in which the vague terms of *neutral* and *non-neutral* sentence would make sense), and also in order to account for the exact role of verbal modifiers and secondary predicates in the system.

References

Kersti Börjars, John Payne, and Erika Chisarik. On the justification for functional categories in LFG. In Miriam Butt and Tracy Holloway King, editors, *Proceedings of the LFG99 Conference*, Manchester, 1999. CSLI.

Daniel Büring. *The Meaning of Topic and Focus — The 59th Street Bridge Accent*. Routledge, London, 1997.

Daniel Büring. On d-trees, beans, and b-accents. *Linguistics & Philosophy*, 26(5): 511–545, 2003.

Miriam Butt and Tracy Holloway King. Structural topic and focus without movement. In Miriam Butt and Tracy Holloway King, editors, *Proceedings of the LFG96 Conference*, Rank Xerox, Grenoble, 1996. CSLI.

Hye-Won Choi. Information structure, phrase structure, and their interface. In Miriam Butt and Tracy Holloway King, editors, *Proceedings of the LFG97 Conference*. CSLI Publications, 1997.

Hye-Won Choi. *Optimizing Structure in Context: Scrambling and Information Structure.* Dissertations in Linguistics. CSLI Publications, Stanford, CA, 1999.

Ileana Comorovski. *Interrogative Phrases and the Syntax-Semantics Interface*. Kluwer Academic Publishers, Dordrecht, Boston, London, 1996.

Mary Dalrymple. *Lexical Functional Grammar*, volume 34 of *Syntax and Semantics*. Academic Press, London, 2001.

Mary Dalrymple and Irina Nikolaeva. *Objects and information structure*. Cambridge University Press, Cambridge UK, 2011.

Henriëtte de Swart and Donka Farkas. *The Semantics of Incorporation: from Argument Structure to Discourse Transparency*. CSLI Publications, 2003.

Katalin É. Kiss, editor. *Discourse-Configurational Languages*. Oxford University Press, Oxford, New York, 1995.

Katalin É. Kiss. *The Syntax of Hungarian*. Cambridge Syntax Guides. Cambridge University Press, Cambridge, 2002.

Katalin É. Kiss. Eseményszerkezet és topik-predikátum tagolás a magyar mondatban [Event Structure and the Topic-Predicate Division in the Hungarian Sentence]. *Nyelvtudományi Közlemények*, 101:160–172., 2005.

Katalin É. Kiss. Focussing as predication. In Valéria Molnár and Susanne Winkler, editors, *The Architecture of Focus*. Mouton de Gruyter, Berlin, 2006.

Anna Gazdik. Multiple Question in French and in Hungarian. An LFG Account. In Miriam Butt and Tracy Holloway King, editors, *Proceedings of the LFG10 Conference*. CSLI Publications, 2010.

Anna Gazdik. *Multiple Question in French and in Hungarian. A Lexical-Functional Analysis with Special Emphasis on the Syntax-Discourse Interface*. PhD thesis, Université Paris Diderot-Paris 7; ELTE Budapest, June 2011.

Beáta Gyuris. *The Semantics of Contrastive Topics in Hungarian*. PhD thesis, ELTE, Budapest, 2002.

Beáta Gyuris. Quantificational contrastive topics with verum/falsum focus. *Lingua*, 119:625–649., 2009.

Joachim Jacobs. Funktionale satzperspektive und illokutionssemantik. *Linguistische Berichte*, 91:25–58., 1984.

László Kálmán. Word Order in Neutral Sentences. *Approaches to Hungarian*, 1: 13–23., 1985a.

László Kálmán. Word Order in Non-neutral Sentences. *Approaches to Hungarian*, 1:25–37., 1985b.

László Kálmán, editor. *Magyar leíró nyelvtan. Mondattan 1*. Tinta Könyvkiadó, Budapest, 2001.

László Kálmán, Gábor Prószéky, Ádám Nádasdy, and György C. Kálmán. Hocus, Focus and Verb Types in Hungarian Infinitive Constructions. In Abraham Werner and Sjaak de Meij, editors, *Topic, Focus and Configurationality*, pages 129–142. Benjamins, Amsterdam, 1986.

Tracy Holloway King. *Configuring Topic and Focus in Russian*. PhD thesis, Stanford University, 1993.

Tracy Holloway King. Focus domains and information structure. In Miriam Butt and Tracy Holloway King, editors, *Proceedings of the LFG97 Conference*, University of California, San Diego, 1997. CSLI Publications.

András Komlósy. Complements and adjuncts. In Ferenc Kiefer and Katalin É. Kiss, editors, *The Syntactic Structure of Hungarian*, pages 91–177. Academic Press, New York, San Diego, 1994.

Tibor Laczkó and György Rákosi. On Locative Dependencies Involving Particle Verbs in Hungarian. Talk presented at the Thematic Workshop on Spatial and Temporal Relations in LFG, April 2011.

Louise Mycock. *A New Typology of Wh-Questions*. PhD thesis, Manchester University, 2006.

David Pesetsky. Wh-in-situ: Movement and unselective binding. In E. Reuland and A. ter Meulen, editors, *The Representation of (In)definites*. MIT Press, Cambridge MA, 1987.

Luigi Rizzi. The fine structure of the left periphery. In Haegeman Liliane, editor, *Elements of Grammar*, pages 281–337. Kluwer, Dordrecht, 1997.

Balázs Surányi. Preverbs, chain reduction, and phases. In Marcel den Dikken and Robert Vago, editors, *Approaches to Hungarian*, volume 11. John Benjamins, Amsterdam, 2009.

Enric Vallduví. *The Informational Component*. Garland, New York, 1992.