# Operator Fronting in Hungarian

Péter Szűcs

University of Debrecen

**Abstract**

This paper investigates the Hungarian Operator Fronting construction, where an element from a subordinate sentence appears in the matrix clause. Two subtypes are distinguished: a standard long-distance dependency and a proleptic version. In the latter, the fronted element is analyzed as a thematic object of the main verb. This thematic object is linked to an embedded grammatical function via obligatory anaphoric binding. The configuration can be integrated into a wider perspective of control from an LFG perspective.

## 1. Introduction[1]

Operator Fronting (OF) in Hungarian is a construction whereby some dependent of a subordinate clause surfaces in the matrix clause. The term itself is aimed to be a theory- and analysis-neutral label and is not to be viewed as a commitment to a particular view of the phenomenon. In the previous literature, the structure has also been called "Focus Raising" (Kenesei 1992, Lipták 1998, Gervain 2009), "Operator Raising" (Gervain 2002) and "Long Operator Movement" (É. Kiss 2002). Here "operator" refers to the fact that the element under question usually bears some discourse function in the main clause while "fronting" captures the noncanonical positioning of this element.

OF is illustrated in (1). (1a) is a standard Hungarian sentence with a subordinate clause, while (1b) is the actual OF example. *János* ('John') is the subject of the embedded clause in both sentences but in (1b) it occurs outside its original place, in the matrix clause preverbal position. Note that in this latter case the fronted constituent bears optional accusative case-marking. Accusative case-marking indicates that the fronted element is OBJ in the main clause. This is related to the possible occurrence of a demonstrative in the standard sentence (1a). These details are going to be important for the forthcoming discussion of the phenomenon.

(1) a. (*Az-t*)   *mondtad, hogy*   <u>*János jön*</u>   *a*   *partira.*
     that-ACC   said.2SG   that(c)[2]   John   comes   the   party.onto
     'You said that John is coming to the party.'

   b. <u>*János(-t)*</u>   *mondtad, hogy*   *jön*   *a*   *partira.*
     John(-ACC)   said.2SG   that(c)   comes   the   party.onto
     '(Of) John you said that he is coming to the party.'

---

[2] The "c" stands for "complementizer". This is to avoid any confusion with the demonstrative in such sentences. If not indicated otherwise, nominative case and present tense assumed in the glosses.

The focus of my research is to answer the following three questions regarding OF:

    i.    What is the relationship between the fronted element and the matrix predicate?

    ii.    What is the relationship between the fronted element and the embedded grammatical function?

    iii.    How does the construction fit into syntactic theory in general, particularly from the perspective of LFG?

The structure of the paper is as follows. Section 2 will give an empirical overview of OF, with special attention to variation in the phenomenon. Here we will see that there are two separate underlying processes and the one resulting in an accusative marked fronted element in (1b) is the theoretically really interesting one. In section 3 I will outline the possible analyses of OF, contrasting an earlier analysis (Coppock (2003)) with the one that I propose. The crucial difference is that I will argue that in the accusative version of (1b), the fronted element is a thematic OBJ of the main predicate. I will support this claim with evidence from several directions. These sections should answer question i) and ii) above. Section 4 will set out to answer question iii), and argues that OF may be related to control-constructions in a broader sense and fits well into the LFG-typology of such phenomena.

## 2. Overview of Operator Fronting

It was Gervain (2002) who realized that OF is not a unitary phenomenon but should be divided into two configurations: using Chomskyan terms, a "movement"-like dependency and a "base-generated" one. That is, if the fronted element retains its original case (nominative in 1b), it originates in the embedded clause and is "raised"/ "moved" into the main clause (much like *wh*-dependencies or topicalization), while "case-switch" (to accusative in 1b) indicates that the it fully belongs to the main clause and is related to the embedded GF through coreference. In the discussion below, I will refer to the first type as "LDD-OF" (long-distance dependency), while the second type is going to be "proleptic-OF" (anticipating the analysis put forth in section 3).

Additionally, Gervain (2002) posits that there is a dialectal divide between native speakers regarding the acceptability of the two versions, but subsequent research (Jánosi 2014) cast doubt on this. In this paper I will treat OF as a dialectally uniform phenomenon.

It is common in the two versions of OF that the distance between the fronted element and its embedded correlate may be large.

(2)    *János(-t)    mondtad, hogy   hallottad,   hogy   jön.*
        John(-ACC)  said.2SG  that(c)  heard.2SG  that(c)  comes
        '(Of) John you said that you heard that he is coming to the party.'

Another common property is that the fronted constituent may bear any grammatical function in the embedded clause, as (3) shows with an OBL. This is usually not discussed in the literature for the case of proleptic-OF. In such cases (i.e. when in proleptic-OF, the embedded GF is not a SUBJ), a resumptive pronoun is obligatory (3b). No such pronoun is possible in (3a).

(3) a. *Londonba   mondtad, hogy   (*oda) mész.*
       London.to    said.2SG that(c) there   goes.2SG

    b. *London-t     mondtad, hogy  oda     mész.*
       London-ACC   said.2SG that(c) there   goes.2SG
       'To London you said that you are going (there).'

The split between LDD- and proleptic-OF may be observed in a wide variety of syntactic contexts. Here I survey a sample of these contexts, for further contrasts, see Gervain (2002, 2009). The data in (4)-(7) also relies on these works.

LDD-OF but not proleptic-OF shows island-sensitivity:

(4) a. *\*János mondtad, hogy  hallottad    a    hírt,    hogy   jön.*
       John    said.3SG that(c) heard.2SG   the  news    that(c) comes.

    b. *János-t   mondtad, hogy  hallottad    a    hírt,    hogy*
       John-ACC said.3SG  that(c) heard.2SG   the  news    that(c)
       *jön.*
       comes.
       '(Of) John you said that you heard the news that he is coming.'

The interpretation of a reciprocal pronoun may reconstruct into the embedded clause, allowing for the backward binding of the experiencer argument (see Pesetsky 1987) only in the case of LDD-OF.

(5) *Egymás      szülei(-\*t)   mondtad,   hogy   elszomorították  a.*
    each.other   parents(-ACC) said.2SG    that(c) saddened.3PL        the
    *fiúkat*
    boys.
    '(Of) Each other's parents you said that they saddened the boys.'

The data in (4) and (5) follows directly from an approach where LDD-OF involves a direct link to the embedded position (through functional identification), while in proleptic-OF the fronted constituent fully belongs to the main clause and the link is less direct instead (as we will see, it is an anaphoric binding relationship).

An important difference is that in the case of quantified fronted elements, the embedded verb may show either singular or plural agreement in the case of proleptic-OF (6b). The possibility of the plural is surprising since these phrases trigger singular agreement in unembedded contexts (7).

(6) a. *Az   összes  fiú mondtad, hogy    jön/      \*jönnek.*
       the  every   boy said.2SG  that(c) come.3SG come.3PL

b. *Az    összes fiú-t      mondtad, hogy   jön/       jönnek.*
    the   every boy-ACC    said.2SG  that(c)  come.3SG   come.3PL
    '(Of) every boy you said that they are coming.'

(7)   *Az    összes  fiú  jön /        *jönnek.*
    the   every   boy come.3SG    come.3PL
    'Every boy is coming.'

This also makes sense if LDD-OF is a strict syntactic dependency ("movement"), while proleptic-OF exploits a coreference relationship.

Finally, the fronted element must be preverbal in the case of LDD-OF, while it can be postverbal in proleptic-OF (although the latter is somewhat stylistically marked, the contrast in (8) is clear).

(8)   *Mondtad János*(-t), hogy     jön.*
    said.2SG  John(-ACC) that(c)     comes
    'You said of John that he is coming.'

We may get an explanation for this also under the assumption to that the fronted constituent is "extracted" in LDD-OF, and such extractions can only target the preverbal area in Hungarian, while the accusative element is a main clause object, which may be either pre- or postverbal.

The properties and the analytical ideas outlined above will be further elaborated in the next section, where I show a possible analysis of OF.

## 3. Analyzing of Operator Fronting

As I have already suggested, it is clear that OF may be licensed in two ways: it could be a strict syntactic dependency (LDD-OF) or a coreference relation (proleptic-OF). In the following sections, I outline how these can be captured in the framework of LFG.

### 3.1  Analyzing LDD-OF

For the purposes of the analysis, I assume a basic phrase structure of Hungarian along the lines of Laczkó (2014), outlined in Figure (1). The sentence is headed by an exocentric S node, which dominates an iterative topic- and quantifier-field, followed by a unique Spec-VP (which hosts focus and various verbal modifiers) and a flat postverbal area. This is of course vastly simplified, a fuller picture would include annotations about information-structure, plus an array of checking equations for proper implementation, see Laczkó (2014).

To analyze LDD-OF, I posit that the phrase-structure rules of Hungarian specify that certain preverbal positions are optionally functionally identified with an embedded GF.
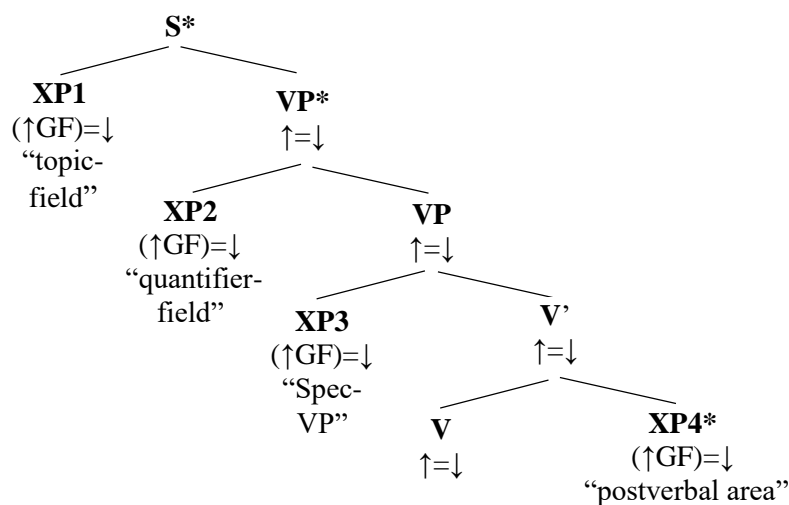
**Figure 1.**
**The basic structure of Hungarian**

Operator Fronted elements can occupy either the XP1 or the XP3 position, so these positions have to be supplied with the following annotation:

(9) a. $(\uparrow\text{LDD-OF-P}\textsc{ath}) = \downarrow$

b. $\text{LDD-OF-P}\textsc{ath} \equiv \text{AF}^+ \qquad\qquad \text{GF}$
$(\rightarrow\textsc{tense})$
$(\rightarrow\textsc{bridge} \neq -)$

(9) indicates that the fronted element may be identified with the GF of any grammatical function embedded in a tensed argument function (AF: SUBJ, OBJ, OBL$_\theta$. OBJ$_\theta$). These tensed argument functions are the subordinate *that*-clauses (the interaction with the associated demonstrative pronouns will be discussed in section 3.2.1 below). The $(\rightarrow\textsc{bridge} \neq -)$ constraint is aimed to exclude identification over a nonbridge-verb, as these block this version of Operator Fronting, as is expected in regular long-distance dependencies.

So for instance in (1b) it is ensured by (9) that *János* ('John') is identified as the SUBJ of the subordinate clause. The clause itself is analyzed as the OBJ of the main verb.

Example (10) illustrates a case where the clause is a SUBJ and the fronted element is the OBJ of this SUBJ.

(10) *Jánost     tilos,      hogy   meghívd.*
John.ACC forbidden    that(c) invite.SBJV.2SG
'You inviting John is forbidden.'

Also, it is possible that some *that*-clauses in Hungarian are functionally OBLs (see e.g. Szűcs (this volume))[3]. For these, the path would involve this argument function.

While this analysis may notationally be different from those already proposed by others (Lipták 1998, Gervain 2002, Coppock 2003), its basic spirit is the same. This is not so with the analysis of the case-switched, proleptic-OF. In the following sections, I will concentrate on that, showing how my proposed analysis differs from the previous ones, particularly Coppock's (2003) LFG analysis.

### 3.1 Coppock (2003)

For the version of (1b) where the embedded subject *János* ('John') bears accusative case in the main clause, Coppock (2003) posits that it functions as an athematic OBJ of the main predicate, much like what we see in the so-called "raising to object"-constructions, see (11) and (12) below.

(11)   *János-t      mondtad, hogy   jön.*
       John-ACC    said.2SG   that(c)  comes
       '(Of) John you said that he is coming.'

(12)   *I believe John to be happy.*

However, unlike "raising to object", the identification of the athematic object in proleptic-OF is not functional, but anaphoric in nature. That is, instead of having strict syntactic identity, the relationship between *Jánost* and the embedded subject is only co-reference. The embedded subject might be pronounced in this scenario, though doing so would be usually dispreferred, Hungarian being a pro-drop language (but see 3b). As indicated in (3), with nonsubject correlates, the pronoun naturally surfaces. Thus for Coppock (2003), (11) has a simplified f-structure like Figure 2. (Information-structure is disregarded.)
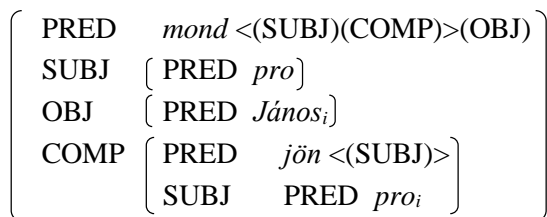
$$
\begin{bmatrix}
\text{PRED} & mond <\!(\text{SUBJ})(\text{COMP})\!>\!(\text{OBJ}) \\
\text{SUBJ} & [\text{PRED } pro] \\
\text{OBJ} & [\text{PRED } János_i] \\
\text{COMP} & \begin{bmatrix} \text{PRED} & jön <\!(\text{SUBJ})\!> \\ \text{SUBJ} & \text{PRED } pro_i \end{bmatrix}
\end{bmatrix}
$$

**Figure 2.**

F-structure for (10) in Coppock's (2003) account.

---

[3] However, for the purposed of the present paper, I remain conservative about the COMP-debate, I assume that this function is available in the inventory of LFG. Nothing crucial depends on this, and if the references paper is along the right track, some details might be recast in a COMP-less approach as well.

A consequence of the anaphoric link is the agreement variation shown in (6b). This may be explained with reference to an ambiguity between INDEX and pragmatic agreement, following the terminology of Wechsler & Zlatić (2003). In their theory, agreement may be of several types: CONCORD (the syntactic side, usually participating in NP/DP-internal agreement), INDEX (based on the semantic content of the nominal, manifest in e.g. subject-verb agreement, anaphora) and pragmatic agreement (based on conceptual and discourse factors).[4] As we have seen in (7), the default INDEX value for quantified noun-phrases in Hungarian is singular. However, we have here an anaphoric binding relationship between the antecedent *Jánost* ('John') and the embedded *pro*. As Wechsler & Zlatić (2003:84) explain, "this invites the possibility of pragmatic/semantic agreement. Hence the range of options for personal pronoun-antecedent agreement are INDEX agreement and pragmatic/semantic agreement". This is exactly what we see in (6b). Despite the singular index, these quantified nominals are conceptually plural, enabling the possible plural agreement.

In addition, Coppock (2003) claims that the singular-plural dichotomy correlates with the distributive-collective interpretation of the embedded predicate. While this is not true in an absolute sense (Gervain 2002:81 observes plural agreement is still possible with exclusively distributive embedded predicates, see (13), there is a preference for plural agreement to be interpreted as collective, as in (13). (The singular version, as far as I can tell, does not display a preference.)

(13)   *Két fiút     mondtál, hogy levegőt   vettek*
       two  boys.ACC said.2SG that(c) air.ACC    took.3PL
       '(Of) two boys you said that they took a big breath.'

(14)   *Két fiút     mondtál, hogy vittek     egy bőröndöt.*
       two  boys.ACC said.2SG that(c) carried.3PL a    suitcase.ACC
       'Of two boys you said that they were carrying a suitcase.'
       → collective reading is preferred

The main problem with this account is that it violates LFG's Semantic Coherence condition, as the semantically meaningful element *János* ('John') is not thematically linked to any predicate. Coppock (2003) realizes this and proposes that the Semantic Coherence condition is only an Optimality Theory-style[5] constraint which may be violated in given circumstances. While this is a possible theory, one may wonder if it is really optimal to downgrade a basic LFG principle to save the analysis of a specific construction. An alternative analysis, without the violation, is certainly preferable. I turn to this in the next section.

---

[4] Also, morphology determines declension.
[5] See e.g. Bresnan (2000).

## 3.2 The proposed alternative: prolepsis

Semantic Coherence is violated in Figure 2 because there is a meaningful (PRED-bearing) element that is an athematic argument, yet is only anaphorically linked to a semantically selected grammatical function. To avoid this situation one may resort to three paths: a) deny that the element in question is semantically contentful; b) posit that the link is functional; c) argue that the element is a thematic argument of the main predicate. Option a) is trivially ruled out since *János* ('John') (and any other fronted element under discussion) is clearly not an expletive. The facts about overt pronouns and agreement variation are clearly opposed to option b). This leaves us with option c): these fronted elements are the *thematic* objects of the main predicates (I will refer to these as "proleptic objects"). In this approach, the f-structure of (11) looks like Figure 3.

$$
\begin{bmatrix}
\text{PRED} & mond <\text{(SUBJ)(OBJ)(COMP)}> \\
\text{SUBJ} & [\text{PRED } pro] \\
\text{OBJ} & [\text{PRED } János_i] \\
\text{COMP} & \begin{bmatrix} \text{PRED} & jön <\text{(SUBJ)}> \\ \text{SUBJ} & \text{PRED } pro_i \end{bmatrix}
\end{bmatrix}
$$

**Figure 3.**
The alternative ("proleptic") account.

This is configuration may be labelled as "prolepsis", which in the definition of Salzmann (2017) is a configuration whereby "a structural complement of the matrix verb is semantically related to the predicate of a finite embedded clause". I will elaborate on the place of prolepsis in syntactic theory in section 4, but first let us see what are the reasons for which one may prefer this account over Coppock (2003).

### 3.2.1 Arguments for a thematic OBJ

The proleptic account avoids the violation of Semantic Coherence, that is an obvious advantage. But as it stands, this is just a technical detour and the advantage melts away if no independent argumentation is put forward for the thematic nature of the OBJ argument. I set out to to provide such argumentation in this section.

One basic question that arises when considering the thematic status of the proleptic object is whether the predicates in question take thematic objects in general. Their availability would increase the plausibility of a thematic analysis. Coppock (2003:136) denies this possibility. While this indeed seems restricted for a verb like *mond* ('say') (but see the discussion below about the status of demonstratives like in sentence (1), other verbs, which may equally

participate in OF, readily allow such objects, as in (15) and (16)[6]. Similar examples may be construed with *fontolgat* ('contemplate'), *állít* ('claim'), *furcsáll* ('find strange'), *jósol* ('predict'), etc.

(15)　*Kétlem　　János　hazamenését.*
　　　doubt.1SG　John　home.going.POSS.ACC
　　　'I doubt John's going home.'

(16)　*Jánost　kétlem,　hogy　hazament.*
　　　John.ACC doubt.1SG　that(c)　home.went.3SG
　　　'(Of) John I doubt that he went home.'

Also, an accusative-marked demonstrative pronoun may occur with every subordinating verb in question, as in (1), repeated here as (17).

(17)　(*Az-t*)　*mondtad, hogy　János　jön　a　partira.*
　　　that-ACC said.2SG　that(c)　John　comes the party.onto
　　　'You said that John is coming to the party.'

This takes us to the debate in Hungarian syntax about the status of such demonstratives. According to the (more or less) standard view of Kenesei (1994), these pronouns are expletives, so they count as athematic arguments in LFG terms. There is an alternative view, dating back to Tóth (2000), which regards the pronouns as contentful (see also Rákosi & Laczkó 2005). While I do not claim that the debate can be considered settled, there are good reasons to adopt this second view.[7] First, similar pronouns systematically occur not only in structural cases, but also in oblique ones, as in (18), for which an expletive-analysis is not plausible (even Lipták 1998, who otherwise endorses Kenesei's view with regards sentences like (17), refers to such oblique pronouns as "argumental referring words").

(18)　*János büszke volt <u>arra,</u>　hogy　győzött.*
　　　John　proud　was　that.onto　that(c) won.3SG
　　　'John was proud that he had won.'

Second, unlike canonical expletives, these pronouns may be associated with discourse functions, e.g. focus in (20).

(19)　*\*Only IT rains.*

---

[6] (15) also illustrates that proleptic-OF is grammatical with a nonbridge-verb. LDD-OF would be ungrammatical here.

[7] The debate is clearly related to the status and analysis of *wh*-scope marking constructions, as in (i). Though a systematic investigation has not been carried out, the present approach is straightforwardly aligned with the "indirect dependency approach", e.g. Dayal (2000).

(i)　<u>*Mit*</u>　*gondolsz, hogy　ki　jön　a　partira?*
　　　what.ACC　think.2SG　that(c)　who　comes　the　party.onto
　　　'What do you think, who is coming to the party?'

(20)     Csak *AZ-T*     *mondtad, hogy*    *János jön*    *a*     *partira.*
          only     that-ACC said.2SG that(c) John    comes the   party.onto
          'You said only that John is coming to the party.'

Third, these pronouns may be coordinated with the undoubtedly thematic objects of the kind shown in (21).

(21)     *Kétlem*     *János hazamenését*        *és*   *(azt),*       *hogy*
          doubt.1SG   John    home.going.POSS.ACC and that.ACC that(c)
          *nem is*      *szólt*     *róla.*
          not even   told.3SG about.it
          'I doubt John's going home and that he didn't even tell us about it.'

Note that the pronoun is optional. What happens is that the OBJ function of the main predicate may be fulfilled by the pronoun. In this case, the clause itself is an adjunct to it. If there is no pronoun, the clause itself bears the respective GF. This gives the prediction that extraction from the clause should only be possible in the absence of the pronoun (in its presence the clause is subject to the Adjunct Island constraint). This appears to be correct, as the following example shows. (22a) is standard LDD-OF, the complement clause being the OBJ of mond ('say'). In (22b), the OBJ is the pronoun, and clause itself is an ADJUNCT. In this latter case, the "extraction" of the fronted element fails.

(22) a. *János*   *mondtad, hogy*    *jön.*
          John     said.2SG that.ACC that(c) comes.
          '(Of) John you said that he is coming.'
    b. *\*János azt*       *mondtad, hogy*       *jön.*
          John    that.ACC   said.2SG that.ACC comes.

Such a pattern actually has already been observed in a number of languages, with similar conclusions, see Bennis (1986) for Dutch and Berman (2001) for German. An example from the latter is shown in (22). *Sagen*'s ('say') OBJ argument may be realized as a clause or as a pronoun in (23a). But as can be seen in (23b), the pronoun-option makes the clause opaque for "extraction", just like in Hungarian.

(23) a. *weil*       *er*   *(es) gesagt hat,*     *dass*   *Hans*   *krank*   *ist*
          because he   it    said   have    that(c) Hans ill      is
          'because he said that Hans is ill'

    b  *\*Was*   *hat*   *er*    *es*     *gesagt,*    *dass*    *er*    *gelesen*   *hat?*
          what    has   he    it      said      that(c) he    read     has
          'What did he say that he read?'

Thus we can conclude that a thematic object for the object argument of these verbs is a possible scenario, be it a pronoun, a regular object or a proleptic one.

More direct evidence for the thematic nature of the object *János* ('John') in (11) is that the construction is incompatible with idiom-chunks, even when they are postverbal (as the preverbal area is associated with discourse-

functions, idiom-chunks are expected to excluded from there, regardless of the thematic status of the argument-slot).

(24) a. *Jánost     elkapta      a     gépszíj.*
    John.ACC caught.3SG  the   driving.belt
    'The driving belt caught John.' → 'John has to work a lot.'

   b. *Mondtad   a    gépszíj-at,     hogy  elkapta    Jánost.*
     said.2SG  the   driving.belt-ACC  that(c) caught.3SG  John.ACC
     '(Of) the driving belt you said that it caught John.' (No idiomatic reading is available.)

Finally, an argument may be construed along the lines of Bresnan (1982:71-72). She observes that the complement clauses of equi-type verbs may be elided given the proper context. This is not possible with raising verbs. The reason is that the stranded object is still semantically interpretable in the first case as the main clause object (*John*) gets a thematic role from *persuade* in (25a), but not from *believe* in (25b). So (25a) is formally incomplete (but reconstructable from discourse) but semantically coherent while (25b) is incomplete and incoherent. As can be seen from (25) proleptic-OF patterns with *persuade*, suggesting that it does get a thematic role from the predicate.

(25) a. *Someone had to wash my car. I persuaded John (to wash my car).*

   b. *Someone stole my car. I believed John *(to have stolen my car.)*

(26)   A: *Szerintem    János  a    legokosabb.*
      in.my.opinion John   the   smartest.
      'I think John is the smartest.'

    B: *De eddig  te  Pétert    mondtad    (hogy  ő    a*
     but so.far you Peter.ACC   said.2SG  that(c) he    the
    *legokosabb).*
    smartest
    'But so far you said (of) Peter (that he is the smartest).'

### 3.2.2 Prolepsis and argument-structure

Now that I have laid down the main ideas of the proleptic analysis, I briefly turn to some details with regards to argument-structure. The basic picture is that the proleptic construction is based on a valency-increasing operation. So the standard lexical entry of a subordinating verb like *mond* ('say') involves two subcategorized GFs, as in (27a), where the agent argument is mapped onto the SUBJ, and the propositional one onto the OBJ. In contrast, the proleptic entry in (27b) has three: the subject (agent), the proleptic object ("subject matter", see Pesetsky 1995) and the clause (proposition). This means that the two OBJs in (27) do not have the same status and the lexical entry in (27a) has undergone a morphosemantic process that rearranged the semantic participants. This may be nicely modelled in Kibort's (2007) conception of LMT, but space-limitations prevent me from demonstrating it here.

(27)  a. <(SUBJ)(OBJ)>
     b. <(SUBJ)(OBJ)(COMP)>

For this alternation to be possible, the added "subject matter"-argument has to be independently available in the conceptual structure of the predicate. Such elements may appear as optional delative-marked adjuncts. In other words, a morphosemantic process "argumentalizes" a nonargument participant.[8] So for instance *érez* ('feel), which only marginally licenses such a participant, is degraded in proleptic-OF too.

(28)  a. *Azt*     *mondtam/??éreztem*  *Jánosról,*  *hogy*  *jön.*
         that.ACC  said.1SG  felt.1SG     John.about  that(c)  comes
         'I said/felt of John that he is coming.'

     b. *Jánost mondtam/??éreztem,*  *hogy*  *jön.*
         John.ACC said.1SG  felt.1SG     that(c)  comes
         'Of John I said that he is coming.'

Another thing that happens with the lexical entry is that a referential identity is forced between the proleptic OBJ and some GF of the embedded clause. This is ensured with an annotation like (29).

(29)    OBJ INDEX= COMP$^{+}$ GF* INDEX

The discussion so far answers the first two questions raised in the introduction: in proleptic-OF and the fronted element is a thematic argument of the main verb and it is related to the embedded grammatical function via anaphoric binding. Now it is time to turn to the third question: how does this analysis relate to a wider syntactic theory?

## 4. Prolepsis in syntactic theory

As already mentioned, an informal definition of prolepsis is given by Salzmann (2017): "a structural complement of the matrix verb is semantically related to the predicate of a finite embedded clause". The term itself originates in rhetoric, meaning "anticipation". Its use in linguistics goes back at least to Higgins (1981). More recently, it has been brought into the theoretical limelight by Davies (2005). It may be used as a contrast to raising constructions: despite its surface similarity, it displays the opposite behavior with respect to a number of standard tests (idiom-chunks, meaning in passive, islands, etc.). Proleptic analyses have been put forward in a number of languages. The basic picture is always the same. A main predicate is associated with three arguments: a subject, a proleptic element and a finite complement clause. Also, the proleptic element is referentially identical with some GF embedded in the complement clause.

Some examples for proleptic analyses are shown below. For English in (30) by Massam (1985), or for Madurese by Davies (2005) in (31). Other examples

---

[8] This likens prolepsis to applicatives, a research avenue yet to be explored.

include German (Salzmann 2017), Greek (Kotzoglou & Papangeli 2007), Japanese, Korean (Yoon 2007).

(30) *I read of Carrol that she was awfully shy.*

(31) *Siti  ngera  Hasan bari'     melle motor.*
     Siti  think  Hasan yesterday buy  car
     'Yesterday Siti thought about Hasan$_i$ that he$_i$ bought a car.'

(30) and (31) also illustrates a split between two types of proleptic constructions. In (30) the proleptic element (*of Carrol*) is an adjunct PP[9], while in (31) *Hasan* serves as a direct object of the main verb (Hungarian OF fits into this second pattern). This latter seems to be rarer cross-linguistically, which is expected if the "argumentalization" is an additional process, as suggested in the previous section.

Now if we compare the proleptic f-structure in Figure 3 with a standard object-equi sentence, the similarity is obvious.

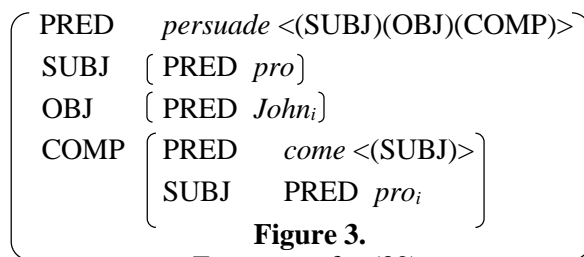(32)    *I persuaded John to come.*

$$\begin{bmatrix} \text{PRED} & persuade <(\text{SUBJ})(\text{OBJ})(\text{COMP})> \\ \text{SUBJ} & [\text{PRED } pro] \\ \text{OBJ} & [\text{PRED } John_i] \\ \text{COMP} & \begin{bmatrix} \text{PRED} & come <(\text{SUBJ})> \\ \text{SUBJ} & \text{PRED } pro_i \end{bmatrix} \end{bmatrix}$$

**Figure 3.**
F-structure for (32).

In both structures, a matrix-clause thematic dependent in an anaphoric binding relationship with an embedded argument. Note that equi is also standardly contrasted to raising, essentially along the same lines that I have mentioned earlier. Thus, I suggest that prolepsis fits into syntactic theory by virtue of being a kind of control-construction[10]: a finite, anaphoric, equi-type control. The next two sections elaborate on this claim.

---

[9] Note however that Chomsky (1981) mentions dialectal sentences like (i), where interrogative word seems to receive accusative case from the main predicate.

   (i)    *Whom do you suggest should become president?*

[10] It is important to note that I use the term "control" in an LFG-sense, which encompasses both raising- and equi-type constructions (Falk 2001:117-118). In both, some sort of identity is established between a main clause- and an embedded clause-dependent. This differs to the GB/MP terminology, where "control" only refers to equi-type constructions. For some details of the raising/equi dichotomy, see section 4.2.

## 4.1 Prolepsis and equi

First, it must be admitted that for prolepsis to qualify as a control-construction, we must adopt a wider concept of control than is usual. Control standardly refers to constructions where the referential dependency is established with the subject of an immediately embedded, non-finite complement clause. As we have seen, these do not apply to prolepsis, as here the embedded clause is finite, the controlled function may be nonsubject (as in (3b)), and longer paths are also possible (as in (2) and (4b)). (It may be noted that the construction in Greek as described by Kotzoglou & Papangeli 2007 is restricted to immediately embedded subjects, so it conforms to standard control to a larger extent.)

However, if we abstract away from these contextual factors, and focus on the mechanism itself, the parallel is undeniable, which justifies the wider perspective of control. As we will see, there is empirical support for this position and it leads into a typology of control-constructions that is fully consistent with the architecture of LFG.

Just like in standard control, the co-reference of the matrix-dependent and the embedded argument is obligatory. Despite being finite, the subject of the embedded clause in (34) (indicated as *pro* here for convenience) must be *János* ('John'). So, the control-relationship is obligatory.

(33)  *I persuaded John$_i$ pro$_{i/*j}$ to come.*

(34)  *János-t$_i$      mondtad,  hogy    pro$_{i/*j}$  jön.*
John-ACC    said.2SG  that(c)          comes
'(Of) John you said that he is coming.'

As summarized in Landau (2013), a core-feature of obligatory control constructions is that they license bound-variable interpretations and sloppy interpretation under ellipsis as in (35)-(38) below. Prolepsis, being an obligatory anaphoric dependency, conforms to these properties of control.

(35)  *Only Bill forced himself to jump.* (Bill = Only x [x forced x to jump].)

(36)  *Csak János mondta    magát,     hogy  (ő) nyert.*
only John   said.2SG    himself.ACC that(c)  he won.3SG
'Only John said (of) himself that he had won.' (John = Only x [x said of x that x won].)

(37)  *Mary encouraged Paul$_i$ PRO$_i$ to attend the ceremony, but not David$_j$ (encourage PRO$_{*i/j}$ to attend the ceremony).*

(38)  *Én  János$_t$$_i$ mondtam,  hogy  jön,   te   pedig  Pétert$_j$*
I   John.ACC said.1SG    that(C) comes you  but    Péter.ACC
*(mondtad,   hogy   PRO$_{*i/j}$ jön).*
said.2SG     that(C)          comes
'(Of) John I said that he comes and you did so of Peter.'

## 4.2 Prolepsis and the typology of control

As already mentioned in footnote 7, the term "control" here is taken to refer to construction where a referential identity is forced between a matrix-clause and an embedded-clause dependent. Also, these constructions are rooted in the lexical items themselves, rather than being structurally enforced, as in *wh*-type structures.

Control may be classified along several axes. An essential one is the raising-equi opposition. While in the former, there is no thematic relationship between the main predicate and the controller argument, there is one in the latter. That is, as commonly held, John is not a "seemer" in (39a), but he is a "trier" in (39b).

(39)  a. *John seemed to win.*

   b. *John tried to win.*

Next, in LFG there are two mechanisms that can establish f-structural identities. One is anaphoric control which is semantic co-reference, as in prolepsis. The other is functional control, a strict, syntactic identity, whereby one f-structural element simultaneously satisfies two grammatical functions.

In principle, these two axes may co-vary. However, one constellation is barred by the LFG-architecture: raising with anaphoric identification. The reason is that this leads to the situation that was the original problem with Coppock's (2003) analysis of OF. The matrix argument remains semantically unintegrated: it is linked semantically neither to the main predicate, nor to the embedded one (it is simply co-referent with the controllee, which is independently licensed in the COMP).

Nevertheless, in equi, the bifurcation into a functional and an anaphoric type is a real option. While in Dalrymple (2001) equi is exclusively linked to anaphoric control, Falk (2001) argues that the picture is more complex and some equi verbs use functional control. His case in point is *try* vs. *agree*. According to Falk (2001:136-139), *try* should be analyzed as having a functionally controlled XCOMP complement, while *agree* has an anaphorically controlled COMP. Support for this is available from passivization (39) or the possibility of partial control (41).[11] In (40a), the controller (and thus the controllee) is simply missing. In contrast, discourse control of the embedded subject is possible in (40b), since it is independently licensed as an f-structural *pro* subject. (Essentially, it is a covert pronoun.)

(40)  a. *\*It was tried to finish earlier.*

   b. *It was agreed to finish earlier.*

---

[11] See Haug (2013) for an LFG account of partial control.

In (41) for *try* functional control forces a strict referential identity of the controller and the controllee, the semantic nature of anaphoric control allows for some latitude for *agree*.

(41) a. *John tried to go outside.* (understood subject of *go*: only John)

b. *John agreed to go outside.* (understood subject of *go*: John + possibly other people)

However, while functional control strictly implies exhaustive control, anaphoric control is not necessarily semantically loose. It may be, as in (40), but exhaustively interpreted strict anaphoric control is possible (in fact, prolepsis is such a scenario). I agree with Haug (2013), who suggests that anaphoric control should be viewed as a continuum from obligatory through "quasi-obligatory" to no-control.

| CONTROL-TYPE | | | Example |
|---|---|---|---|
| **Thematicity of controller** | **Nature of identification** | **Finiteness** | |
| Equi (thematic) | Anaphoric identification | Finite complement | prolepsis |
| | | Non-finite complement | canonical control ("*agree*-type") |
| | Functional identification | Finite complement | Turkish object control (?), see 4.2.2 |
| | | Non-finite complement | canonical control ("*try*-type") |
| Raising (non-thematic) | Anaphoric identification | Finite complement | not expected |
| | | Non-finite complement | not expected |
| | Functional identification | Finite complement | Copy Raising/ Hyperraising see 4.2.1 |
| | | Non-finite complement | canonical raising |

**Table 1.**
An LFG-typology of control.

Putting all these dimensions together (raising/equi, anaphoric/functional, finite/nonfinite embedded clause), we may arrive at the above typology of control-constructions from an LFG-perspective.[12] In the remainder of the paper, I give a brief overview of those constructions in the taxonomy that have not been mentioned so far.

### 4.2.1 Finite raising

Although raising, just like equi, is traditionally associated with nonfinite clauses, there are well-described cases of raising from finite clauses. Instances of this with an empty embedded position are labelled Hyperraising, while cases with an overt resumptive pronoun are called Copy Raising. The schematic structure of these are shown in (42).

(42)  a. $DP_i$ [$IP_{finite}$ resumptive$_i$]       b. $DP_i$ [$IP_{finite}$ _____$_i$]

These two structures are amenable to a parallel analysis, the overtness of the resumptive pronoun seems to boil down to the status of the pro-drop parameter in the given language (see Ademola-Adeoye 2010).

A Copy Raising example is shown in (43), from Igbo (Ura 1998), a Hyperraising one in (44), from Lubukusu (Carstens & Diercks 2013). Space limitations prevent me from elaborating on the details, but the cited works provide evidence from expletives, idioms and scope considerations that these are indeed raising structures.[13]

(43)  *Ézè$_i$ di    m    [kà    ọ$_i$    hũrũ Adá]*.
      Eze  seems  to me  that(c) he    saw Ada
      approx.: 'Eze seems to me that he saw Ada.'
(44)  *Chisaang'i  chilolekhana  mbo    chikona*
      animal       seem          that(C) sleep
      approx.:'The animals seem that they are sleeping.'

### 4.2.2 Finite equi, functionally identified

Ince (2006) describes a Turkish object-control structure where the embedded clause has both tense and agreement features, thus qualifying as a finite clause. Idiom-chunks are excluded from this sentence-type, so it must be equi. There is no mention of any referential flexibility in the data, so I tentatively categorize this as functional identification. (Further investigation may alter this conclusion.)

---

[12] A possible extension of the typology is to integrate the direction of the control-relation, i.e. forward vs. backward control. For discussion, see Haug (2017).

[13] The proper analysis of the related English sentences like (i) is not settled at this point, see Asudeh & Toivonen (2012) vs. Landau (2011).

(i)  *Richard seems like he smokes.*

(45) *Ben Ali-yi      yarın      balığı  yiyecek      sanıyordum.*
     I    Ali-ACC  tomorrow  fish    eat.fut.3SG  assumed.1SG
     'I thought that Ali will eat the fish tomorrow.'

## 5. Summary

In this paper I set out to investigate the properties and the analysis of the Hungarian Operator Fronting construction. My main questions concerned the relationship between the fronted element, the main predicate and the embedded correlate as well as the theoretical implications of the phenomenon.

It was established that two subtypes of OF should be distinguished: LDD-OF and proleptic-OF. LDD-OF is a standard long-distance dependency: the fronted element is not related to the main predicate and is linked to its embedded position via a standard functional uncertainty-equation. In proleptic-OF, the fronted constituent becomes the direct thematic object of the main predicate (contra Coppock 2003) and is linked to the embedded GF via obligatory anaphoric binding. This may be regarded as a finite, equi-type anaphoric control relationship, which can be neatly placed in the LFG-conception of control-constructions.

As one of my reviewers noted, Asudeh (2005) already anticipated that the interaction of resumption, raising, control and finiteness constitute a promising avenue for typological research in LFG.[14] This paper may be seen as some fulfillment of this anticipation.

## References

Ademola-Adeoye, Feyisayo Fehintola. 2010. *A Cross-linguistic Analysis of Finite Raising Constructions*. Doctoral dissertation, University of KwaZulu-Natal.

Asudeh, Ash. 2005. Control and semantic resource sensitivity. *Journal of Linguistics* 41: 465–511.

Asudeh, Ash, Ida Toivonen. 2012. Copy raising and perception. *Natural Language and Linguistic Theory* 30(2): 321-380.

Bennis, Hans. 1986. *Gaps and dummies*. Dordrecht: Foris.

Berman, Judith. 2001. On the cooccurrence of *es* with finite clause in German: an LFG analysis. In Meurers, Walt D., Kiss Tibor (eds.), *Constraint-Based Approaches to Germanic Syntax*. Stanford: CSLI Publications. 7-30.

Bresnan, Joan (ed.). 1982. *The Mental Representation of Grammatical Relations*. Cambridge, MA: MIT Press.

---

[14] His theory of control is based on a resource management semantics and arrives at this conclusion: "(…) Finite control can therefore be seen as the control analogue of copy raising. The analysis has typological implications for control mechanisms and extends the relation between control and raising in a new dimension" (Asudeh 2005: 509).

Bresnan, Joan. 2000. Optimal syntax. In Dekkers, Joost, Frank van der Leeuw, Jeroen van de Weije (eds.), *Optimality Theory: Phonology, Syntax and Acquisition*. Oxford University Press. 334–385.

Carstens, Vicki, Michael Diercks. 2013. Parameterizing case and activity: hyper-raising in Bantu. In Kan, Seda, Claire Moore-Cantwell, Robert Staubs (eds.), *Proceedings of the 40th Annual Meeting of the North East Linguistic Society*. Amherst: University of Massachusetts Graduate Linguistic Student Association. 99-118.

Chomsky, Noam (1981). *Lectures on Government and Binding*. Dordrecht: Foris.

Coppock, Elizabeth. 2003. Sometimes it's hard to be coherent. In Butt, Miriam, Tracy H. King (eds.), *Proceedings of the LFG03 Conference.* Stanford: CSLI Publications. 126-143.

Dalrymple, Mary 2001. *Lexical Functional Grammar*. San Diego: Academic Press.

Davies, William D. 2005. Madurese prolepsis and its implications for a typology of raising. *Language* 81: 645–665.

Dayal, Veneeta. 2000. Scope marking: cross-linguistic variation in indirect dependency. In Lutz, Uli, Gereon Müller, Armin von Stechow (eds.), *Wh-scope marking*. John Benjamins Publishing Company. 157-193.

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

Falk, Yehuda N. 2001. *Lexical-Functional Grammar*. Stanford: CSLI Publications.

Gervain Judit. 2002. *Linguistic Methodology and Microvariation in Language: The Case of Operator-raising in Hungarian*. MA Thesis, University of Szeged.

Gervain Judit. 2009. Resumption in Focus(-raising). *Lingua* 119: 687-707.

Haug, Dag. 2013. Partial control and the semantics of anaphoric control in LFG. In Butt, Miriam, Tracy H. King (eds.), *Proceedings of the LFG13 Conference*. Stanford: CSLI Publications. 274–294.

Haug, Dag. 2017. Backward control in Ancient Greek and Latin participial adjuncts. *Natural Language and Linguistic Theory* 35(1): 99–159.

Higgins, Roger. 1981. Proleptic objects and verbs of perception in Zacapoaxtla Nahuat. *Texas Linguistic Forum* 18: 69-88.

Ince, Atakan. 2006. Direct complement clauses as object control structures in Turkish. In Bainbridge, Erin, Brian Agbayani (eds.), *Proceedings of the thirty-fourth Western Conference On Linguistics*. Department of Linguistics, California State University, Fresno. 208-221.

Jánosi Adrienn. 2013. *Long split focus constructions in Hungarian with a view on speaker variation*. Doctoral dissertation, CRISSP.

Kenesei István. 1994. Subordinate clauses. In Kiefer Ferenc, É. Kiss Katalin (eds.), *The syntactic structure of Hungarian*. San Diego: Academic Press. 141-165.

Kibort, Anna. 2007. Extending the applicability of Lexical Mapping Theory. In Butt, Miriam, Tracy H. King (eds.), *Proceedings of the LFG08 Conference*. Stanford: CSLI Publications. 250–270.

Kotzoglou, George & Dimitra Papangeli. 2007. Not really ECM, not exactly control: The "quasi-ECM" construction in Greek. In Davies, William D., Stanley Dubinsky (eds.), *New Horizons in the Analysis of Control and Raising*. Dordrecht: Springer. 111–131.

Laczkó Tibor. 2014. The essentials of an LFG-analysis of Hungarian finite sentences. In Butt, Miriam, Tracy H. King (eds.), *Proceedings of the LFG14 Conference*. Stanford: CSLI Publications. 325-345.

Landau, Idan. 2011. Predication vs. aboutness in copy raising. *Natural Language and Linguistic Theory* 29: 779-813.

Lipták Anikó. 1998. A magyar fókuszemelések egy minimalista elemzése. In Büky László, Maleczki Márta (eds.), *A mai magyar nyelv leírásának újabb módszerei 3*. Szeged: University of Szeged. 93-116.

Pesetsky, David. 1987. Binding Problems with Experiencer Verbs. *Linguistic Inquiry* 18(1): 126-140.

Pesetsky, David. 1995. *Zero syntax: Experiencers and Cascades*. Cambridge, Massachusetts/London, England: The MIT Press.

Rákosi György, Laczkó Tibor. 2005. Verbal category and nominal function – Evidence from Hungarian Subject Clauses. In Butt, Miriam and Tracy H. King (eds.), *Proceedings of the LFG05 Conference*. Stanford: CSLI Publications. 353-370.

Salzmann, Martin. 2017. Prolepsis. In Everaert, Martin, Henk van Riemsdijk (eds.), *The Wiley-Blackwell Companion to Syntax, 2nd edition*. Wiley-Blackwell. 3203-3245. http://martinsalzmann.com/files/Salzmann2017_Prolepsis_final.pdf

Tóth Ildikó. 2000. *Inflected Infinitives in Hungarian*. Tilburg: TILDIL Dissertation Series.

Wechsler, Stephen, Larisa Zlatić. 2003. *The Many Faces of Agreement: Morphology, Syntax, Semantics, and Discourse Factors in Serbo-Croatian agreement*. Stanford: CSLI Publications.

Yoon, James H-S. 2007. Raising of major arguments in Korean and Japanese. In Davies, William D., Stanley Dubinsky (eds.), *New Horizons in the Analysis of Control and Raising*. Dordrecht: Springer. 71-107.