



Focus Phenomena in a Parallel LFG Grammar of Hungarian and English

Tibor LACZKÓ
György RÁKOSI

Institute of English and American Studies
University of Debrecen
laczkot@delfin.unideb.hu
rakosigy@hotmail.com

Abstract. The paper gives an overview of the grammar of parallel Hungarian and English focus phenomena and outlines an account in the framework of *Lexical-Functional Grammar* (LFG). Focus constructions are subject to variation with respect to their grammatical properties both within and across the two languages. The LFG framework is well-suited to represent this variation in a systematic way and, at the same time, to capture at the level of *information-structure* the shared contribution that various focus constructions make.

Keywords: focus, information structure, LFG, variation

1. Introduction

For the last two decades, there has been an increased interest in generative grammar in what has come to be known as *interface phenomena*. The issue is the proper treatment of the flow of information between the core syntactic module(s) and other representational levels, especially lexical-conceptual structure, morphological structure, semantic structure and phonological structure. The study of the interfaces is the central objective of the Minimalist Program, and it has also received increased attention in such non-transformational generative theories as LFG.

In this paper, we provide an overview of one such interface phenomenon: the grammatical encoding of focus. Our immediate goal is threefold. First, in Section 2, we briefly describe the grammar of focus. Focusing primarily serves to divide

the clause into a part that expresses new information (focus) and another one which expresses old information (background). It is a typical interface phenomenon in being a grammatical device that encodes what is essentially a non-grammatical, discourse-governed contrast between old and new information. Its proper linguistic treatment is a non-trivial issue for this very reason. Our second aim, to be addressed in Section 3, is to narrow down on the grammar of focus marking in Hungarian and English. In particular, we investigate the division of labour between syntax and phonology in the encoding of focus across different focus constructions, as well as the presence or absence of a direct semantic impact of focus. Third, in Section 4 we present the outlines of an LFG-theoretic account of focus phenomena, building on recent advances of the theory which expand the classical LFG architecture by the introduction of *information-structure*. We present our conclusions and an outlook on our future research plans in Section 5.

2. Focus phenomena and their linguistic representation

2.1. Defining focus

In the widest sense, focus is the part of the clause that is new or asserted. The rest of the clause, which includes old and often presupposed information, is generally referred to as the background.¹ In essence, focusing thus serves to partition the clause in two, a function most manifest in question-answer pairs like (1) below:

- (1) “*What did Mary buy?*”
“*Mary bought A HAT*”².”

The new information that the answer provides is that what Mary bought is *a hat*, this noun phrase being in focus. The rest of the clause serves as the background and is presupposed old information, since we already know from the question that Mary bought something.

It is a relatively strong universal that focus always receives prosodic marking. The universal is formulated in Reinhart (1995) as follows:

¹ We do not attempt here to overview the huge literature on focus and related phenomena, and concentrate on works and issues that are directly relevant to our purposes. For comprehensive overviews of the literature, we refer the reader to É. Kiss (1998), Szendrői (2001, 2004) and Vallduví (1992, 1993).

² Capitals are used to mark focused constituents. Note that there may be mismatches between focus content and prosodic marking, see the discussion below (2).

(2) **Stress-focus correspondence** (Reinhart 1995)

The focus of an utterance always contains the main stress of the utterance.

As the formulation suggests, the prosodic marking can in certain cases cover only part of the focused constituent. In the answer in (1) above, for example, the indefinite article is unaccented, even though the whole noun phrase *a hat* is in focus. There may be exceptions to this universal inasmuch as a focused constituent need not always contain the main stress in certain languages (see Szendrői 2004 for some discussion), but (2) is certainly valid for English and Hungarian. In both languages, focussed constituents must be marked prosodically, and, as we will see, variation across focus constructions is in part variation in the presence or absence of direct syntactic coding.

In one currently popular approach, focus is a unitary phenomenon in the sense that its sole function is to introduce new information (cf. especially Reinhart 1995, Szendrői 2001, 2004). É. Kiss (1998), however, makes strong arguments that we must distinguish between what she calls *information* and *identificational focus*. This distinction is grammaticalized since different constructions may express either one or the other focus type. Clause-final focus in English (1) is a representative of the first strategy: as information focus, it merely presents new information, which, however, is not meant to be exhaustive. In other words, the dialogue in (1) is well-formed even if Mary bought other things than a hat.

Clefting, on the other hand, is argued by É. Kiss to be an operation that codes identificational focus. Identificational focus is stronger semantically than informational focus, and its impact is described by É. Kiss (1998: 245) as follows:

An identificational focus represents a subset of the contextually or situationally given elements for which the predicate phrase can potentially hold; it is identified as the exhaustive subset of this set.

The difference between the two focus types is illustrated by the following test³:

- (3) “*Mary bought A HAT.*”
 “*No, she bought a coat, too.*”
- (4) “*It was A HAT that Mary bought.*”
 “*No, she bought a coat, too.*”

In both (3) and (4), the reply is interpreted as a negation of exhaustivity. The identificational focus coded by the cleft in (4) asserts exhaustivity, and consequently, this semantic feature can be meaningfully negated in the reply. (3)

³ É. Kiss (1998: 251) credits Donka Farkas (p.c.) with the authorship of this test.

involves only information focus, which does not include exhaustivity. The reply therefore is inappropriate.

A related distinction is made in the LFG literature by Choi (1997, 2001). She distinguishes between *contrastive* and *completive focus* by using the binary feature *+/- prominent in discourse*, the former being +PROM and the latter being -PROM. This dichotomy is very similar to É. Kiss' distinction, but the overlap between the two is only partial. É. Kiss (1998) argues at length that identificational focus is often contrastive, but it need not always be (see op. cit. for details).

Importantly, whether contrastive or not, identificational focus always changes the truth-conditions of the sentence by the necessary assertion of exhaustivity. This type of focus must therefore be represented in semantic structure, whereas information focus does not necessarily have a direct semantic impact (though it does affect information structure). We will therefore assume É. Kiss' dichotomy for the purposes of this paper.

2.2. Focus in the GB/Minimalist tradition

Szendrői (2001, 2004) argues at length that the representation of focus is a non-trivial issue in transformational grammarian models in general, and in Minimalism in particular. The fundamental problem is that focus is prosodically coded (see (2) above), and sometimes that is the only grammatical marking that it receives (as in the case of end focus in English, cf. (1)). In transformational grammar, however, the phonological/prosodic module (PF) has always been an output level for syntax with no feedback. What is more, PF cannot communicate directly with post-syntactic interpretive levels in standard models.

The problem then is this. We know that focus has an effect on the interpretation of the sentence, but when its only coding is through stress, there is no obvious way to capture this interpretive effect in the model. The standard account is to introduce a focus feature lexically upon the head of the focussed phrase, which will force the phrase to move into the specifier position of a FocusP, where it is interpreted as the focus of the clause⁴. Though this solution is technically adequate, it is counterintuitive in treating focus as a lexical feature (and introducing it, as it seems, in an arbitrary manner).

Szendrői (2001, 2004), following Reinhart (1995), proposes instead to modify the architecture of the theory. We can properly account for focus facts if we treat focus as essentially a PF-feature (and do not introduce it in the lexicon), and then allow PF to communicate directly with the interpretive module: LF. This solves the above problem at the cost of modifying the grammatical model via what counts as a radical change from the perspective of standard approaches.

⁴ Bródy (1995) is a by-now classical analysis of Hungarian focus marking along these lines.

2.3. Focus and the LFG architecture

Lexical-Functional Grammar assumes representational modularity, just like transformational grammars. The crucial difference is that the architecture is much less restricted both in allowing for the application of a greater number of modules (called *structures*) and for a less constrained interaction among them, as dictated by the exigencies of grammar writing. An immediate consequence is that in LFG, there is no principled problem whatsoever posed by sending information from, say, *phonological structure* to *semantic structure*, thereby creating a suitable environment for the representation of focus.

The confines of this paper do not allow us to introduce the LFG architecture in full, and we refer the interested reader to Bresnan (2001), Dalrymple (2001) and Laczkó et al. (2010) for overviews. Diagram 1 is included below to give a pictorial overview of the classical LFG architecture.

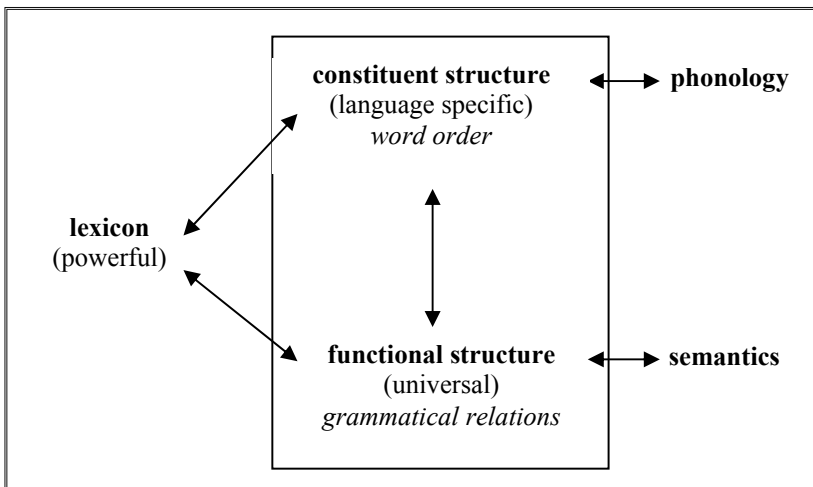


Diagram 1. *The classical LFG architecture*

The core syntactic modules are *c(onstituent)-structure*, which represents information about possibly language specific word order and constituency patterns, and *f(unctional)-structure*, which encodes featural information concerning the grammatical relations and predicate-argument structures underlying the surface strings of expressions. Importantly, this is a parallel, and not a procedural architecture. As a result, information flow is allowed in principle in any direction. In particular, phonology and c-structure can communicate in either order, but there is no principled constraint against phonology serving as input to f-structure or s(semantic)-structure.

In the classical LFG model, the discourse functions topic and focus were represented as syntactic features at f-structure. This treatment is adequate and in fact necessary in the case of, for example, syntacticized focus constructions like the English cleft-construction. Clefting is a way of encoding focus directly in syntax, with a clear truth-conditional effect, as we have seen in 2.1. In the classical architecture, f-structure is the primary input to s-structure, thus f-structural encoding guarantees that focus information will reach s-structure.

End-focus (as in 1 and 3 above), however, is not syntacticized and is not truth-conditional, and should, therefore, not be encoded in f-structure. To be able to represent non-syntacticized focus and to capture the meaning contribution that any focus construction has, a separate module, *i(nformation)-structure* has been introduced into the theory (see, a.o., Choi 1997, 2001, King 1997, King & Zaenen 2004). I-structure contains special attributes that represent the information content of the clause. In this paper, we will use three such features: ID-FOC for identificational focus, INF-FOC for information focus, and GROUND for background.

The result is an enriched architecture. King (1997), for example, proposes the following extended model (only the relevant aspects of the approach being included)⁵:

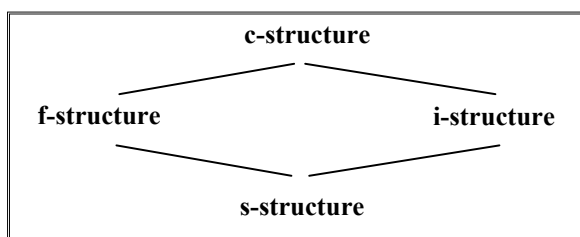


Diagram 2. *The extended LFG architecture of King 1997*

The focus phenomena that we discuss in Section 3, we argue, require a richer architecture to be properly represented. Our proposal is summarized in Diagram 3. Space limitations force us here to focus on designated aspects of this model, and we leave a full presentation, as well as a more complete justification, to another occasion.

⁵ King (1997) does not assume a mapping from f-structure to i-structure because a discourse function constituent does not always coincide with an f-structure constituent. The elements of the background, for example, typically do not form a constituent. See King (1997) for details.

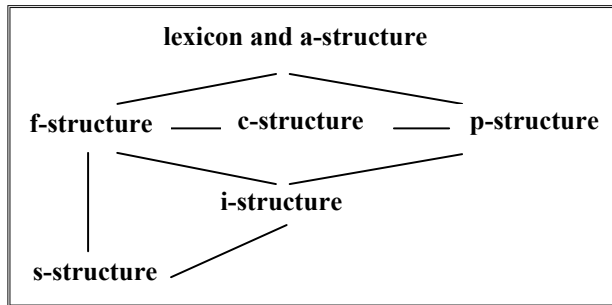


Diagram 3. *Our proposed architecture*

3. Parallel focus phenomena in Hungarian and English

Hungarian, unlike English, is known to be a discourse-configurational language, which encodes discourse functions directly in syntax. In contrast, English is regarded as configurational inasmuch as the basic syntactic functions (subject, object, indirect object) must be inserted into designated syntactic positions – a constraint that is not operational in Hungarian syntax at a descriptive level. Nevertheless, as we will point out directly, focus marking is not always configurational in Hungarian, neither is it always only prosodic in English.

3.1. Focus in the left periphery

Hungarian has long been known to be a language in which the primary means of focus marking is configurational. In particular, a focused constituent has to occupy an immediately preverbal position. As a result, focused constituents are in complementary distribution with so-called verbal modifiers (particles, bare nouns and certain adverbials that form a complex predicate with the verb). The examples below are presented by vertically aligning the preverbal positions in the respective sentences.

- (5) a. *János le- tette a könyv-et a padló-ra.*
 John down- put the book-ACC the floor-onto
 ‘John put down the book on the floor.’
- b. *János A KÖNYVET tette le a padlóra.*
 ‘It was THE BOOK that John put down on the floor.’
- c. *JÁNOS tette le a könyvet a padlóra.*
 ‘It was JOHN that put down the book on the floor.’
- d. *A PADLÓRA tette le János a könyvet.*
 ‘It was ON THE FLOOR that John put down the book.’

As the English translations suggest, preverbal focus is exhaustive and is therefore identificational focus in É. Kiss' terminology. Since it is a grammaticized discourse function, it is represented in f-structure. Note, however, that to trace its presence, f-structure needs input both from c-structure and p-structure (for not every preverbal constituent is focused, only those that themselves receive the main stress of the clause).

One type of focus constructions in English (a left-dislocated constituent associated with a pitch accent) also occurs preverbally, albeit in a clause-initial rather than in an immediately preverbal position. É. Kiss (1998: 251) shows that this construction codes information focus, as is justified by the exhaustivity test introduced in 2.1:

- (6) “A HAT, Mary picked for herself.”
 “No, she picked herself a coat, too.”

Focus left-dislocation is configurationally coded (and, naturally, prosodically marked). Being information focus, it does not affect the truth-conditions of the sentence, and, consequently, we assume that it is primarily an i-structure phenomenon, which does not receive f-structure encoding.

3.2. End focus

The most frequent type of focus in English is clause-final. This is a consequence of the fact that the focused constituent must contain the main stress of the clause (cf. 2), and the nuclear stress rule places the main stress on the rightmost constituent in English (Reinhart 1995, Szendrői 2004).⁶ É. Kiss (1998) shows that this type of focus also exist in Hungarian, see (7b) below.

- (7) a. *Mary gave a book TO JOHN.*
 b. *Mari ad-ott egy könyv-et JÁNOS-NAK.*
 Mary give-PAST.3SG a book-ACC John-DAT
 ‘Mary gave a book TO JOHN.’

In Hungarian, this type of focus has a marked nature, as the default focus marking is preverbal (3.1).

⁶ End focus can be shifted in English to a non-final constituent:

(i) *Mary gave A BOOK to John.*

Reinhart (2006) presents psycholinguistic and acquisition evidence that shifted focus is indeed more marked (i.e., a more costly operation) than neutral end focus. É. Kiss (1998) remarks that such a focus shift in the postverbal field is rare in Hungarian. In fact, the direct equivalent of (i) verges on ungrammaticality, the focused NP would be placed in the regular preverbal focus position.

End focus is information focus in both languages, recall the discussion of (3) in 2.1. As such, we do not introduce it into f-structure, but represent it directly at i-structure.

3.3. Clefting

A cleft construction is essentially a reduced biclausal structure that serves to express identificational focus (see É. Kiss 1998). The focused constituent is transposed into a matrix clause that has an expletive subject in English:

- (8) a. *It was JOHN who/that walked into the room.*
 b. *It was INTO THE ROOM that John came.*

Clefting is the primary device of coding identificational focus in English.

Clefts are also productive in Hungarian, although they are probably less frequently used than in English. This is, again, due to the fact that information focus is generally placed locally into the preverbal position. Note that it is the matrix focus position that is targeted in clefts, too:

- (9) a. *JÁNOS volt az, aki/*hogy be-jött.*
 John was that.NOM who.NOM/that in-came
 'It was JOHN who came into the room.'
 b. **A SZOBÁ-BA volt az, ahova/hogy János be-jött.*
 the room-into was that.nom where.to/that John in-came
 'It was INTO THE ROOM that John came.'

As (9a-b) testify, the Hungarian cleft construction is not full grammatical equivalent to its English counterpart. First, the subordinate clause can only be introduced by a *who*-pronoun, but not by a *that*-type complementizer. Second, the Hungarian construction does not license the focusing of oblique phrases (9b)⁷.

Nevertheless, both languages have an obviously grammaticalized and truth-conditionally relevant cleft-construction to be employed for expressing focus. Therefore both cleft-constructions are to be represented in f-structure as well as in other relevant representational modules.

⁷ We plan to develop a comprehensive and comparative analysis of English and Hungarian cleft sentences in future work.

4. Outlines of an LFG account

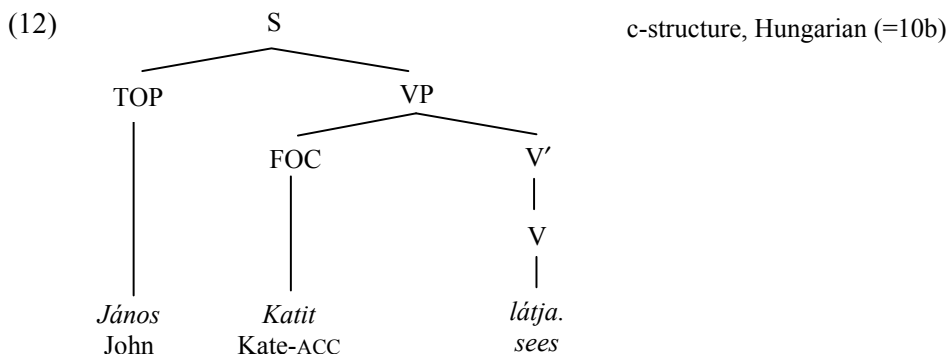
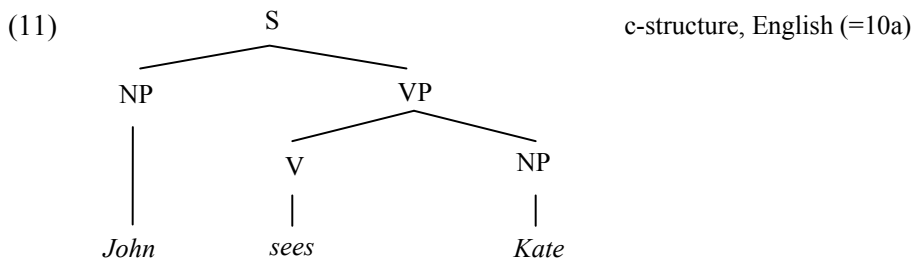
Space limitations confine us to offering only the outlines of an LFG account. In particular, we will concentrate on the treatment of the following English and Hungarian focus constructions:

- (10) a. *John sees KATE.*
 b. *János KATI-T lát-ja.*
 John Kate-ACC see-3SG
 ‘It is KATE that John sees.’

As discussed in the previous section, end focus is the most frequent focus construction in English, whereas Hungarian generally employs preverbal focus. We have also seen that the two constructions also differ in their function: (10a) contains information focus, whereas (10b) has truth-conditionally relevant identificational focus.

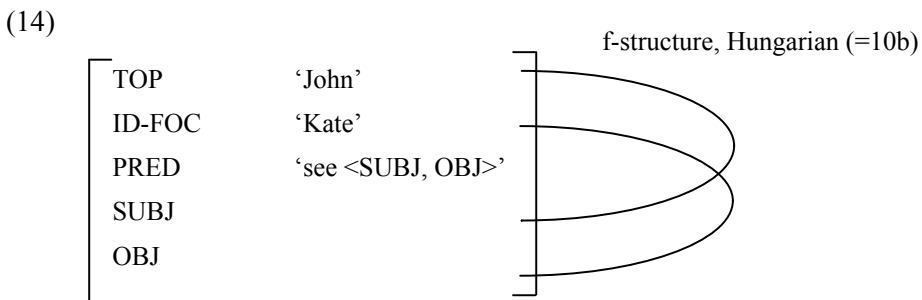
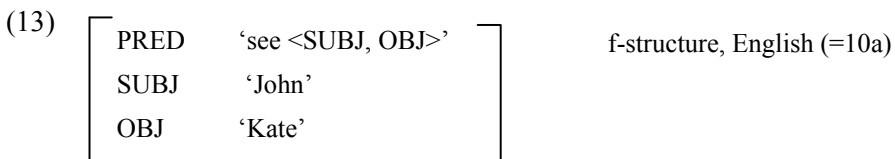
In what follows, we present an LFG-style analysis of the two constructions, focusing on what is most directly relevant to the comparison: c-structure, f-structure, and i-structure. For expository purposes, we omit features that have no significant contribution to the presentation.

Let us start with the two respective c-structures.



The obvious difference between the two structures is that whereas the discourse functions Topic and Focus are directly coded in Hungarian syntax, they are not in English. English places the subject and the object in a designated syntactic position, and though subjects are frequently topics, and clause-final constituents often receive focus interpretation, neither correspondence is necessary.

The functional structure of both constructions is the same with respect to the predicate-argument structure of the two respective verbs, and with respect to the assignment of the basic syntactic functions to the arguments. Where they differ is that the Hungarian f-structure will include an identificational focus and a topic feature, for both are grammaticized in Hungarian. Quite unlike their English counterparts in (10a), which are not represented at f-structure.



The arches in the Hungarian f-structure mark the relevant dependencies: the value of the TOPIC attribute is identical to the value of the SUBJECT attribute, and the same holds for ID-FOC and OBJ, respectively. Features that are not immediately necessary for us (agreement features, tense, etc.) have been omitted.

Finally, i-structure is utilized to represent the impact focusing has on the information structure associated with the utterance of the two test sentences. Here the only difference between the two constructions is that one contains information focus, and the other identificational focus.

$$(15) \left[\begin{array}{ll} \text{INF-FOC} & \text{'John'} \\ \text{GROUND} & \left\{ \begin{array}{l} \text{'Kate'} \\ \text{'sees'} \end{array} \right\} \end{array} \right] \quad \text{i-structure, English (=10a)}$$

$$(16) \left[\begin{array}{ll} \text{ID-FOC} & \text{'John'} \\ \text{GROUND} & \left\{ \begin{array}{l} \text{'Kate'} \\ \text{'sees'} \end{array} \right\} \end{array} \right] \quad \text{i-structure, Hungarian (=10b)}$$

We follow King (1997) in assuming that the value of the background attribute (ground) is an unordered set. This is so because i-structure is not identical to a semantic representation, and what matters is how focusing partitions the clause in two.

5. Summary

In this paper, we have briefly overviewed parallel focus phenomena in English and Hungarian, and argued that they can be analyzed in the framework of Lexical-Functional Grammar in a principled manner. The parallel architecture of LFG is well-suited to represent such interface phenomena as focus, especially when the classical LFG architecture is augmented by the introduction of information structure.

The empirical coverage and the analysis presented is necessarily partial, but we believe we have succeeded in demonstrating the power of the LFG model. Our most important research goal is to work out the details of the analysis, and apply the result directly to the LFG-based computational grammar that we are currently developing⁸.

Our research plans include the following areas to be addressed in the next phase of our project. First, the empirical coverage needs to be extended to include, among other phenomena, multiple focus constructions, single and multiple *wh*-questions, as well as a detailed analysis of topicalization. Second, we need to elaborate the mapping mechanisms that determine the information flow among the modules we discussed in Section 4, paying special regard to cases of mismatch between the different types of representations.

⁸ Continuously updated information about this computational linguistic enterprise can be read on our homepage at <http://hungram.unideb.hu>.

Acknowledgements

The authors gratefully acknowledge that the research that has been reported here has been supported by OTKA (*Hungarian Scientific Research Fund*), grant number: K 72983. Tibor Laczkó further acknowledges support from the Research Group for Theoretical Linguistics of the Hungarian Academy of Sciences at the Universities of Debrecen, Pécs and Szeged.

References

- Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell Publishers.
- Bródy, Michael. 1995. Focus and checking theory. In: Kenesei, István (ed.), *Levels and Structures, Approaches to Hungarian 5*, 31-43, Szeged: JATE.
- Choi, Hye-Won. 1997. Information structure, phrase structure and their Interface. In: Butt, Miriam & Tracy Holloway King (eds.), *Proceedings of the LFG07 Conference*. Stanford: CSLI Publications.
online: <http://csli-publications.stanford.edu/LFG/2/lfg97.html>
- Choi, Hye-Won. 2001. Phrase structure, information structure and resolution of mismatch. In Sells, Peter (ed.), *Formal and Empirical Issues in Optimality Theoretic Syntax*. Stanford, 17-62, CA: CSLI Publications.
- Dalrymple, Mary. 2001. *Lexical Functional Grammar. Syntax and Semantics Volume 34*. New York: Academic Press.
- É. Kiss, Katalin. 1998. Identificational focus versus information focus. *Language* 74 (2): 245-273.
- King, Tracy Holloway. 1997. Focus domains and information structure. In: Butt, Miriam & Tracy Holloway King (eds.), *Proceedings of the LFG07 Conference*. Stanford: CSLI Publications. online:
<http://csli-publications.stanford.edu/LFG/2/lfg97.html>
- King, Tracy Holloway & Annie Zaenen. 2004. F-structures, information structure, and discourse structure. Talk presented at *the Winter School in LFG and Computational Linguistics*, University of Canterbury, Christchurch, New Zealand. Extended abstract available online:
<http://csli-publications.stanford.edu/LFG/9/lfg04kingzaenen-abs.html>
- Laczkó, Tibor, György Rákosi & Ágoston Tóth. 2010. HunGram vs. EngGram in ParGram. On the comparison of Hungarian and English in an International Computational Linguistic Project. In: Hegedűs, Irén & Sándor Martsa (eds.), *CrossSections. Volume 1: Selected papers in linguistics from the 9th HUSSE conference*, 81-95. Institute of English Studies, Faculty of Humanities, University of Pécs: Pécs.
- Reinhart, Tanya. 1995. *Interface strategies. OTS Working Papers*. Utrecht: UIL-OTS.
- Reinhart, Tanya. 2006. *Interface Strategies. Optimal and costly computations*. Cambridge, MA: The MIT Press.

- Szendrői, Kriszta. 2001. *Focus and the syntax-phonology interface*. PhD dissertation. London: UCL.
- Szendrői, Kriszta. 2004. Focus and the interaction between syntax and pragmatics. *Lingua* 114: 229-254.
- Vallduví, Enric. 1992. *The Informational Component*. New York: Garland.
- Vallduví, Enric. 1993. *Information packaging: A survey*. MS. Report number HCRC/RP44. Human Communication Research Centre, University of Edinburgh.