Chapter 32

LFG and Finno-Ugric languages

Tibor Laczkó
Károli Gáspár University of the Reformed Church in Hungary

The chapter discusses some salient, sometimes competing, LFG analyses of a variety of (morpho-)syntactic phenomena in Finno-Ugric languages, with occasional glimpses at alternative generative approaches and at some related phenomena in languages belonging to Samoyedic, the other major branch of Uralic languages. We concentrate on clausal c-structure representational issues, verbal modifiers, focused constituents, negation, copula constructions, argument realization, subject-verb agreement, differential object marking, evidentiality and a set of noun phrase phenomena related to event nominalization. It argues that LFG provides an appropriate and suitably flexible formal apparatus for a principled analysis of all the phenomena in all the Finno-Ugric languages discussed here. In addition, it shows that the analysis of some of these phenomena can also contribute to LFG-internal theorizing.

1 Introduction

1.1 General remarks on Finno-Ugric languages

Finno-Ugric is one of the two branches of Uralic, the other branch being Samoyedic. In Figure 1 we show the major branches of the Uralic family tree and those leaves (languages) that are discussed, or at least mentioned, in this chapter. This figure is in accordance with the general remarks in the introductory chapter of Miestamo et al. (2015) on the representation of the Finno-Ugric branch.¹ We use

¹We are thankful to Anne Tamm for helpful discussions of certain family tree issues.
the names of the individual languages as they appear in that volume. The authors point out that, although there are several alternative approaches to this branch, most of them share the view that the following language groups are valid genealogical units: Samoyedic, Ugric, Permic, Mari, Mordvin, Saamic and Finnic. However, the details of the relationships among certain languages are subject to variation across these competing approaches.

For the sake of a complete picture, we have included the Samoyedic branch as well. In the Northern branch there are two major sub-branches: Enets-Nenets and Nganasan. From the Enets-Nenets sub-branch Tundra Nenets will be discussed and compared with some Finno-Ugric languages in Section 7.1.2 with respect to differential object marking. The only living representative of the Southern branch is Selkup, also mentioned in Section 7.1.2. Saamic languages also have a variety of sub-branches. From these languages Inari Saami will be discussed in Section 5.2 on copula constructions and in Section 7.1.1 on subject-verb agreement.

As regards the geographical distribution of the languages indicated in Figure 1, Estonian is primarily spoken in Estonia, Hungarian is spoken in Hungary, Finnish and Inari Saami are spoken in Finland, and all the other languages are spoken in Russia.

Several languages belonging to the Finno-Ugric branch of Uralic languages have a considerable number of properties that have contributed to linguistic research in LFG. On the one hand, these phenomena provide empirical or typological evidence for theoretical generalizations. On the other hand, they exhibit cases in which LFG is well-suited for the development of principled analyses. Such phenomena include, but are not limited to, discourse-functionality, negation, wh-questions, copular clauses, particle-verb constructions, event nominalization, possessive constructions, the nature and inventory of grammatical functions, evidentiality, rich inflectional morphology, partitives, duals and complex agreement patterns.

In this chapter we can only concentrate on those phenomena in Finno-Ugric languages that have been analyzed in an LFG framework in such a way that the summary of the given analysis within the limitations of space serves the purposes

---

2 Several languages in this figure are also referred to by alternative names in some other works, e.g. Khanty = Ostyak, Mansi = Vogul, Udmurt = Votyak, Mari = Cheremis; see the discussion of Dalrymple & Nikolaeva (2011) in Section 7.1.2, for instance. When we cite authors, we keep the version of the name of a language that they use.

3 For a recent, fundamentally similar Uralic family tree representation indicating all the languages (including those that are extinct by now), see Maticsák (2020).
Figure 1: The (simplified) family tree of Uralic languages
of the chapter, as outlined in the foregoing paragraph. Consequently, this determines which languages appear in the chapter. Given that Hungarian is the most intensively and extensively researched Finno-Ugric language in LFG, the discussions of LFG analyses of Hungarian phenomena outnumber the discussions of phenomena in other Finno-Ugric languages. For further information on related and additional phenomena and other Uralic or Finno-Ugric languages, the interested reader is referred to the following comprehensive sources: Abondolo (1998), Dryer & Haspelmath (2013), Miestamo et al. (2015) and de Groot (2017). The online journal *Finno-Ugric Languages and Linguistics* (http://full.btk.ppke.hu) regularly contains generative papers on Finno-Ugric languages. In addition, Tamm & Vainikka (2018) present an overview of generative works on Finnish and Estonian syntax.

As regards comprehensive analyses of several phenomena in Hungarian, Laczkó 2021 offers a synthesis of his earlier LFG(-XLE) accounts of the following phenomena in Hungarian finite clauses: sentence structure, verbal modifiers, operators, negation and copula constructions. He posits all this in the context of a critical overview of alternative Chomskyan and lexicalist approaches to these phenomena. Tamm (2004c) develops a comprehensive LFG approach to the relations between Estonian aspect, verbs and case.

The following databases on Uralic languages are useful resources about their syntactic properties: the Uralic language typological data set at bedlan.net/data/uralic-language-typological-data-set the Selkup and Kamas corpora at www.slm.uni-hamburg.de/ine, and the typological database of Ugric languages at en.utdb.nullpoint.info.

---

4 For introductions to LFG in Hungarian, see Laczkó (1989) and Komlósy (2001). The following works also have introductory sections to LFG: Szabó (2017) in Hungarian and Tamm (2004a) in Estonian.

5 In Section 9 we make brief references to additional works on Uralic languages in general and Finno-Ugric languages in particular that we cannot discuss here for limitations of space.

6 See, for instance, Brattico (2019) on Finnish word order, É. Kiss (2020) on pronominal objects in Ob-Ugric, and Asztalos (2020) on focus in Udmurt.

7 In her review, Anne Tamm has kindly provided the following information about the history of syntactic research on Estonian. “For a long while since the mid-20th century, there was more work on Estonian syntax than on Finnish syntax. Keeping abreast with western mainstream linguistics in the 60s, 70s and early 80s resulted in numerous formal syntactic works and a tradition of understanding syntax that is, in spirit, rather similar to LFG approaches. Rätsep (1978), for instance, is a lexicalist analysis of patterns of argument structures and their alternations; this work has certainly been influential in the context Uralic syntax. Tamm (2012c) provides an overview of the treatment of verb classes in this and related works, these early generative-style lexicalist works are available in Estonian only […]. Almost all LFG work on Estonian expands that work in some way.”

8 In his XLE work he further develops Laczkó & Rákosi’s (2008–2019) implemented Hungarian grammar.
1.2 The structure of the chapter

In accordance with the scopes of LFG works on Finno-Ugric languages, the significantly larger part of this chapter (Section 2-Section 7) concentrates on the investigation of clausal phenomena, and this is followed by the discussion of salient LFG analyses of some noun phrase phenomena (Section 8). In Section 2 we discuss clausal c-structure representational issues by focusing on a variety of LFG approaches to Hungarian. In Section 3 we concentrate on verbal modifiers in Hungarian and Estonian in general and on their radically different relations to focus in these languages in particular. In Section 4 we offer a brief overview of an LFG-XLE analysis of negation in Hungarian by also pointing out its potential contribution to the treatment of negation phenomena cross-linguistically. In Section 5 we discuss LFG accounts of copula constructions in Hungarian, Inari Saami and Finnish. In Section 6 we deal with LFG treatments of some aspects of argument realization in Finnish and Estonian. In Section 7 we concentrate on a selection of morphosemantic phenomena: (i) subject-verb agreement in Inari-Saami and Finnish; (ii) differential object marking in Uralic with particular attention to Finno-Ugric languages; (iii) the grammaticalized expression of evidentiality in Udmurt and Estonian. In Section 8 we present a summary of a variety of LFG approaches to noun phrase phenomena in Hungarian: (i) c-structure issues; (ii) event nominalization, and we add a short section on the morpho-syntax of possessive noun phrases in Finnish and Hungarian. In Section 9 we make brief references to additional relevant LFG(-related) works on Finno-Ugric languages that space limitations have prevented us from discussing. In Section 10 we conclude.

2 C-structure representation in clauses

de Groot (2017) presents a very useful tabular comparison of the major word order properties of 21 Uralic languages. In Table 1 we present the parts of his table that are relevant for our current purposes.

As the table shows, in these languages word order is predominantly free (except for Enets and Nenets). The two major patterns are SVO and SOV with roughly the same frequency. In seven languages there is a designated preverbal focus position (and in one of them, Komi, there is an additional postverbal Foc position). In three languages the Foc position is clause final. This is the general word order picture. Below we fundamentally concentrate on Hungarian because several alternative LFG c-structure analyses have been proposed for this
Table 1: Word order properties of 21 Uralic languages (part of Table 11, de Groot 2017: 548)

<table>
<thead>
<tr>
<th>Language</th>
<th>word order</th>
<th>major pattern</th>
<th>focus position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finnish</td>
<td>free</td>
<td>SVO</td>
<td></td>
</tr>
<tr>
<td>Estonian</td>
<td>free</td>
<td>SVO</td>
<td>clause final</td>
</tr>
<tr>
<td>Votic</td>
<td>free</td>
<td>SVO</td>
<td></td>
</tr>
<tr>
<td>Ingrian</td>
<td>free</td>
<td>SVO</td>
<td>clause final</td>
</tr>
<tr>
<td>Veps</td>
<td>free</td>
<td>SVO</td>
<td>clause final</td>
</tr>
<tr>
<td>Karelian</td>
<td>free</td>
<td>SVO</td>
<td></td>
</tr>
<tr>
<td>South Saami</td>
<td>free</td>
<td>SOV/SVO</td>
<td></td>
</tr>
<tr>
<td>North Saami</td>
<td>free</td>
<td>SVO</td>
<td></td>
</tr>
<tr>
<td>Skolt Saami</td>
<td>free</td>
<td>SVO</td>
<td></td>
</tr>
<tr>
<td>Erzya</td>
<td>free</td>
<td>SVO</td>
<td></td>
</tr>
<tr>
<td>Mari</td>
<td>free</td>
<td>SOV</td>
<td>Foc V</td>
</tr>
<tr>
<td>Komi</td>
<td>free</td>
<td>SOV/SVO</td>
<td>Foc V / V Foc</td>
</tr>
<tr>
<td>Udmurt</td>
<td>free</td>
<td>SOV</td>
<td>Foc V</td>
</tr>
<tr>
<td>Hungarian</td>
<td>free</td>
<td>SOV/SVO</td>
<td>Foc V</td>
</tr>
<tr>
<td>Khatry</td>
<td>free</td>
<td>SOV</td>
<td>Foc V</td>
</tr>
<tr>
<td>Mansi</td>
<td>free</td>
<td>SOV</td>
<td></td>
</tr>
<tr>
<td>Nenets</td>
<td>not free</td>
<td>strict SOV</td>
<td>Foc V</td>
</tr>
<tr>
<td>Enets</td>
<td>not free</td>
<td>strict SOV</td>
<td>Foc V</td>
</tr>
<tr>
<td>Nganasan</td>
<td>free</td>
<td>SOV</td>
<td></td>
</tr>
<tr>
<td>Selkup</td>
<td>free</td>
<td>SOV</td>
<td></td>
</tr>
<tr>
<td>Kamas</td>
<td>free</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

language. In addition, we make some comparative remarks on Finnish and Estonian.

Hungarian is a classic example of a discourse configurational language: see É. Kiss (1995), for instance.9 The crucial empirical generalizations about Hungarian sentence structure are as follows. The fundamental sentence articulation is

9On sentence structure and discourse-functionality in Finnish in non-LFG frameworks, see Vilkuna (1995) and Brattico (2019), for instance. According to Vilkuna (1995), there is a preverbal K (contrast) and also a T (topic) position in Finnish. While fundamentally these two positions are also available in Estonian, on the basis of their experimental and corpus investigation, Sahkai & Tamm (2018b, 2019) claim that other types of constituents can also occur in the preverbal domain. While Hungarian exhibits strong discourse-configurationality, Estonian is only weakly discourse-configurational: see Sahkai & Tamm (2018b: 416–417). Hiitam (2003) argues that topic is to be defined semantically and not configurationally in this language. In
topic-predicate (also called topic-comment in a variety of approaches). In the topic field, the order of topics and sentence adverbs is free. In the preverbal domain, quantifiers follow the topic field. In neutral sentences\(^{10}\) there is a designated immediately preverbal position for a special constituent type: ‘verbal modifier’ (vm). This is a conventionally used cover term for a range of radically different categories sharing the syntactic property of occupying this designated preverbal position. Preverbs (also known as verbal particles or coverbs),\(^ {11}\) bare nouns, designated XP arguments, etc. are all assumed to be vms. Basically, the word order of postverbal elements is also free. In a non-neutral sentence the (heavily stressed) focused constituent occupies the immediately preverbal position, and, as a consequence, the vm has to occur postverbally, i.e. the vm and the focus are in complementary distribution. How to capture this complementarity is a crucial cross-theoretical issue. The two salient solutions are as follows. (i) There is only a single designated preverbal position for which focused constituents and vms compete. (ii) There are two distinct positions for the two elements: focus and vm. In this approach it needs to be explained why these two elements cannot co-occur.

Börjars et al. (1999) offer some general considerations against functional projections like TopP and FocP (à la Government and Binding Theory and the Minimalist Program) for languages like Hungarian and some hints at a possible LFG alternative with an extended verbal projection in which word order regularities are capturable by dint of Optimality Theoretic (OT) style constraints. They claim that the assumption that discourse functions are not necessarily associated with the specifier positions of functional projections allows an analysis of Hungarian in which quantifier phrases and topics are positioned within an extended verbal projection, avoiding the postulation of functional projections without heads. They propose that Hungarian sentences are VP projections, as in (1),\(^ {12}\) and they suggest that the immediately preverbal occurrence of the focused constituent should be captured in terms of OT constraints. In this work, there is no discussion of vms and their complementarity with focused phrases.

---

\(^{10}\)The standard description of a neutral sentence is that it does not contain negation or focus, it is not a wh-question, and it has level prosody.

\(^{11}\)Other Ob-Ugric languages have developed verbal particles to a lesser extent, see Zsirai (1933). For more information on Uralic (aspectual) verbal prefixation and verbal particles, see Kiefer & Honti (2003). For an analysis of Estonian sentence-final particles with focus, see Tamm (2004c: 224–242), discussed in Section 3.2.

\(^{12}\)The superscripts in V\(^ 1 \) and V\(^ 2 \) indicate bar-levels.
Adopting the basic representational assumptions and ideas of Börjars et al. (1999), in their OT-LFG framework, Payne & Chisarik (2000) develop an analysis of Hungarian preverbal syntactic phenomena: the complementarity of constituent question expressions, focused constituents, the negative marker and verbal modifiers.

Gazdik (2012), capitalizing on Gazdik & Komlósy (2011), outlines an LFG analysis of Hungarian finite sentence structure, predominantly driven by discourse functional assumptions and considerations. She postulates two sentence structure types, and she assumes that both structures are available to both neutral (N) and non-neutral (NN) sentences, which are distinguished by their different prosodic behaviours. (2) shows one of the two structures. Here the immediately preverbal XP has a presentational-focus-like function in N sentences and the standard identificational focus function in NN sentences. The other structure differs in one important respect: the preverbal element is a vm, and the vm and the verb are dominated by V'. The vm receives the usual phonological-word-initial stress in N sentences and the focus stress in NN sentences.
Laczkó (2014b), after a detailed critical overview of previous LFG approaches, postulates the skeletal sentence structure in (3).\(^{13}\) He argues against assuming an IP for the structural-categorial representation of Hungarian sentences and he argues for S as the core category.\(^{14}\) He proposes a CP/S alternative that is closest in spirit to É. Kiss’ (1992) special GB approach.\(^{15}\)

\[
\text{(3) \hspace{1cm} CP}
\]

\[
\hspace{1cm} C \quad S^*
\]

\[
\hspace{1cm} XP(τ) \quad S
\]

\[
\hspace{1cm} XP(τ) \quad VP^*
\]

\[
\hspace{1cm} XP(q) \quad VP
\]

\[
\hspace{1cm} XP(sp) \quad V'
\]

\[
\hspace{1cm} V \quad XP^*
\]

Adopting one of the most crucial aspects of É. Kiss’s (1992) analysis, he assumes that vms and focused constituents target the Spec,VP position. He employs disjunctive functional annotations to capture this preverbal complementarity.\(^{16}\)

Consider the following generalization. ‘The daughters of S may be subject and predicate’ (Bresnan 2001: 112). In his analysis, Laczkó proposes that this generalization should be modified in the following way.

\[
\text{(4) The daughters of S may be subject/topic and predicate.}
\]

He points out that this modification receives independent support from the following rule from Bresnan & Mchombo (1987).\(^{17}\)

---

\(^{13}\)In (3) \(τ\) stands for topic (position), \(q\) stands for quantifier (position), \(sp\) stands for the specifier position. \(S^*\) and \(VP^*\) encode the possibly iterative left-adjunction of \(XP(τ)\) and \(XP(q)\) to \(S\) and \(VP\), respectively.

\(^{14}\)In LFG IP and S are taken to be parametric options in Universal Grammar.

\(^{15}\)For a comparison of these GB and LFG approaches, see Laczkó (2020).

\(^{16}\)For details and the discussion of what other elements are assumed to compete for the Spec,VP position, see Section 3.1 and Section 4.

\(^{17}\)On the basis of (5), subject and/or topic is even more appropriate than subject/topic in (4).
Laczkó argues that a VP can contain a subject if the XP in \([S \ X P \ VP]\) is a topic. This requires all other occurrences of VP to be subjectless. In this scenario, the following three parametric options seem to emerge across languages: (i) strictly VP-external subject, as in English; (ii) VP-internal subject in a designated position, as in Russian\(^{18}\); (iii) VP-internal subject without a designated position, see Hungarian.

This section has demonstrated that LFG provides a suitably flexible formal apparatus by the help of which the sentence structures of typologically different languages can be described in a principled manner with respect to discourse functional configurationality.

3 Verbal modifiers and focus

In this section we discuss analyses of verbal modifiers in Hungarian (Section 3.1) and Estonian (Section 3.2).

3.1 Hungarian

As has been pointed out in Section 2, the crucial (cross-)theoretical question to address in the case of Hungarian is how to account for the preverbal complementarity of focused constituents and verbal modifiers. Compare the examples in (6). (6a) is a neutral sentence and the vm oda ‘to.there’, which is categorically a preverb, immediately precedes the verb. By contrast, (6b) is a non-neutral sentence, and in it the vm can neither precede nor follow the focused constituent (in SMALLCAPS) in the preverbal domain.

\[(6) \quad \text{Hungarian:} \]
\[
a. \quad \text{János minden-t oda adott Mari-nak.} \quad \text{John.nom everything-acc vm gave Mary-dat} \quad \text{‘John gave everything to Mary.’} \\
\]
\[
b. \quad \text{János minden-t (*oda) MARI-NAK (*oda) adott oda.} \quad \text{John.nom everything-acc vm Mary-dat vm gave vm} \quad \text{‘John gave everything TO MARY.’} \\
\]

\(^{18}\)See King (1995), for instance.
The cross-theoretic question is whether we should assume that the two constituents fight for one and the same syntactic position or that they occupy two distinct positions. With the salient exception of É. Kiss (1992, 1994), the GB/MP mainstream assumes two distinct positions, and employs a variety of principles that block the simultaneous occurrence of constituents in these positions: see Brody (1990) and É. Kiss (2004), for instance, and also see Laczkó (2021) for a comparative overview of different analyses of the complementarity of focused constituents and verbal modifiers in Hungarian.

Several LFG approaches have a similar view, see Ackerman (1987, 1990), Payne & Chisarik (2000), Mycock (2006, 2010), the basic idea being that vms get semantically and morphologically incorporated into the verb.\(^\text{19}\) In Section 2 we also pointed out that Gazdik (2012) has a special proposal. She employs two distinct sentence structures, both having neutral and non-neutral versions. The main point here is that the basic vm vs. focus contrast is treated in two different structural dimensions. Thus, this can be regarded as an extreme instance of assuming that the two elements do not fight for the same syntactic position.

By contrast, Laczkó (2014b) argues that focus constituents (ordinary foci, the immediately preverbal \textit{wh}-phrases and negated constituents)\(^\text{20}\) and vms (of various types) target the same Spec,VP position, hence their complementarity. In (7) we repeat the relevant part of his overall sentence structure shown in (3) in Section 2.

\[
(7) \quad \begin{array}{c}
\text{VP} \\
\text{XP(sp)} \quad V' \\
V \quad \text{XP*}
\end{array}
\]

Laczkó (2014b) employs disjunctive functional annotations to capture the complementarity of the elements he assumes to compete for this position.

As we pointed out in Section 2, vms come in several varieties: preverbs, idiom chunks, secondary predicates, designated reduced or full arguments. Preverbs are the central and theoretically by far the most challenging members of this

\(^{19}\)At first sight, it can be taken to be a supporting fact that the vms of the preverb type and the verb make up one phonological word, i.e. it is only (the first syllable of) the preverb that receives word-initial stress. However, even XP vms follow the same pattern (in which the following verb loses its word-initial stress).

\(^{20}\)On the details of negation in Hungarian, see Section 4.
heterogeneous group, because their combination with the finite verb, often called particle-verb construction (pvc), exhibits both lexical and syntactic properties (and the former motivate the incorporation analysis). Their most salient lexical characteristics are as follows. The preverb can affect the argument structure of the main verb, pvcs are often non-compositional, and both non-compositional and compositional pvcs can undergo productive derivational processes like event nominalization. However, the preverb and the main verb are strictly separable syntactically under clearly definable circumstances. For instance, as exemplified in (6) above, a focused constituent, as a rule, immediately precedes the main verb, and in such cases the preverb must occur postverbally.

In several recent LFG approaches, for instance Forst et al. (2010), Laczkó & Rákosi (2011), Rákosi & Laczkó (2011), Laczkó (2013a) and Laczkó (2014b), it is assumed that preverbs and other types of vms uniformly occupy a distinct preverbal syntactic position (typically Spec,VP), as opposed to the vm-incorporation analysis, which is primarily motivated by the preverbal complementarity of vms, focused and wh-constituents.

Forst et al. (2010) propose an LFG-XLE treatment of a variety of particle-verb constructions in English, German and Hungarian. Their main claim is that non-compositional and non-productive pvcs should be treated radically differently from compositional and productive pvcs. The former are best analyzed along lexical lines with the help of XLE’s concatenation device. By contrast, the authors argue that the productive pvc types call for a syntactic treatment. One of the most important motivations for this sharp distinction is that productive pvcs can be analyzed ‘on the fly’, i.e. automatically and straightforwardly, in the syntax, without previously and lexically encoding them. Their solution is complex predicate formation in the syntax by applying XLE’s restriction operator.21

Laczkó & Rákosi (2011) and Rákosi & Laczkó (2011) explore the tenability and implementational applicability of the approach proposed by Forst et al. (2010) by each developing an LFG-XLE analysis of two different pvc types. Laczkó & Rákosi (2013) posit this approach in a cross-linguistic and cross-theoretical context. As opposed to previous LFG accounts, Laczkó (2013a) argues that compositional pvcs should also be treated lexically in a manner similar to the treatment of non-compositional pvcs. He points out that one of the advantages of this uniform lexical treatment is that classical LFG’s view of the distribution of labour between

---

21For formal details, see Forst et al. (2010).
the lexical and the syntactic components of grammar can be maintained, at least in this domain. He also shows how various morphological processes (often consecutively) involving pvcs can be handled (e.g. causativization, nominalization, and preverb reduplication), which may cause potential problems for a syntactic analysis of compositional pvcs.

Laczkó (2014b) captures the preverbal complementarity of focused constituents and vms by assuming that they fight for the same Spec,VP position. He encodes this by associating the disjunctive sets of annotations in (8) with this position. The first disjunct of the main disjunction says that a constituent bearing any grammatical function can have the focus discourse function. The second disjunct handles vms. Laczkó employs XLE’s check feature device here.\(^{22}\)

\[
(8) \quad \{ (\uparrow \text{gf}) = \downarrow \\
(\uparrow \text{focus}) = \downarrow \\
\mid (\downarrow \text{check}_{\text{vm}}) = c + \\
\{ \uparrow = \downarrow \\
\mid (\uparrow \text{gf}) = \downarrow \} \}
\]

The check feature in (8) is used for all types of vms. It requires the presence, in Spec,VP, of an element lexically marked with the defining counterpart of this feature. Preverbs are intrinsically associated with this feature, i.e. in their lexical forms they are associated with the defining member of the check_{vm} feature pair, and they receive the functional (co-)head annotation, see the first disjunct in the second major disjunct. All the other types of vms are specified for this status by individual verbs. It depends on the verb whether it selects a vm, and, if so, which argument (bearing any subcategorized grammatical function) will be singled out, see the second disjunct in the second major disjunct.\(^{23}\)

---

\(^{22}\)The essence of this device is that check features come in pairs: there is a defining equation and it has a constraining equation counterpart. These check feature pairs, which can be used both in c-structure representations and lexical forms, can ensure that two elements will occur together in a particular configuration or a particular element occurs in a particular position. The check feature in (8) is of the latter type.

\(^{23}\)Laczkó also assumes that a wh-phrase (or, in multiple wh-questions the immediately preverbal wh-phrase) also fights for the Spec,VP position, so he adds another disjunction to (8) to capture this, by using additional (interrogative) check features: for details, see Laczkó (2014b). In addition, he assumes that negated constituents also occupy this position. Furthermore, he postulates that in the type of predicate negation in which there is no focused constituent, the negative marker also targets this position. Therefore, he adds two more disjuncts, see Section 4.
Laczkó (2014a) outlines an LFG analysis of a variety of vms other than pre-verbs: bare nouns,\textsuperscript{24} ob\l\ XP arguments, xcomp arguments and idiom chunks. The crucial aspect of this analysis is that in the lexical form of the verb taking any one of these vm types it is specified that either the verb occurs in a sentence containing a focused constituent or else its designated complement must occupy the Spec,VP position.\textsuperscript{25}

### 3.2 Estonian

Tamm (2004c) presents a detailed description of pvcs in Estonian, and she outlines an LFG analysis. She points out that Estonian separable particles are basically comparable to their Hungarian counterparts, the most important difference being that aspectual particles typically occupy the clause final focus position. Tamm distinguishes three basic uses of Estonian particles, and she discusses the particle ära, which can perform all the three functions. Consider her examples.

(9) Directional (deictic) use of ära, Estonian:

\[
\begin{align*}
\text{ära} & \quad \text{veerema} \\
\text{away} & \quad \text{roll} \\
\text{`roll away'}
\end{align*}
\]

Tamm points out that verbs combining with ära in this use have an implicit path argument that is only optionally realized overtly. The closest Hungarian counterpart is el ‘away’ (as in el-gurul ‘roll away’).

\textsuperscript{24}Viszket (2004) offers a detailed empirical description of a whole range of bare noun phrases in Hungarian. In neutral sentences these constituents can only occur immediately preverbally, in the vm position. In her LFG account of the syntax of bare noun phrases, Viszket adopts Laczkó’s (1995; 2000b) [+vm] feature and she also introduces a special [+vm] feature. Her new feature, when associated with a predicate in its lexical form, bans the occurrence of a bare NP in the vm position; practically, it prevents such a constituent from occurring in neutral sentences. Viszket identifies seven major types of predicates that need to be provided with this feature in their lexical forms. For instance, the verbs of pvcs, the predicates of certain idioms and certain predicates with resultative xcomps belong here. These types also have the [+vm] feature. In addition, there are predicates without the [+vm] feature that also need [+vm]. For example, nominal and adjectival predicates, and verbs that always need word-initial stress belong here. On partitive mass and plural NPs in Estonian, corresponding to bare nominal vms in Hungarian, see Tamm (2007a,c).

\textsuperscript{25}In her review, Anne Tamm points out that there are similarities between Laczkó’s analysis of Hungarian particle verbs and the analysis of Estonian particle verbs and aspect in Rätsep (1969) written in Estonian, which Tamm (2012c: 62–63, 72–75) has summarized, or Rätsep’s (1978) account of government structures of complex verbs in Estonian.
(10) Completive use of ära, Estonian:

Naaber suri ära.
neighbour die.PST.3SG ÄRA
‘The neighbour died.’

Verbs that combine with ära in this use have a theme or patient argument, obligatorily realized as a subject or an object. The closest Hungarian equivalents are meg ‘PFV’ (as in meg-hal ‘PFV-die’) and el ‘away’ (as in el-olvad ‘away-become.melted’).

(11) Bounding use of ära, Estonian:

Ta suudles tüdruku ära.
s/he kiss.PST.3SG girl.GEN ÄRA
‘S/he did the kissing of a girl.’

This sentence is appropriate in the following situation, for instance. Someone makes a bet to kiss a girl, and when this goal is achieved, the result can be reported by using this PFC. The closest Hungarian counterparts are meg ‘PFV’ (meg-ebédel ‘have/eat up one’s lunch’) and ki ‘out’ (as in ki-alussza magát out-sleep oneself.ACC ‘sleep one’s share, as much as needed’).

Tamm assumes that ära in its directional use has a PRED feature, and she gives the following lexical representation (Tamm 2004c: 231).

(12) ära P (↑ PRED) = ‘AWAY((↑ SUBJ))’
{(xcomp ↑) b1) ∨ ((xcomp ↑) b2)}

This encodes that the particle functions as the PRED of the lexical verb, and it has a subject argument. In addition, it has disjunctive existential constraints on the boundedness (b) attributes.

Tamm assumes that ära in its completive use also has a PRED feature, see her lexical form in (13) (Tamm 2004c: 232).

(13) ära P (↑ PRED) = ‘UP, COMPLETELY((↑ SUBJ))’
{(xcomp ↑) b1) = max ∨ ((xcomp ↑) b2)}

As opposed to its previous two uses, Tamm assumes that ära in its bounding use has no PRED feature, and it only encodes b and focus specifications, see (14) (Tamm 2004c: 229).

(14) ära Prt (↑ b1) = max
(↑ b2) = min
(↑ focus b1) = max
(↑ focus b2) = min
The particle in this use contributes f-structure information about the aspectual features of the clause, see the first two annotations, and it also encodes that this boundedness is the focused information, see the last two annotations.

In addition, verbal predicates can also carry aspectual information in their lexical forms. For instance, Tamm assumes that suudlema ‘kiss’, see (11) for instance, has the following lexical representation.

(15) suudlema  V  (↑ PRED) = ‘KISS〈(↑ OBJ)↑ OBJ)〉’
   (↑ b2)

This verb has an existential constraint on b2, which can be unified with the \textsc{min} value of the b2 of the particle in (14). Finally, the partitive and total case-markers on object arguments also encode aspectual information, so the entire aspectual feature value set of an Estonian sentence comes from three main sources via unification: verbs, aspectual particles and partitive/total case markers.\footnote{On the aspectual interaction of various verb types and partitive/total case in Estonian, see the discussion of Tamm’s (2006) analyses in Section 6.2.}

3.3 Concluding remarks

At the end of Section 3 we can make the following concluding remarks.

Hungarian \textsc{vm} phenomena are relevant from both cross-theoretical and LFG-specific perspectives in two important respects.

First, the focus-vm complementarity is a general generative theoretical issue. As the foregoing discussion shows, LFG provides a flexible formal platform even for alternative analyses significantly different in nature, which may be due to partially different views of the relevant components of the architecture of LFG.

Second, the behaviour of Hungarian \textsc{pvcs}, representing the major class of \textsc{vms}, is of great importance in the realm of complex predicates across typologically different languages, see Alsina et al. (1997) in general and Ackerman & Lesourd (1997) in that volume, in particular. The mixed lexical-morphological and syntactic properties of compositional and productive as well as non-compositional and unproductive \textsc{pvcs} pose a substantial challenge for both syntactically and lexically oriented generative theories, including LFG. From their entirely lexicalist perspective, Ackerman et al. (2011) give a taxonomic overview of a variety of approaches to complex predicates in LFG and HPSG. They point out that the classical models of the two theories rejected argument-structure-changing operations in the syntax, including complex predicate formation: see Bresnan (1982) and Pollard & Sag (1987). However, some more recent views in both theories admit syntactic complex predicate formation: see Alsina (1992, 1997), Butt (2003) and Müller (2006). By contrast, Ackerman et al. (2011), in their Realization-Based
Lexicalism (RBL) model, reject complex predicate formation in the syntax, and, as a trade-off, they admit analytic, i.e. multiple-word, forms of predicates in their lexicon as a marked option. As regards the treatment of Hungarian pvcs, Ackerman (2003) develops an RBL analysis. Forst et al. (2010), Laczkó & Rákosi (2011), Rákosi & Laczkó (2011) and Laczkó & Rákosi (2013), in their LFG-XLE framework, handle the productive types in the syntax by means of the RESTRICTION operator. By contrast, Laczkó (2013a), in the same framework, argues that both productive and unproductive pvcs need a lexical treatment.

As regards Estonian, Tamm’s (2004c) analysis has demonstrated that LFG also provides an appropriate formal apparatus for capturing the interplay of discourse functionality and the complex, multidimensional aspectuality system of this language.

4 Negation in Hungarian

Miestamo et al. (2015) discuss negation in Uralic in a comprehensive and systematically comparative fashion. They show that 17 Uralic languages employ negative auxiliaries. Hungarian, Khanty, Mansi and Estonian are exceptions in that they have no such auxiliaries. Of all these languages, we are only aware of a few LFG analyses of negation in Hungarian (most of them being rather sketchy and covering only some aspects of negation phenomena).

Laczkó (2014c) develops the first comprehensive LFG-XLE approach to the following six major types of clausal (aka predicate) and constituent negation in Hungarian: (i) ordinary constituent negation (the negated constituent is focused); (ii) universal quantifier negation without (another) focused element (= ordinary constituent negation, i.e. the negated universal quantifier is focused); (iii) universal quantifier negation with focus (= there is a preverbal focused constituent following the negated universal quantifier); (iv) predicate negation, without focus, the negative particle precedes the verb; (v) predicate negation, with focus, the negative particle precedes the verb; (vi) predicate negation, with focus, the negative particle precedes the focus. He proposes the following structural analysis.

---

27 There is a publicly accessible database on negation in Ob-Ugric and Samoyedic languages at https://www.univie.ac.at/negation/index-en.html.
28 Payne & Chisarik (2000), in their OT-LFG framework, also sketch an analysis of some of these types. For a critical overview, see Laczkó (2014c).
29 In (16) NEG stands for the (category of the) negative particle and the abbreviations in square brackets indicate the types of negation: [uQN] = universal quantifier negation, [EPN] = (VP-)external predicate negation, [CN] = constituent negation, [IPNH] = (VP-)internal predicate negation, phrasal adjunction, [IPNH] = (VP-)internal predicate negation, head-adjunction. The curly brackets signal the complementarity of [CN] and [IPNH].
In XLE grammars three devices are used for the encoding of negation: (i) the negative morpheme (whether bound or free) can be represented as a member of the \textsc{adjunct} set; (ii) it can encode the \textsc{neg +} feature value; (iii) it can encode the \textsc{pol neg} feature value. Laczkó (2015) points out that these devices are not used uniformly or consistently across the XLE grammars of various languages. He makes the following proposal. Type (i) is most appropriate when a language uses a free morpheme for the expression of negation, a negative particle. Type (ii) is best for bound negative morphemes. Type (iii) is most natural for encoding the scope of negation. In this proposed system, he develops an LFG-XLE analysis of Hungarian negative concord items.

In Laczkó’s (2014c) approach negated constituents also occupy the \textsc{spec,VP} position, see \textsc{cn} in (16). In addition, in the case of clause negation, the negative particle is also assumed to be in \textsc{spec,VP} when there is no focused constituent there, see \textsc{ipnph}. In his rules, Laczkó assumes that the negative particle has the category \textsc{neg}, and he uses a special XLE-style phrasal categorial label for negated constituents: \textsc{xpneg}. His \textsc{xpneg} rule is given in (17).

(17) \textsc{xpneg} \rightarrow \textsc{neg} \textsc{xp} \\
\downarrow \in (\uparrow \textsc{adj})
On the basis of these assumptions and rules, he adds the following two disjuncts to the disjunction in Spec,VP established so far, handling focused constituents and vms, shown in (8) in Section 3.1.30

\[(18) \{ \ldots \mid \text{XPneg} \mid \text{NEG} \} \]
\[
\begin{align*}
(↑\text{GR})=&\downarrow & \downarrow \in (↑\text{ADJ}) \\
(↑\text{FOCUS})=&\downarrow & (↑\text{FOCUS})=&\downarrow
\end{align*}
\]

As this section has shown, LFG provides an inventory of appropriate formal devices for analyzing complex negation phenomena in languages like Hungarian. At the same time, the treatment of these negation phenomena motivates examining the nature of the relevant formal devices carefully.

5 Copula constructions

5.1 Hungarian

The two major general LFG strategies for the treatment of copula constructions (ccs) across languages are represented by Butt et al. (1999) and Dalrymple et al. (2004). In the former approach, ccs are treated in a uniform manner functionally. The copula is always assumed to be a two-place predicate. It subcategorizes for a subject (subj) argument, which is uncontroversial in any analysis of these constructions, and the other constituent is invariably assigned a special, designated function designed for the second, ‘postcopular’ argument of the predicate: predlink. As opposed to this approach, in Dalrymple et al.’s (2004) view, the subj & predlink version is just one of the theoretically available options. In addition, they postulate that the copula can be devoid of a pred feature (and, consequently, argument structure) and in this use it only serves as a pure carrier of formal verbal features: tense and agreement. Finally, it can also be used as a one-place ‘raising’ predicate, associating the xcomp function with its propositional argument and also assigning a non-thematic subj function. When the postcopular constituent has the predlink function, it is closed in the sense that its subject argument is never realized outside this constituent. The xcomp and the predlink types involve two semantic and functional levels (tiers): the copula selects the relevant constituent as an argument. By contrast, when the copula is a mere formative, the two elements are at the same level (tier): the postcopular constituent is the real predicate and the copula only contributes morpho-syntactical features.

30Based on their prosodic and semantic behaviour, he assumes that both types of negative elements are focused constituents.
In LFG’s formal system, they are functional coheads. All this is summarized in Table 2.

Table 2: Three types of copular constructions, Dalrymple et al. (2004)

<table>
<thead>
<tr>
<th>role of the postcopular constituent</th>
<th>open</th>
<th>closed</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>main pred, the copula is a formative: functional coheads (single-tier)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(B)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>xcomp of the copula main pred: ‘be〈(↑ xcomp)〈↑ subj)’ (double-tier)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(C)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>predlink of the copula main pred: ‘be〈(↑ subj)〈↑ predlink)’ (double-tier)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As regards the treatment of copula constructions, Laczkó (2012) develops the first comprehensive LFG analysis of the following five most important types of cc in Hungarian: (i) attribution or classification; (ii) identity; (iii) location; (iv) existence; (v) possession. He subscribes to the view, advocated by Dalrymple et al. (2004) and also by Nordlinger & Sadler (2007), among others, that the best LFG strategy is to examine all cc individually, and to allow for diversity and systematic variation both in c-structure and in f-structure representations across and even within languages. This means that he rejects Butt et al.’s (1999) and Attia’s (2008) uniform predlink approach at the f-structure level. Table 3 summarizes the most important aspects of his analysis.31

Here we can only highlight the most crucial ingredients of this approach, concentrating on the ‘copula’s function’ and ‘argument structure’ columns in the table. In the attribution/classification type the copula has no pred feature and, thus, no argument structure, cf. column (A) in Table 2. The versions of the copula in all the other four cc types are two-place predicates. In the identity and possession types the second argument is assumed to have the predlink function, cf. column (C) in Table 2, while in the location and existence cc types it bears the obl_loc function, which is a variant of the closed type of postcopular constituents in column (C) in Table 2. Thus, in Laczkó’s (2012) analysis the copula has five dis-

---

31 The following abbreviations are used in Table 3: cop = copula, attr/class = attribution/classification, pr3: cop = is the copula present in the present tense and 3rd person paradigmatic slots? pr3: neg = how is negation expressed in pr3? vm = which element (if any) occupies the vm position in neutral sentences? s = subj, pl = predlink, interch = the two arguments’ grammatical functions are interchangeable in the 3rd person, spec = specific, def = definite, foc = focus, agr = agreement.
tinct lexical forms, which encode their respective sets of properties indicated in Table 3.

5.2 Inari Saami and Finnish

Toivonen (2007) analyzes subject-verb agreement phenomena in Inari Saami with a brief comparison with the corresponding Finnish phenomena, see Section 7.1.1. In her general approach, she also proposes an LMT (Lexical Mapping Theory: Findlay & Kibort forthcoming [this volume]) analysis of Inari Saami possessive constructions, again with a brief comparison with the Finnish counterparts. The empirical generalizations that she starts with, and which are relevant here, are as follows. (i) The possessed item is the subject. (ii) The possessed item bears nominative case. (iii) The possessor bears locative case. Consider one of her examples in (19), illustrating these facts.

(19) Inari Saami:
Muste lah tun.
I.LOC are.2SG you.NOM.SG
‘I have you.’

Toivonen assumes that the Inari Saami copula in this function is a two-place predicate with a theme (possessum) argument and a location (possessor) argument that receive the [-r] and the [-o] intrinsic specifications, respectively, and they are mapped onto SUBJ and OBL, respectively: see (20).
Toivonen compares Inari Saami and Finnish possessive constructions. For her comparison from the perspective of agreement, see Section 7.1.1. Here we concentrate on the GFS of the arguments of the possessive copulas of the two languages. Compare Toivonen’s Inari Saami example in (19) above with her corresponding Finnish example in (21).

(21) Finnish:
    Minulla on sinut.
    LADE is.3sg you.acc.sg
    'I have you.'

She makes the following generalizations about Finnish possession ccs. The possessum is either in nominative case (ordinary noun phrases) or in accusative case, see (21), and it has the OBJ function. The possessor is an oblique case-marked noun phrase, and it has the SUBJ function.

These two sections have shown that the behaviour of copula constructions in Hungarian, Inari Saami and Finnish exhibits remarkable variation, especially in the case of possession ccs. We can make the following concluding observations. On the one hand, the LFG framework, in this case, too, provides appropriate formal tools for feasible analyses of these construction types. On the other hand, the complexity of these phenomena can be used to argue for particular approaches in the inventory of LFG’s alternative formal devices in this particular domain.

6 Aspects of argument realization

6.1 Finnish

Pylkkänen (1997) develops an event-structure-based linking approach to Finnish causatives. She claims that her theory is minimalistic in two respects. On the one hand, in formalizing the relationship between event participants it minimizes reference to the thematic role properties of these participants (e.g. agent, theme and cause) by referring to events themselves. The basic assumption is that if one eventuality causes another then the participants of the former always rank higher
than those of the latter. On the other hand, an adequately developed system of inferring prominence relations obviates the need for argument structure, the level of representation mediating between event structure and grammatical functions. Pylkkänen’s system of inferring prominence from lexical semantic representations capitalizes on the following two assumptions proposed by Parsons (1990): (i) thematic roles are relations between events and individuals; (ii) causation is a relation between events. As a consequence, the thematic hierarchy is treated as applying at the level of individual events and not at the level of predicates. From this it follows that a predicate can have more than one thematic hierarchy: as many thematic hierarchies as events. All participants can be organized into a prominence hierarchy by ranking the individual thematic hierarchies with respect to each other. This ranking is regulated by Parsons’ second assumption: the causal relations between events. In essence, if \( e_1 \) causes \( e_2 \), then \( e_1^{\Theta_H} \) (the thematic hierarchy of \( e_1 \)) is ranked higher than \( e_2^{\Theta_H} \) (the thematic hierarchy of \( e_2 \)). Consider Pylkkänen’s two hierarchies in (22) and (23).

(22) Thematic Hierarchy: agent/experiencer > other > theme

(23) Event Hierarchy: cause(\( e_1, e_2 \)) \( \rightarrow \) \( e_1^{\Theta_H} > e_2^{\Theta_H} \)

Then linking constraints provide the mapping between the prominence hierarchy resulting from (22) and (23) and the following grammatical function hierarchy.

(24) \( \text{subj} > \text{obj} > \text{obj}_\theta > \text{obl} \)

In order for the linking constraints to be unifiable, Pylkkänen converts Parsons’ logical forms into attribute-value matrices. Consider her f-structure and event structure representation of (25), one of her examples, in (26).

(25) Finnish:
Matti kävel-yttä-ä koiraa.
Matti.NOM walk-CAUS-3SG dog.PAR
‘Matti walks the dog.’

In the event structure there are ranked participants. \text{IND} \ means ‘index’, which is a ‘pointer’ to an event participant, and \text{RANK} \ indicates the prominence of the participant concerned. In the case of (25), the \text{RANK1} \ participant is realized as the subject, while the \text{RANK2} \ and \text{RANK3} \ participants are realized as the object.
Tibor Laczkó

6.2 Estonian

Tamm (2006) develops an LFG analysis of the interaction of transitive telic verbs and aspectual case in Estonian. In this language the objects of telic verbs can bear either partitive (PAR) case or total (TOT) case. The choice between partitive and total is regulated by the aspectual features of the sentence, compare Tamm’s examples in (27) and (28).\[^{32}\]

(27) Estonian:
Mari kirjutas raamatu ühe aastaga.
Mari.NOM write.pst.3sg book.tot one.gen year.com
‘Mari wrote a/the book in a year.’

(28) Estonian:
Mari kirjutas raamatut terve aasta.
Mari.NOM write.pst.3sg book.par whole.tot year.tot
‘Mari was writing a/the book for a whole year.’

Tamm shows that the sentence in (27), with its object in total case, has a perfective interpretation, and the sentence in (28), with its object in partitive case, is imperfective, as is supported by the types of the adjuncts in them: ‘in a year’ vs. ‘for a year’. In addition, Tamm shows that Vendlerian achievement verbs like

\[^{32}\]The lexical entries for the Estonian case-markers encoding aspectual features are modeled as semantic (Butt & King 2004) and constructive cases (Nordlinger & Sadler 2004), and they provide the formal tools for Tamm’s analysis. On the terminology of Finnic core cases, see Tamm (2011a, 2012c). On partitives in Finnish, see Vainikka & Maling (1996).
võitma ‘defeat’ are compatible with objects in partitive case in Estonian, although the sentences they occur in are perfective by default, see her example in (29).

(29) Estonian:
Mari vōitis Jürit.
Mari.NOM defeat.PST.3SG George.PAR
‘Mary defeated George.’

In her analysis, Tamm introduces the boundedness aspectual feature: $B$ with two values: $\text{MIN}$ and $\text{MAX}$. She associates this feature both with the lexical forms of the two transitive verb types seen above and with the lexical representations of case markers in the following way. Her basic generalization is that ‘write’-type verbs are boundable, and ‘defeat’-type verbs are bounded. In the lexical form of the former boundedness is encoded as an existential constraint, while in the lexical form of the latter it is encoded as a defining equation: the $B$ feature has the $\text{MIN}$ value, see (30) and (31), respectively.

(30) kirjutama ‘write’... \(\uparrow B\)

(31) võitma ‘defeat’... \(\uparrow B\) = MIN

As regards case, the total case-marker, attached to an object noun phrase, introduces the $\text{MAX}$ value for $B$, while the partitive case-marker specifies $B$ as $\neq \text{MAX}$. These values are encoded with inside-out function application, see (32) and (33).

(32) \(\text{tot} \quad (\uparrow \text{CASE}) = \text{TOT} \)
\(\quad ((\text{OBJ} \uparrow) B) = \text{MAX} \)

(33) \(\text{par} \quad (\uparrow \text{CASE}) = \text{PAR} \)
\(\quad ((\text{OBJ} \uparrow) B) \neq \text{MAX} \)

In this system, a ‘write’-type verb requires that the sentence should be marked for boundedness, and its underspecified $B$ feature admits either of the two object cases. For instance, Tamm gives the following lexical representations for the verb and the object in (27).

(34) kirjutas \(V \quad (\uparrow \text{PRED}) = \text{WRITE}((\uparrow \text{SUBJ})(\uparrow \text{OBJ}))'\)
\(\quad (\uparrow \text{TENSE}) = \text{PST} \)
\(\quad (\uparrow \text{PERS}) = 3 \)
\(\quad (\uparrow \text{NUM}) = \text{SG} \)
\(\quad (\uparrow B) \)
(35) **raamatu** N  
\((↑ \text{PRED}) = \text{‘BOOK’}\)  
\((↑ \text{CASE}) = \text{TOT}\)  
\((\text{OBJ} ↑ \text{B}) = \text{MAX}\)

On the basis of this, her f-structure representation of (27) is as follows.

\[
\begin{array}{c}
\text{PRED} \ '\text{WRITE}⟨\text{SUBJ, OBJ}⟩' \\
\text{B} \ \text{MAX} \\
\text{TNS} \ \text{PST} \\
\text{NUM} \ \text{SG} \\
\text{PERS} \ 3 \\
\end{array}
\]

\[
\begin{array}{c}
\text{SUBJ} \ \\
\text{PRED} \ '\text{MARÍ}’ \\
\text{CASE} \ \text{NOM} \\
\ldots \\
\end{array}
\]

\[
\begin{array}{c}
\text{OBJ} \ \\
\text{PRED} \ '\text{BOOK}' \\
\text{CASE} \ \text{TOT} \\
\text{NUM} \ \text{SG} \\
\ldots \\
\end{array}
\]

Obviously, the f-structure representation of (28) would be different from (36) in one important respect: the value of B would be \(\neq \text{MAX}\) on the basis of (33).

By contrast, the value of the B feature of a ‘defeat’-type verb is \(\text{MIN}\), which only allows compatibility with an object in partitive case, given that total case encodes the opposite value: \(\text{MAX}\). Tamm offers the following lexical representations for the verb and the object in (29), and she points out that there is no value clash with respect to the B feature.

(37) **võitis** V  
\((↑ \text{PRED}) = \text{‘DEFEAT}⟨(↑ \text{SUBJ}) (↑ \text{OBJ})⟩’\)  
\((↑ \text{TNS}) = \text{PST}\)  
\((↑ \text{PERS}) = 3\)  
\((↑ \text{NUM}) = \text{SG}\)  
\((↑ \text{B}) = \text{MIN}\)

(38) **jürít** N  
\((↑ \text{PRED}) = \text{‘George’}\)  
\((↑ \text{CASE}) = \text{PAR}\)  
\((\text{OBJ} ↑ \text{B}) \neq \text{MAX}\)

As another argument-realization topic, Torn (2006) discusses the status of certain non-core arguments and adjuncts of verbal predicates in Estonian. She points out that fundamentally there are two approaches to these constituents. One of them regards non-core arguments as oblique case-marked indirect objects,
separating them from adjuncts, while the other lumps the two groups together as adverbials. Torn subscribes to the first approach.

By way of illustration, Torn shows that in this language participants of an event that are indirectly affected are realized by noun phrases bearing the same ‘local’ case suffixes as are used to express spatial adverbial dependents: see her examples in (39) and (40).

(39) Adverbial allative, Estonian:
Mees istus diivanile.
man.NOM sat sofa.ALL
‘A man sat onto the sofa.’

(40) Oblique allative, Estonian:
Ema andis lapsele raha.
mother.NOM gave child.ALL money.PAR
‘The mother gave money to the child.’

Torn says that diivanile ‘onto the sofa’, a noun phrase in allative case, is an un-governed adverbial constituent in (39), while lapsele ‘to the child’, a noun phrase in allative case here, too, expresses the indirectly affected argument of the ditransitive verb andma ‘give’ in (40). In her terminology, diivanile in (39) is an adverbial modifier, and lapsele in (40) is an object adverbial.

Torn offers the following three arguments for distinguishing object adverbials from adverbial modifiers. (i) A verbal predicate selects a particular governed case for its object adverbial and not a semantically compatible set of cases. (ii) An object adverbial constituent can serve as an antecedent in an obligatory control construction. (iii) It is a functional similarity between object adverbials on the one hand, and subjects and objects on the other, that they can be involved in systematic case alternations. Such alternations can never involve adverbial modifiers.

Torn adopts LFG’s LMT classification of governable grammatical functions. In this setting, she assumes that locative case-marked noun phrases can have either the OBL or the ADJUNCT function.
7 Morpho-syntactic phenomena

7.1 Agreement

7.1.1 Subject-verb agreement in Inari Saami and Finnish

Toivonen (2007) examines verbal inflectional morphology in Inari Saami. She develops her analysis by concentrating on the Saami copula leðe ‘to be’. In this language, as in various Northern Uralic languages, the number feature has three possible values: singular, dual and plural. It is another special property of this language that there can be either full agreement or partial agreement between the subject and the verb. Animate and specific subjects trigger the former, see (41), inanimate subjects trigger the latter, see (42).

(41) Inari Saami:
    Meecist lava uábbi já viljá.
    forest.LOC are.3du sister.NOM and brother.NOM
    ‘In the forest are my sister and brother.’

(42) Inari Saami:
    Riddoost láá kyehti keeðgi.
    beach.LOC are.3pl two rock
    ‘On the beach are two rocks.’

Subject noun phrases headed by unspecific human nouns and animal nouns can trigger either full or partial agreement. (43) illustrates the unspecific human case.

(43) Inari Saami:
    Táálust lava/láá kyehti ulmuu.
    house.LOC are.3du/are.3pl two person
    ‘There are two people in the house.’

Toivonen presents the paradigms of the copula in this three-way number and dual agreement system as in Table 4. She develops an LFG analysis with fully specified and underspecified lexical forms of verbal predicates. Consider her representations of four morphological forms of the copula in (44–47).

(44) lava V
    (↑ PRED) = ‘BE’
    (↑ TENSE) = PRS
    (↑ MOOD) = INDICATIVE
    (↑ SUBJ NUM) = DU
    (↑ SUBJ PERS) = 3
    (↑ SUBJ HUM) = +

xxviii
Table 4: Agreement paradigms for ‘to be’

<table>
<thead>
<tr>
<th></th>
<th>full</th>
<th>partial</th>
</tr>
</thead>
<tbody>
<tr>
<td>SG</td>
<td>lam</td>
<td>lii</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>lah</td>
<td>lii</td>
</tr>
<tr>
<td>3</td>
<td>lii</td>
<td>lii</td>
</tr>
<tr>
<td>DU</td>
<td>láán</td>
<td>láá</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>leppee</td>
<td>láá</td>
</tr>
<tr>
<td>3</td>
<td>lava</td>
<td>láá</td>
</tr>
<tr>
<td>PL</td>
<td>lep</td>
<td>láá</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>leppeδ</td>
<td>láá</td>
</tr>
<tr>
<td>3</td>
<td>láá</td>
<td>láá</td>
</tr>
</tbody>
</table>

(45) lam V
  (↑ PRED) = ‘BE’  
  (↑ TENSE) =PRS  
  (↑ MOOD) = INDICATIVE  
  (↑ SUBJ NUM) = SG  
  (↑ SUBJ PERS) = 1  
  (↑ SUBJ HUM) = +  

(46) lii V
  (↑ PRED) = ‘BE’  
  (↑ TENSE) =PRS  
  (↑ MOOD) = INDICATIVE  
  (↑ SUBJ NUM) = SG  

(47) láá V
  (↑ PRED) = ‘BE’  
  (↑ TENSE) =PRS  
  (↑ MOOD) = INDICATIVE  

Toivonen makes crucial use of the principle of morphological blocking as developed by Andrews (1990). The basic idea is that if a subject noun phrase is compatible with more than one verb form, it will select the variant that exhibits the largest number of its own feature values. This explains, for instance, why human subjects do not freely co-occur with láá or why singular subjects cannot co-occur with láá. The answer to the first question is that láá has no [+human] feature, see (47). The answer to the second question is that there are more specific forms of the copula in that they also encode the [+singular] feature, compare (47) with (45) and (46).
Tibor Laczkó

Toivonen also briefly compares the Inari Saami agreement system with the corresponding Finnish system. She points out that Finnish has no grammatical dual. In addition, Finnish does not exhibit partial agreement. Furthermore, animacy has not been grammaticalized in standard Finnish. It is another significant difference that in Inari Saami, verb agreement is always triggered by grammatical subjects, while in Finnish several independent conditions need to be simultaneously satisfied for agreement to take place. First, in Finnish, as well as in Estonian, only nominative NPs trigger agreement, compare Toivonen’s examples in (48) and (49).

(48) Finnish:
Autot ajavat yleensä kovaa moottoriteillä.
cars.NOM drive.3PL generally hard motorways.ADE
‘Cars generally drive fast on the motorways.’

(49) Finnish:
Linja-autoja kulkee nykyisin joka sunnuntai.
buses.PAR run.3SG nowadays every Sunday
‘Nowadays, buses run every Sunday.’

In (48) the nominative subject triggers agreement, while in (49) the subject is in partitive case and the verb takes 3SG default agreement.

A Finnish verb also has default agreement in existential and possessive constructions. (50) illustrates the latter type.

(50) Finnish:
Koulussa on uudet opettajat.
school.INE is.3SG new.NOM.PL teachers.NOM
‘The school has new teachers.’

In this example, although the (post-verbal) subject is nominative, it is not in its preverbal canonical position; therefore, here, too, the verb displays 3SG default agreement.

As regards their possessive constructions, Inari Saami and Finnish differ in two significant respects. On the one hand, in Inari Saami possessive constructions pronouns are in nominative case, while in Finnish the corresponding pronouns take accusative case, compare (51) and (52). On the other hand, the possessor is always in nominative case in Inari Saami, it has the subject function, and it always triggers agreement, while in Finnish the possessor is either in nominative

33See Hiietam (2003), for instance.
case (ordinary noun phrases) or in accusative case, and the verb always carries 3sg default agreement, compare (53) and (54).

(51) Inari Saami:
Muste lah tun.
I.LOC are.2SG you.NOM.SG
‘I have you.’

(52) Finnish:
Minulla on / *olen sinut.
I.ADE is.3SG / is.1SG you.ACC.SG
‘I have you.’

(53) Inari Saami:
Muste lava puásui já peenuv.
I.LOC are.3DU reindeer.NOM and dog.NOM
‘I have a reindeer and a dog.’

(54) Finnish:
Minulla on / *olen poro ja koira.
I.ADE is.3SG / is.1SG reindeer.NOM and dog.NOM
‘I have a reindeer and a dog.’

Toivonen makes the following concluding generalization about Finnish possessive constructions. There is no normal agreement in them, because the possessor is not the subject, and because the subject possessor is not in nominative case. This is why 3sg default agreement is employed.

Toivonen offers a comparative overview of the agreement systems of Inari Saami and Finnish shown in Table 5.

### 7.1.2 Aspects of differential object marking in Uralic

Coppock & Wechsler (2010) point out that there is object agreement in Nenets, Enets, Nganasan and Selkup in the Samoyedic family and in Mordvinian (Finno-Volgaic), Hungarian (Ugric), Ostyak and Vogul (both Ob-Ugric) in the Finno-Ugric family. These languages exhibit remarkable variation with respect to the feature specifications of their object agreement. In Hungarian and Samoyedic there are two conjugation paradigms: subjective and objective, and the latter is used in the case of definite and third person objects. In Ob-Ugric languages there is a subjective conjugation and three objective conjugation paradigms, one
for each possible number value of the object (singular, dual and plural). In Mordvinian there is genuine agreement for both person and number between the verb and the object. Coppock & Wechsler (2010) concentrate on Northern and Eastern Ostyak, Hungarian and Samoyedic.34

In Northern Ostyak the verb agrees with its object in number but not in person: see (55) and (56). An additional factor is that the object has to be topical, otherwise the subjective conjugation is used.

(55) Northern Ostyak:
Ma täm kälang wel-sə-l-am.
I this reindeer kill-pst-plObj-1sgSubj
‘I killed these reindeer.’

(56) Northern Ostyak:
Xũnsi năng mũng-iluw xälśa want-lə-l-an?
when you we-acc where see-prs-plObj-2sgSubj
‘When did you see us where?’

Coppock & Wechsler (2010) postulate the following diachronic analysis of these facts.

At the first stage third person pronouns were incorporated (↓ pred)=‘pro’ and (↓ index pers)=3 with the three number values (↓ index num)=n. This was combined with the topicality condition: (↓ σ df)=topic.

34 Also see Coppock & Wechsler (2012) on Hungarian.
At the second stage the pred ‘pro’ was dropped.

The authors claim that it is reasonable to assume that at this stage person specification was present because Eastern Ostyak still manifests this stage.

At the third stage the person specification was lost in Northern Ostyak, see (59), but this did not happen in Eastern Ostyak.

As a result, these objective conjugation suffixes became usable with first and second person objects, too.

Coppock & Wechsler (2010) also show that Hungarian has two conjugations that are conditioned by the definiteness of the object by using the following examples. The general pattern is that definite objects trigger the objective agreement type, see (60), and indefinite objects require the subjective type, see (61).

(60) Hungarian:
Lát-om a madar-at.
see-prs.1sg.def the bird-acc
‘I see the bird.’

(61) Hungarian:
Lát-ok egy madar-at.
see-prs.1sg.indf a bird-acc
‘I see a bird.’
In addition, the objective agreement type is sensitive to the person feature of the object: in the pronominal domain only third person pronouns trigger it, see (62), while first and second person pronouns require the subjective conjugation, see (63). Coppock & Wechsler (2010) refer to this as the third person restriction in this language.

(62) Hungarian:
Lát-ják őt/őket.
see-PRS.3PL.DEF it/them
‘They see it/them.’

(63) Hungarian:
Lát-nak engem/téged/minket.
see-PRS.3PL.INDF me/you/us
‘They see me/you/us.’

It is another property of the Hungarian object agreement system that it is not sensitive to the number value of the object.

Coppock & Wechsler (2010) propose the following diachronic analysis. At the first stage, just like in the case of Northern and Eastern Ostyak, third person pronoun incorporation took place, see (57) above. The second stage was also the same: the pred ‘pro’ was dropped and the topicality condition retained, see (58). This is the present-day Eastern Ostyak system. At the third stage the number constraint was dropped, but the person restriction was retained, see (64) and compare it with (59) characterizing Northern Ostyak.

(64) $V_{aff}$
\[
(\uparrow OBJ) = \downarrow \\
(\downarrow PRED) = ‘PRO’ \\
(\downarrow_\sigma DF) = \text{TOPIC} \\
(\downarrow INDEX PERS) = 3 \\
(\downarrow INDEX NUM) = N \quad \text{where } N \in \{\text{SG, DU, PL}\}
\]

Finally, at the fourth stage the topicality constraint was reanalyzed as a definiteness constraint, see (65).

(65) $V_{aff}$
\[
(\uparrow OBJ) = \downarrow \\
(\downarrow PRED) = ‘PRO’ \\
(\downarrow_\sigma DF) = \text{TOPIC} (\uparrow OBJ DEF) = c + \\
(\downarrow INDEX PERS) = 3 \\
(\downarrow INDEX NUM) = N \quad \text{where } N \in \{\text{SG, DU, PL}\}
\]
Dalrymple & Nikolaeva (2011) investigate differential object marking (DOM) by exploring syntactic, semantic and informational structural differences between marked and unmarked objects in a wide range of genetically and typologically different languages. As regards Uralic, they concentrate on Tundra Nenets in the Samoyedic subfamily and on Ostyak (Khanty), Vogul (Mansi) and Hungarian in the Finno-Ugric subfamily.\(^{35}\)

Dalrymple & Nikolaeva (2011) develop a formal theory of information structure and its place in the architecture of LFG. In this theory information structure is closely related to semantic structure. It is a favourable aspect of this approach that it makes possible a simple specification of the informational structural status of an argument by providing a DF feature value in its semantic structure.

In Tundra Nenets there is only a single object function: OBJ. First and second person (pronominal) objects do not agree with the verb, just like in Hungarian, see (63) above. Third person objects optionally agree with the verb. If there is agreement, the object has the TOPIC DF, while no such function is associated with it in the absence of agreement. Dalrymple & Nikolaeva (2011) model this in the following way.

\[
Agreement \text{ with third person topical objects:} \\
(\uparrow \text{OBJ} \text{ PERS}) = 3 \\
((\uparrow \text{OBJ})_\sigma \text{ DF}) = \text{TOPIC}
\]

This specification encodes that the semantic structure contributed by the third person object is associated with the topic role in information structure.

Dalrymple & Nikolaeva (2011) also distinguish a language type in which there are two object functions: OBJ and OBJ\(\theta\). They claim that Ostyak belongs to this type, in addition to Mongolian, Chatino and Hindi, among others. The OBJ\(\theta\) function in these languages is only available to patient/theme arguments. Dalrymple & Nikolaeva (2011) make the following empirical generalizations. Although Ostyak has two object functions, they cannot co-occur in a sentence, because this language does not have a double object construction. In the case of verbs such as ‘give’ there are the following two possibilities: either the goal or the theme must have an OBL function. When the goal has a dative oblique function, the theme has two object choices. If it is topical, it has the agreeing OBJ function, and if it is not topical, it has the non-agreeing OBJ\(\text{ THEME} \) function.

Dalrymple & Nikolaeva (2011) compare the Nenets type and the Ostyak type of DOM in the following way.

\(^{35}\)On aspecual DOM in Estonian, see the discussion of Tamm (2006) in Section 6.2.
Dalrymple & Nikolaeva (2011) also point out that in the Ob-Ugric branch of Finno-Ugric languages Vogul follows the same DOM pattern as Ostyak: object marking is information structure driven: topicalization by means of object agreement. The authors hypothesize that this also held for Proto-Ob-Ugric. There are no attested semantic restrictions on agreeing objects in Ob-Ugric. As shown above, object agreement works differently in Hungarian. First and second person pronouns never trigger agreement, just like in Tundra Nenets, see above. Third person object agreement is not regulated by information structure: it is triggered by definiteness. It is only definite third person objects that trigger agreement irrespective of their discourse function status.

The authors suggest that earlier Hungarian was closer to Ostyak and Vogul, and in modern Hungarian definiteness marking is an innovation, after the development of the grammatical category of definiteness and the appearance of grammatical articles. Their reconstruction of the relevant linguistic historical processes is as follows. They assume that the Ob-Ugric system of DOM, which is exclusively based on information structure, is the most archaic type, and probably it can be hypothesized for Proto-Eastern-Uralic, i.e. the Proto-Uralic dialects from which the Samoyedic and Ugric languages developed. At a later stage, agreement became reduced to third person topical objects in Samoyedic and Proto-Hungarian as a consequence of the fact that third person was frequently associated with secondary topicality. By contrast, first and second person pronouns occupy the highest position on a scale of topic-worthiness. Dalrymple & Nikolaeva (2011) suggest that the Samoyedic languages (Nenets, Selkup and Nganasan) and Old Hungarian grammaticalized the tendency that first and second person
pronouns are likely primary topics and unlikely secondary topics. Thus, they cannot correspond to the primary object, given that in these languages it tends to be strongly associated with secondary topic. No such restrictions hold for third person objects. Hungarian and (possibly) Selkup represent the next historical stage, at which the grammatical marking of third person topical objects is extended to non-topical definite objects. According to Dalrymple & Nikolaeva (2011) this change manifests the spreading of grammatical marking to non-topical objects that exhibit topic-worthy features with the concomitant loss of relatedness to information structure.\footnote{Dalrymple & Nikolaeva make the following footnote comment. ‘An alternative explanation was recently suggested by Coppock & Wechsler (2010), who claim that object agreement in proto-Uralic was initially restricted to third person topical objects. It later spread to all topical objects in Northern Ostyak and Vogul, whereas Samoyedic languages preserve the original situation. This suggestion provides an elegant analysis of feature loss as a mechanism of historical change: Northern Ostyak lost the specification that restricted topical agreement to third person objects (the (↑ obj pers)=3 specification for agreeing verbs). However, the causal mechanism of this development remains unclear: it presupposes the spread of marking to unlikely contexts’ (Dalrymple & Nikolaeva 2011: 201).}

This section on DOM has shown how complex these phenomena are in Uralic languages in general and in Finno-Ugric languages in particular. It has also demonstrated that LFG’s well-developed modular architecture provides the necessary and appropriate formal devices to capture both the synchronic differences between languages and the diachronic processes in a principled manner.

### 7.2 Evidentiality

Asudeh & Toivonen (2017) propose a modular LFG approach to evidentiality, which is a well-established morpho-syntactic category in a considerable number of languages, for instance, Tariana, Cherokee, Cheyenne, Quechua and Tuyuca. These languages employ fully grammaticalized evidentiality morphology, which encodes the source and reliability of speakers’ knowledge. Other languages, e.g. English, do not have such evidentiality marking, and they use alternative means to express sources of evidence or degrees of certainty about evidence (apparently, I saw that..., etc.). For the description of grammaticalized evidentiality they use the following f-structure features: [DIRECT ±], [VISUAL ±], [REPORTED ±], which also express semantic content to be captured as modifiers on events in Glue Semantics. In languages like English (with non-grammaticalized evidentiality) predicates like sound and seem optionally encode evidentiality information for the semantic component of the theory. The authors argue that LFG’s modular architecture is especially well-suited to capturing the systematic similarities and
differences between grammaticalized and non-grammaticalized ways of expressing evidentiality across languages.

Szabó (2021) points out that in the family of Uralic languages both evidentiality systems can be found. For instance, the Finnic, the Saamic and the Mordvinian languages and Hungarian do not have grammaticalized evidentiality. By contrast, Estonian, Livonian, Mari, Komi, Udmurt as well as the entire Ob-Ugric and Samoyedic branches employ grammaticalized evidentiality.

Szabó (2021) sketches an LFG approach to grammaticalized evidentiality in Udmurt. She shows that there are two past tense paradigms in this language, and the 2nd past is used to express the source of information, among other aspects of morpho-syntax. Therefore, this verb form is multiply ambiguous. Szabó (2021: 82) captures this by proposing that the 2nd past contributes the following attribute-value pair to the f-structure of a sentence.37

(69) \[
\text{SOURCE RES} \lor \text{PFV} \lor \text{HEAR} \lor \text{FOLK} \lor \text{MIR} \lor \text{INFER} \lor \text{NON-V}
\]

As (69) shows, in this domain the f-structure is multiply ambiguous with all these disjunctive values for SOURCE, and the assumption is that it is basically the context that disambiguates.

Tamm (2008) shows that in Estonian partitive case-marking has either epistemic modality or aspectual use. In the former, it encodes incomplete evidence (cf. grammaticalized evidentiality marking), and in the latter, it presents an event as incomplete. The lack of partitive-marking indicates complete evidence and complete event, respectively. In this language both verbs and object arguments can be marked for partitive. Tamm proposes the lexical form in (70) for the aspectual partitive case marker on the object, and the lexical forms in (71) and (72) for the impersonal and personal evidentiality markers on verbs, respectively.

(70) \[
(\uparrow \text{CASE}) = \text{PARTITIVE} \\
((\text{OBJ} \uparrow) \text{EVENT}) \neq \text{COMPLETE}
\]

(71) \[
[-\text{ta-vat}] \\
(\uparrow \text{FORM}) = \text{PARTITIVE EVIDENTIAL} \\
(\uparrow \text{MODE OF COMMUNICATION}) = \text{INDIRECT} \\
(\uparrow \text{EVIDENCE}) \neq \text{COMPLETE} \\
(\uparrow \text{VOICE}) = \text{IMPERSONAL}
\]

37Where res = resultative, pfv = perfective, hear = hearsay, folk = folklore, mir = mirative, infer = inferential, non-v = non-volitional.
32 LFG and Finno-Ugric languages

(72) [-va-t]  (↑ FORM) = PARTITIVE EVIDENTIAL
       (↑ MODE OF COMMUNICATION) = INDIRECT
       (↑ EVIDENCE) ≠ COMPLETE
       (↑ VOICE) = PERSONAL

Tamm sketches a Discourse Representation Theory-based semantic description associated with the f-structure representation.

For further discussions and analyses of evidentiality, see Szabó (2017) on Udmurt, and Tamm (2004c, 2012a) on Estonian. On partitives, also see Tamm (2012b).

8 Noun phrase phenomena in Hungarian

8.1 C-structure issues

As we show below, Hungarian noun phrases have been analyzed as either NPs or DPs in LFG approaches. Both views are fully legitimate in this framework, given that the standard LFG inventory of functional categories contains D (in addition to I and C).\(^{38}\) It is a crucial property of possessive noun phrases in this language that the possessor can be expressed in either nominative or dative case, and the two variants occupy distinct syntactic positions. Despite this fact, only one of them can occur in any single possessive noun phrase, that is they are in complementary distribution, as opposed to the possible co-occurrence of 's and of possessors in English.

Chisarik & Payne (2003: 189) use an NP approach to the representation of Hungarian and English noun phrases, see the structures they assume for (73) and (75)\(^{39}\) in (74) and (76), respectively.

(73) Hungarian:
    a király-nak a lány-a
    the king-DAT the daughter-POSS.3SG
    ‘the king’s daughter’

\(^{38}\)It is not unusual to find alternative categorial analyses of the same construction types in LFG. For instance Bresnan (2001) treats finite English sentences that contain no auxiliaries (e.g. Mary opened the door) as having the category S, while Dalrymple (2001) employs an IP approach.

\(^{39}\)Notice that Hungarian possessive noun phrases belong to the head-marking type.
They provide the following justifications for these representations. On the one hand, the dative possessor, see (74), can function as a predeterminer to coordinated NPs as in their example in (77).

(77) Hungarian:
   a király-nak [NP [ a fi-a és [ a lánya-a ]]]
   the king-DAT the son-Poss.3SG and the daughter-Poss.3SG
   ‘the king’s son and daughter’

On the other hand, the nominative possessor stands in complementary distribution with the definite article, just like the ‘s possessor in English.

The following remarks can be made on this approach. First, the coordination facts can also be captured in a DP analysis in which the dative possessor is in Spec,DP and Chisarik and Payne’s NP is a D’, where the definite article is the D head and the other constituent is (the head of) an NP.40 Second, it would

40 See Laczkó’s (1995) DP structure in (80) below.
need some justification to assume that a word-level functional category (D) is in complementary distribution with a phrasal category (NP).\footnote{It seems to be a further minor complication that the functional category D is used in an unusual way: it does not head and project a DP.} Third, in the case of pronominal nominative possessors there is no complementary distribution with the definite article; moreover, they must co-occur, compare (78) and (79).\footnote{(78) shows the grammaticality properties of this construction type in standard Hungarian. However, Szabolcsi (1994) documents a dialectal version in which even such non-pronominal nominative possessor constructions follow the pattern exemplified in (79).}

\begin{align*}
\text{(78)} & \quad \text{Hungarian:} \\
& \quad \begin{array}{l}
\text{(*a) János lány-a} \\
\text{the John.NOM daughter-POSS.3SG} \\
\text{‘John’s daughter’}
\end{array} \\
\text{(79) & \quad \text{Hungarian:} \\
& \quad \begin{array}{l}
\text{* (az) ő lány-a} \\
\text{the he.NOM daughter-POSS.3SG} \\
\text{‘his daughter’}
\end{array}
\end{align*}

Motivated by Szabolcsi’s (1994) seminal GB analysis, Laczkó in Laczkó (1995) and all subsequent work adapts a DP approach.\footnote{Without adopting theory-specific details like moving the nominative possessor from Spec,NP to Spec,DP, where it acquires dative case, as in Szabolcsi’s (1994) GB analysis.} The essential aspects of his structural representation of (73) and (75) would be as in (80) and (81), respectively.

\begin{align*}
\text{(80) & \quad \text{DP} \\
& \quad \begin{array}{c}
\text{DP D’} \\
\end{array} \\
& \quad \begin{array}{c}
\text{D’ D NP} \\
\end{array} \\
& \quad \begin{array}{c}
\text{D NP a N} \\
\end{array} \\
& \quad \begin{array}{c}
a N lánya \\
\end{array} \\
& \quad \begin{array}{c}
királynak
\end{array}
\end{align*}
This approach avoids the complications mentioned in connection with Chisarik & Payne’s (2003) NP analysis.

8.2 Event nominalization

8.2.1 Argument structure inheritance

Following Grimshaw (1990) and Szabolcsi (1994), among others, Laczkó in Laczkó (1995) and in all relevant subsequent work assumes that complex event nominals (cens) derived by the -ás/-és suffix (henceforth: Ás suffix) inherit the argument structure of the input verb, as opposed to simple event nouns (sens) and result nouns (reses). The most important properties of Hungarian cens are as follows; see also Laczkó (2000b, 2003, 2009a).

When an Ás noun has both a simplex form and a complex form containing a perfectivizing preverb, the latter is always a cen and the former is very often ambiguous: cen vs. sen and/or res. Compare the examples in (82).

(82) Hungarian:

a. Anna vizsgátat-ás-a
    Anne.NOM examine-ás-poss.3sg
    ‘Anne’s examination’
    cen: Anne = patient
    sen: Anne = examiner or examinee
b. Anna le-vizsgáztat-ás-a (a professzor által)
   Anne.NOM PFV-examine-ÁS-POSS.3SG the professor by
   'the examination of Anne (by the professor)'
   CEN: Anne = patient

   c. Anna vizsgá-ja
   Anne.NOM exam-POSS.3SG
   'Anne’s exam'
   SEN: Anne = examiner or examinee

(82a) contains a derived nominal without a perfectivizing preverb, and it can be used as either a CEN with an argument structure or as a SEN without an argument structure (with only a lexical conceptual structure). In the former use Anna is interpreted as the patient argument of the nominal predicate, in the latter use it is interpreted as a participant in an examination situation, whether the examiner or the examinee. By contrast, in (82b) the derived nominal contains a perfectivizing preverb, and it can only be analysed as a CEN with obligatory argument structure and Anna must be interpreted as the patient argument. In (82c) the head is an underived noun and it can only be a SEN.

The expression of the arguments of the derived nominal predicate is as obligatory as in the case of the input verb.

(83) Hungarian:
   A vizsgáztat-ás két órá-ig tart-ott.
   the examine-ÁS-POSS.3SG two hour-for last-PST.3SG
   'The examination lasted for two hours.' (SEN)

(84) Hungarian:
   *A le-vizsgáztat-ás két órá-ig tart-ott.
   the PFV-examine-ÁS-POSS.3SG two hour-for last-PST.3SG
   'The examination lasted for two hours.' (CEN)

As (83) shows, when no complement is present, an otherwise ambiguous (CEN/SEN) nominal must be interpreted as a SEN. (84) demonstrates that an ‘only CEN’ nominal cannot occur without its obligatory internal argument(s). The external argument can be suppressed optionally, see (82b) above.

CENS cannot be pluralized, see (85).
When an adjunct in the DP with a derived nominal head is expressed by a postpositional phrase, this PP has to be ‘adjectivalized’ either by combining it with a formative element, one of the present participial forms of the copula: való ‘being’, glossed as való, or by attaching the adjectivizing suffix -i (glossed as AFF) to the postposition. In such cases, the való version is only compatible with the cen reading of an otherwise ambiguous nominal predicate, while the -i variant retains the ambiguity, cf. (86a) and (86b). This is Szabolcsi’s (1994) famous való-test for unambiguously identifying cens in Hungarian.⁴⁴

The core arguments of cens can receive a variety of [−r] gfs in several LFG approaches to Hungarian, see Section 8.2.2. Non-core arguments are typically expressed by case-marked DPs and postpositional phrases, and they are mapped onto obl functions. Adjuncts can also be expressed by case-marked DPs and PPs. In addition, they can be realized by APs, especially when the input verb would take an AdvP for the same kind of modification, e.g. váratlan-ul ‘unexpected-ly’ (Adv) vs. váratlan ‘unexpected’ (A). For empirical generalizations about the major (structural and categorial) ways of realizing obl and adjunct functions in cen constructions and LFG analyses, see Laczkó (1995, 2003).

The Hungarian event nominalization phenomena presented above are relevant for theorizing in generative grammar in general and in LFG in particular for the following reasons. Grimshaw’s (1990) influential proposal substantially distinguishing cens from sens and reeses is based on English data, primarily on -tion nominalization. In this language, however, these derived nouns are genuinely

⁴⁴ Also see Laczkó & Rákosi (2007).
ambiguous and, therefore, it is often difficult to employ Grimshaw’s diagnostics, e.g. (non-)pluralizability, to definitely tell the cen and sen uses apart. Due to this fact, Grimshaw’s theory has been criticized from a variety theoretical perspectives, see Laczkó (2000b) and the references therein. By contrast, in Hungarian there are clear morphological and syntactic indicators, and the diagnostics can be applied reliably and unambiguously. This situation has motivated some LFG practitioners to investigate event nominalization thoroughly and, among other things, to develop various LMT analyses of argument realization in this domain, see Section 8.2.2.

8.2.2 Functional issues

A variety of inventories of gfs in Hungarian DPs with cen heads and a consequent variety of LMT analyses have been proposed, see Table 6.  

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DP_DAT</td>
<td>POSS</td>
<td>SUBJ</td>
<td>SUBJ/POSS</td>
</tr>
<tr>
<td>DP_NOM</td>
<td>POSS</td>
<td>NCOMP</td>
<td>SUBJ/POSS</td>
</tr>
<tr>
<td>DP_OBL/PP</td>
<td>OBL</td>
<td>OBL</td>
<td>OBL</td>
</tr>
</tbody>
</table>

Laczkó (1995) uses gfs standardly employed in noun phrases (POSS and OBL). Assuming that POSS is a semantically unrestricted function, he develops an LMT approach in which there is a poss Condition that is the nominal domain counterpart of the subj Condition in the verbal domain. The subj Condition requires that every (verbal) predicator must have a Subject, see Bresnan (1990), for instance. Laczkó’s (1995: 85) poss Condition states: ‘every event nominal predicator must have a Possessor’.

Rather exceptionally in the generative literature on Hungarian noun phrases, Chisarik & Payne (2003) assume that the two possessor constituents bear distinct gfs, both of which are taken to be semantically unrestricted. The dative realizes the subj function in the nominal domain, while the nominative expresses a new, DP-specific function: NCOMP. SUBJ is considered to be discourse-related, while NCOMP is not.

Charters (2014) proposes a new df in Hungarian possessive DPs: anchor.
Laczkó (2004) assumes that both the dative possessor and the nominative possessor can overtly realize either the subj or the poss gfs, both of which are regarded as semantically unrestricted. Furthermore, the subj argument can also be expressed by an LFG-style PRO. Given this nature and distribution of these gfs, Laczkó’s LMT analysis can adopt the subj Condition from the verbal domain. In addition, his approach can formally handle (anaphoric) control into possessive DPs in Hungarian with the standard LFG mechanism even in the case of cens derived from transitive verbs, which Laczkó’s (1995) system cannot do. Consider the following examples.

(87) Hungarian:

a. Péter elkezdte a kiabál-ás-t.
   Peter.nom started the shout-ÁS-ACC
   ‘Peter started the shouting.’

b. Péter elkezdte a dal énekl-és-é-t.
   Peter.nom started the song.nom sing-ÁS-POSS.3SG-ACC
   ‘Peter started the singing of the song.’

In Laczkó’s (1995) system, the f-structure of the DP in (87a) contains a POSS PRO, which is anaphorically controlled by the matrix subject, and in (87b) a dal ‘the song’ has the POSS function, and (in the absence of any other available gf for the agent controllee) Laczkó is forced to assume that control takes place in a different dimension. By contrast, in Laczkó’s (2004) approach there is a controlled PRO subj in both cases, and in (87b) a dal ‘the song’ has the POSS function. Laczkó’s (2004) subj & poss theory receives further independent support from Laczkó & Rákosi (2019), who argue that this gf inventory is necessary for the adequate LFG handling of certain binding facts in Hungarian DPs. Laczkó (2008b, 2009b), in response to Kenesei (2005), proposes that both T participial constructions and CEN constructions should have a dual PRO & suppression analysis for an adequate treatment of binding and control phenomena.

8.3 Possessives

8.3.1 Finnish

Toivonen (2000) develops an analysis of the morpho-syntax of Finnish possessive noun phrases. This language has the widely attested poss pro-drop in the case of first and second person possessors, see a 1SG example in (88), and Toivonen’s lexical representation of the pronoun and the possessive suffix (glossed as POSS) in (89) and (90), respectively.
In the third person there is an interesting split between the possessive pronoun and the possessive suffix when the latter provides the PRED feature (i.e. in the case of pro-drop). The pronoun must not be bound by the matrix subject, while the POSS-PRO must, cf. (91) and (92).

Furthermore, the 3sg.poss suffix cannot agree with a non-human possessor:

Toivonen captures these facts by means of the following lexical forms.\(^{46}\)

\(^{46}\)SB stands for obligatorily subject bound.
Toivonen also compares corresponding possessive noun phrase constructions in Estonian and Northern Saami. Toivonen (2001) provides a historical context for her analysis in Toivonen (2000), and she also discusses dialectal variation in Finnish with respect to these phenomena. Her proposal involves the erosion of features other than \textsc{pred} ‘pro’, which makes it very similar to Coppock & Wechsler’s (2010) analysis of Ostyak and Hungarian in Section 7.1.2.

8.3.2 Hungarian

Laczkó (2001) develops an LFG approach to the inflectional phenomena in Hungarian possessive DPs in the spirit of Item and Arrangement morphology.\footnote{By contrast, Laczkó (2018) proposes a Word and Paradigm approach, arguing that it has considerable implementational advantages.} Consider the following examples.

(97) Hungarian:

\begin{enumerate}
\item a. \textit{toll-a-i-nk}
  \begin{itemize}
  \item the pen-poss-pl-1pl
  \item ‘our pens’
  \end{itemize}
\item b. \textit{toll-a-i}
  \begin{itemize}
  \item the pen-poss-pl.3sg
  \item ‘her pens’
  \end{itemize}
\item c. \textit{toll-a}
  \begin{itemize}
  \item the pen-poss.3sg
  \item ‘her pen’
  \end{itemize}
\item d. \textit{hajó-i}
  \begin{itemize}
  \item the ship-poss.pl.3sg
  \item ‘her ships’
  \end{itemize}
\end{enumerate}

Laczkó postulates the following sets of functional annotations in the lexical forms of -\textit{a} and -\textit{i}, the main point being that the same morphological form (morph) can encode fewer or more features depending on what other morphs it is combined with, see the optional features in (98).
9 Further reading

Limitations of space have prevented us from discussing additional phenomena in Finno-Ugric languages and their analyses. Below we provide references to further works that we recommend to the interested reader.


48 For instance, see Dalrymple & Lødrup (2000) and Lødrup (2012) for COMP, and Alsina et al. (2005) and Patejuk & Przepiórkowski (2016) against COMP, and the references in these papers.
Tibor Laczkó


10 Conclusion

In this chapter we have discussed some salient, sometimes competing, LFG analyses of a variety of (morpho-)syntactic phenomena in Finno-Ugric languages, with occasional glimpses at alternative generative approaches, on the one hand, and at some related phenomena in languages belonging to Samoyedic, the other major branch of Uralic languages, on the other hand. We have dealt with clausal c-structure representational issues, verbal modifiers, focused constituents, negation, copula constructions, argument realization, subject-verb agreement, differential object marking, evidentiality and a set of noun phrase phenomena related to event nominalization.

On the basis of the interim conclusions at the end of various sections, we can make the following overall concluding remarks at the end of this chapter. On the one hand, LFG provides an appropriate and suitably flexible formal apparatus for a principled analysis of all the phenomena in all the Finno-Ugric languages discussed here. The range of these phenomena is considerably wide and varied, see above, containing several cases that pose serious challenges for generative grammar at large, for instance, the treatment of complex predicates, negation, copula constructions, discourse functions, agreement and event nominalization. On the other hand, the analysis of some of these phenomena can also contribute to LFG-internal theorizing, see, for instance, the choice between LFG treatments of complex predicates involving PIDs and clause negation.
Acknowledgements

Work on this chapter was supported by a research grant obtained from the Faculty of Humanities and Social Sciences, Károli Gáspár University of the Reformed Church in Hungary (Theoretical and Experimental Research in Linguistics, reg. no. 20736B800). We are grateful to our three anonymous reviewers, who at a later stage identified themselves as Anne Tamm, Ida Toivonen and Péter Szűcs. The comments of all our reviewers have greatly contributed to improving both the content and the presentational aspects of this chapter. Our special thanks go to Anne Tamm for generously providing us with a large amount of information about phenomena in several Uralic languages that are comparable to the phenomena whose LFG analyses are discussed in this chapter and for calling our attention to a great number of additional relevant works on Uralic languages. As usual, all remaining errors are our sole responsibility.

Abbreviations

Besides the abbreviations from the Leipzig Glossing Conventions, this chapter uses the following abbreviations.

| ADE | adessive case (marker) |
| ARA | Estonian particle |
| ÁS | Hungarian event nominalizer suffix |
| INE | inessive case |
| PAR | partitive case |
| TOT | total case |
| VALÓ | Hungarian adjectivalizing participle |
| VM | verbal modifier |

References


Sahkai, Heete & Anne Tamm. 2018a. Estonian V2 is prosodic? Poster presentation at the LFG ’18 Conference.


Tamm, Anne. 2004a. Eesti ja ungari keele verbiaspekti modelleerimise probleeme [Problems of modeling Estonian and Hungarian verbal aspect]. In Marju Ilves & János Pusztay (eds.), *Folia estonica tomus xi. Észt-magyar összevetés*


Tamm, Anne. 2012a. Finnségi tárgyeset mint a kategóriákat áthidaló eset egyik altípusa: Az argumentumon megjelenő aspektusjelölő toldalék [Accusative in Finnic as a subtype of cases across categories: The aspect marker appearing on arguments]. In Gábor Alberti, Judit Kleiber & Judit Farkas (eds.), *Vonzásban és


Tibor Laczkó

