SOMETIMES IT'S HARD TO BE COHERENT

Elizabeth E. Coppock

Stanford University

Proceedings of the LFG03 Conference

University at Albany, State University of New York

Miriam Butt and Tracy Holloway King (Editors)

2003

CSLI Publications

http://csli-publications.stanford.edu/

Abstract

The paper provides an analysis of the construction in Hungarian known as *focus raising*, illustrated below.

János-t mond-t-am hogy jön
John-ACC say-PAST-1SG that come.3SG
'It is John that I said is coming.'

In focus raising, the focus of the matrix clause is identified with a grammatical function in the embedded clause. Focus raising is particularly interesting because of the case marking on focus-raised subjects. Rather than having nominative case (unmarked), they appear suffixed by the accusative marker, *-t*, as shown in the example. This fact is analyzed in the paper as follows.

- Accusative NPs such as *Jánost* in the example function as objects; this is evidenced by verb-object definiteness agreement, and the possibility of a reflexive pronoun in the same position.
- On the other hand, the object function is argued not to be thematic (i.e., semantically selected by the matrix predicate), based on its ability to be realized as an expletive pronoun, and the long-distance nature of focus raising.

A pattern of variation among Hungarian speakers observed by Gervain (2002) shows that this athematic object function may host either functional or anaphoric binders for the embedded clause subject. Anaphoric binding from an athematic position leads to a violation of the semantic Coherence condition (Dalrymple 2001, p. 243), but this condition still appears to apply in other cases. Therefore, an Optimality-Theoretic treatment is proposed, in which the semantic Coherence condition is a violable constraint.

Focus raising: *wh-* movment without the *wh-*

In Hungarian, just as in English, question *wh-* phrases such as *mikor* 'when' are normally fronted to the left periphery of the sentence, as in (1). Unlike in English, non-*wh-* phrases like *holnap* 'tomorrow' may front to the same position, as shown in (2).[1] (The fronted constituent is indicated in bold face.)

(1) **Mikor** mond-ott, hogy jön?
when say-PAST.3SG that come.3SG
'When did he say he would come?'

(2) **Holnap** mond-ott, hogy jön.
tomorrow say-PAST.3SG that come.3SG
'He said that he would come *tomorrow*'

(3) *Tomorrow did he say that he would come.

The phenomenon shown in (2) is often called *focus raising*, and has been analyzed in numerous works, including Zolnay (1926), Marácz (1989), É. Kiss (1987, 1990), Kenesei (1994), Horvath (1986, 1995, 1998), Lipták (2001), and perhaps most recently, (Gervain, 2002).

The focus-raising construction shares many properties with the *wh-* question construction in Hungarian. In both constructions, the fronted element is expressed in the focus position of the matrix clause, immediately preceding the verb. In addition, the extraction path may cross several finite clause boundaries, as exemplified in (4), where *ket dolgot* 'two things' fronts across two CPs (É. Kiss, 1987, p. 125).

(4) **Két dolgot** hallottam, [$_{CP}$ hogy János megígért Máriának, [$_{CP}$ hogy megtesz]]
two thing.ACC hear.PAST.1.SG. that John promised Mary.DAT that does
'It was two things that I heard that John promised Mary that he would do.'

Moreover, when the matrix predicate in a focus-raising or question construction is a verb, it must be a *bridge verb*, such as the translational equivalents of *want* and *say*; as shown in (6), extraction may not cross the complement to a non-bridge verb such as *figyel* 'warn' (É. Kiss, 2002).[2]

(6) *A kollégáim **egy diák** figyelmeztettek hogy keres-ett
the my:colleagues a student.NOM warned:me.3PL that seek.3SG-PAST
'My colleagues warned me that a student was looking for me.'

Finally, focus raising, like question formation, is sensitive to island constraints, such as the Complex Noun Phrase Constraint (É. Kiss, 1987). For example, (7b) is ungrammatical because extraction may not cross the complex noun phrase *azt, hogy érkezik* 'it, that he would arrive'.

(7) a. János mondta [azt, [hogy holnap érkezik]]
John said it-ACC that tomorrow arrives
'It John who said it, that he would arrive tomorrow.'

[1] David Beaver (p.c.) points out that *Tomorrow, he said that he would come* is, on the other hand, grammatical, so English does have a type of long-distance focus movement. My point is not to compare Hungarian and English in this paper, but the fact remains that Hungarian and English differ in that long-distance focus-movement and question formation have identical syntax in Hungarian, but not in English.

[2] Focus-raising may also cross adjectival predicates with clausal complements, however (Kenesei, 1994, p. 317):

(5) **Emmá-t** fontos hogy megl´atogas-s-ad
Emma-ACC important that visit-SBJ-2SGDEF
'It is Emma whom it is important that you visit.'

b. *János **holnap** mondta [azt, [hogy érkezik]]
John tomorrow said it-ACC that arrives
'It was tomorrow that John said it, that he would arrive.'

These parallelisms indicate that Hungarian focus-raising and question formation instantiate a unified phenomenon.

NPs bearing a variety of grammatical functions in an embedded clause may undergo focus raising, as shown in (8), where a dative argument, an inessive adjunct, and a direct object, respectively, are shown fronted via focus-raising.

(8) a. Mária **János-nak** akarta, hogy a csomag-ot kézbesítsék
Mary John-DAT wanted that the parcel-ACC deliver.3PL
'As for Mary, it was John that she wanted the parcel to be delivered to.'

b. Legjobban **eb-ben a kérdes-ben** szeretném, hogy megegyezz-ünk
most this-in the question-in would:like.1SG that agree.SBJ-1PL
'It is on this question that I would like most that we agree.'

c. Mindenki **Máriá-t** várta, hogy megválasztják
everyone Mary-ACC waited that elect.3PL
'It was Mary that everybody expected that they elect.'

Usually, the fronted element bears the morphological case that it would bear if it were expressed in the lower clause. Thus, the non-fronted version of (8a) is as in (9), where the dative noun phrase *Jánosnak* still bears dative case.

(9) Mária akarta, hogy a csomagot **János-nak** kézbesitsék.
Mary wanted that the parcel.ACC John-DAT deliver.3PL
'As for Mary, it was John that she wanted the parcel to be delivered to.'

However, there is one class of exceptions to this general pattern, both in focus-raising and in question formation. When the subject of the embedded clause raises to the matrix focus position, the fronted phrase often (optionally for some speakers, obligatorily for others) bears accusative case:

(10) **Péter-t** mond-t-a, hogy jön
Péter-ACC say-PAST-1SG.DEF that come.3SG
'It is Peter who he/she said is coming.'

Péter would of course get nominative, not accusative, downstairs:

(11) János mond-t-a, hogy **Péter** jön
I say-PAST-1SG.DEF that Peter.NOM come.3SG
'John said that Peter is coming'

It is this accusative marking that I seek to explain here.

§1 is devoted to determining whether, and in what sense accusative-marked, focus-raised elements originating from subject position (henceforth, *Pétert*) are objects, as the case-marking suggests. I will conclude that while *Pétert* is in fact an object of the matrix clause, it is an athematic object.[3] In §2, I will argue for one dialect of Hungarian that this athematic object position is anaphorically, rather than functionally, identified with the embedded clause subject. This implies that the phenomenon is a raising construction with anaphoric control, a possibility that is predicted typologically, but unattested.

[3]An athematic GF is one that is not semantically selected by the PRED that governs it; notationally, an athematic GF does not appear inside the angle brackets enclosing the argument list of a PRED.

1 Objecthood and selection

The accusative case marking on *Pétert* in (10) indicates that it functions as a (direct) object (OBJ), since OBJs normally go with accusative case in Hungarian. That hypothesis is shown schematically in (i). If that were true, this construction would exemplify a mismatch of function and meaning; even if *Péter* is functionally the OBJ of the matrix clause, intuitively it is semantically interpreted only as the subject of the embedded clause (Judit Gervain, p.c.). The alternative proposal would be that *Péter* is not functionally an OBJ, even though it formally bears accusative case, as shown schematically in (ii).

(12) (i) $\begin{bmatrix} \text{FOC} & \boxed{1}\begin{bmatrix}\text{CASE} & \text{ACC}\end{bmatrix} \\ \text{OBJ} & \boxed{1} \\ \text{PRED} & \text{'say...OBJ...'} \end{bmatrix}$ (ii) $\begin{bmatrix} \text{FOC} & \begin{bmatrix}\text{CASE} & \text{ACC}\end{bmatrix} \\ \text{PRED} & \text{'say...'} \end{bmatrix}$

Even if *Pétert* functions syntactically as an argument in the matrix clause, we are not obligated to consider it a semantic argument of the verb. Assuming that it is functionally an object is consistent with an analysis conforming to schema (i) below, where the verb syntactically and semantically selects an object (i.e., the object is thematic). That assumption would also be consistent with an analysis like (ii), where the verb syntactically but not semantically selects an object (i.e., the object is athematic).

(13) (i) $\begin{bmatrix} \text{OBJ} & \begin{bmatrix}\text{CASE} & \text{ACC}\end{bmatrix} \\ \text{PRED} & \text{'say}\langle\ldots\text{OBJ}\ldots\rangle\text{'} \end{bmatrix}$ (ii) $\begin{bmatrix} \text{OBJ} & \begin{bmatrix}\text{CASE} & \text{ACC}\end{bmatrix} \\ \text{PRED} & \text{'say}\langle\ldots\rangle\text{OBJ'} \end{bmatrix}$

I will argue in favor of analysis (ii) in (13), in which the accusative-marked, focus-raised element functions syntactically as an argument of the matrix clause, but is not a semantic argument of the verb.

1.1 Objecthood

1.1.1 Definiteness agreement

Patterns of definiteness agreement suggest that *Pétert* is indeed an object of the matrix verb. In Hungarian, verbs agree in definiteness with their direct object.[4] Thus, the verb form of *lát* 'see' is different in *I saw a bird*, as in (15) from what it is in *I saw the bird*, as in (14).

(14) Lát-om a madar-at
see-1.SG.DEF the bird-ACC
'I see the bird'

(15) Lát-ok egy madar-at
see-1.SG.INDEF an bird-ACC
'I see an bird'

(The indefinite conjugation is also used when the verb has no object.)

When a definite NP such as the proper name *Péter* (accusative *Péter-t*) focus-raises, the matrix verb appears in the definite conjugation (*mond-t-a* 'say-PAST-3SG.DEF'), as in (10). Noun phrases with plural numeric determiners such as *két fiu-t* in (16) and question words such as *ki-t* in (17) are indefinite, so the matrix verb goes in the indefinite conjugation (*mond-ott* 'say-3SG.PAST.INDEF') in these cases.

[4] See Bartos (1997) for discussion of the precise semantic characterization of the agreement; semantic definiteness is not obviously sufficient.

(16) **Két fiu-t** mond-ott hogy jön.
two boy-ACC say-3SG.INDEF.PAST that come.3SG
'It was two boys that he/she said were coming.'

(17) **Ki-t** mond-ott, hogy jön?
who-ACC say-3SG.INDEF.PAST that come.3SG
'Who did he/she say is coming?'

If the choice between the definite and indefinite conjugation for a verb is based on the properties of its object, as is normally assumed, then these accusative NPs are objects.

1.1.2 Reflexives

Another argument for the objecthood of accusative-case marked, focus-raised objects is the fact that reflexive pronouns coreferential with the matrix subject may appear in this position, as in (18).

(18) Önmagá-t$_i$ mondta Péter$_i$ hogy szeret-i Mari-t
themself-ACC say-PAST-3SGDEF Peter.NOM that love-3SG.DEF Mary-ACC
'It is himself$_i$ that Peter$_i$ said loves Mary.'

(This example is well-formed only as an answer to the Hungarian equivalent of the question, "Who did Peter say loves Mary?", or as a correction to an assertion that Peter said someone else loves Mary.) As I will argue, the reflexive pronoun *önmagá(t)* must be bound by a syntactic coargument; an argument function syntactically selected by the same PRED. Therefore, the accusative NP in (18) occupies an argument function.

Under the non-object analysis of accusative marked focus raised subjects schematized in (ii) of (12), the reflexive pronoun in (18) is the embedded clause subject, bound by the subject of the matrix clause. If *önmagát* could be bound by the subject of a higher clause, we would predict ambiguity in (19); the reflexive anaphor should be able to take either *János* or *Mari* as an antecedent. (Note that pronouns in Hungarian are not specified for gender.)

(19) János$_j$ mond-t-a hogy önmagát$_{i,*j}$ szeret-i Mari$_i$
John.NOM mond-PAST-3SG.DEF that themself love-3SG.DEF Mary.NOM
'John said that Mary loves herself.'
'*John said that Mary loves himself.'

In fact, in this case *önmagát* may only refer to *Mari*, the potential binder that is a coargument.

Towards the same point, if the non-object analysis were right, then the matrix subject should be able to bind a reflexive pronoun in the subject of the embedded clause, even if focus raising has not taken place. There is a nominative reflexive pronoun in Hungarian (*önmaga*), and this pronoun cannot be bound by the subject of the matrix clause either, as in (20).

(20) *János mond-t-a hogy önmaga szeret-i Mari-t
John say-PAST-3SG.DEF that themself love-3SG.DEF Mary-ACC
'It is John who said that himself loves Mary.'

Thus, the reflexive pronoun must be bound within its clause. Example (18) therefore provides evidence that accusative-marked focus raised subjects have a grammatical function in the matrix clause.[5]

[5]Thanks to Amy Dahlstrom for suggesting this avenue of inquiry. Thanks also to Agnes Mihalik for providing these native speaker judgments.

1.2 Semantic selection

Again, if accusative-marked focus-raised embedded-clause subjects such as *Pétert* in (10) are objects, they are not necessarily semantically selected objects. Indeed, I will argue that they are not.[6]

1.2.1 Dependency is not lexically enforced

First, let us suppose the alternative: accusative-marked fronted subjects are semantically selected. Then we have a lexical entry for bridge verbs such as *mond* 'say' that includes the following lexical specification:

(21) (↑PRED)='say⟨SUBJ OBJ COMP⟩'

Supposing this were the lexical entry, we would predict (22) to be possible, in the absence of any other mechanism for ruling it out. (22) is ungrammatical because the matrix object function cannot be identified functionally or anaphorically with any participant in the embedded clause.

(22) *Péter-t mond-t-am, hogy János jön
Péter-ACC say-PAST-1SG that John.NOM come.3SG
'[roughly] Peter did I say that John is coming'

The putative lexical entry in (21) is like the lexical entry for an equi control verb, which by definition thematically selects the controller along with a complement clause (normally XCOMP). An equi verb such as *try* enforces identity between the controller and the controllee through a stipulation of referential equality between the matrix controller and the SUBJ of the embedded clause, in the verb's lexical entry (Dalrymple, 2001). This precise solution will not do for us, because focus-raising is possible over long distances, as exemplified earlier in (4). Thus, the identity stipulation would have to involve a functional uncertainty equation allowing the controllee to be embedded within multiple COMPs. However, the possible extraction paths for focus-raising mirror precisely the possible extraction paths for *wh*- questions. For example, it is not possible to focus-raise out of a complex noun phrase, as demonstrated earlier in (7b). This parallelism shows that the range of dependencies in focus-raising constructions is not determined lexically, by bridge verbs, but by the aspect of the grammar that regulates long-distance dependencies. Supposing that bridge verbs semantically select for accusative-marked focus-raised subjects prevents us from making use of the grammar to derive the constraints on extraction.

In contrast, if we assume that accusative-marked fronted subjects are not semantically selected by the verb, then we can use the grammar to impose identity between the higher-clause object and an embedded subject. If the object is athematic, then the representation for (10) will consist at least of (23).

(23)
$$\begin{bmatrix} \text{PRED} & \text{'say}\langle \text{SUBJ COMP} \rangle\text{OBJ'} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'John'} \end{bmatrix} \\ \text{COMP} & \begin{bmatrix} \text{PRED} & \text{'come}\langle \text{SUBJ} \rangle\text{'} \end{bmatrix} \end{bmatrix}$$

If OBJ is not functionally identified with any function that is semantically selected, then the condition in (24) will be violated.

[6] I assume with little argument that these accusative-marked elements are indeed subcategorized for by the verb if they are objects. However, an analysis using nonsubcategorized objects may be correct for the cognate object construction (e.g. *Suzy smiled a pretty smile*) as well as certain objectlike temporal adverbials (e.g. *I slept two hours*), so it may be fruitful to explore that possibility.

(24) *Semantic Coherence* (Dalrymple, 2001, p. 243)
A meaning derivation for an utterance is semantically coherent if the meaning derivation produces a meaning for the utterance with no additional unused premises remaining.

This principle is independently motivated as a means of ruling out examples like (25).[7,8]

(25) a. *John rained.

b. *John seems that Mary is nice.

For simplicity of exposition, we may obtain the effect of the semantic coherence condition for our purposes without using glue semantics, if we assume that every instance of a PRED attribute corresponds to a glue semantic resource, and that semantic government corresponds to the using up of premises. Using this correspondence, we may state the semantic Coherence condition as follows:

(26) *Semantic Coherence* (non-glue version):
Every governable function containing a PRED value must be semantically governed.

(27) *Semantic government* (for our purposes)
f is semantically governed iff f is functionally linked to a position inside the angle brackets in the semantic form value of a PRED attribute.

In the examples in (25), the SUBJ function contains a PRED attribute, yet is not semantically governed. In this way, examples (25) fail to satisfy (26).[9]

We can use this principle to make it necessary that the athematic object in (10) *János-t* 'John-ACC' is linked to some function of the embedded clause in the following way. If, as in (22), the athematic object is not linked to any participant in the embedded clause, the semantic Coherence condition is violated.

The possibility of identifying the matrix focus with the embedded subject also follows from independent principles of the grammar. We may assume a rule for regular question formation whereby an element from a possibly embedded clause raises to the preverbal focus position, such as the following:[10]

(28)
FP	→	(NP)	F′
		(↑FOCUS)=↓	↑=↓
		(↑COMP* GF)=↓	

Thus, principles of the grammar (the semantic Coherence condition and the regular rule of question formation) can be made responsible for the identification between the matrix focus and some embedded function.

In fact, these principles alone are insufficient for deriving an f-structure representation of (10) in which the matrix FOC, the matrix OBJ, and the embedded SUBJ are all functionally identified with one another.

[7]Under usual assumptions about the lexical entry for *seem*, the examples in (25) are also ruled out by the requirement that the PRONFORM attribute have the value IT. So, the semantic Coherence condition technically isn't needed to rule these out. However, supposing that a more principled theory of the distribution of *it* vs. *there* were to replace this lexical stipulation, the semantic Coherence condition would be needed in order to rule that sentence out.

[8]It appeared to me at first that the English "copy raising" construction represented below involves an athematic subject that is merely anaphorically integrated: *Richard seems like he's in trouble* (Potsdam and Runner, 2001). Asudeh (2002) gives an analysis of the copy-raising construction in LFG, in which the subject is in fact governed by *like*, so there is no semantic Coherence violation in this type of example according to that analysis.

[9]The semantic Coherence condition may be generalizable in such a way that it applies to discourse functions TOPIC and FOCUS. In fact, such a condition is proposed in Falk (2001), under the name of the Extended Coherence Condition: "All functions in an f-structure must be incorporated into the semantics. Argument funcitons are subject to the Coherence Condition. Overlay functions [incl. focus –EEC] must be identified with arguments or adjuncts; adjuncts must be in f-structures containing PREDs."

[10]See Szendrői (2001) for a good argument for a focus projection, as well as Brody (1990, 1995) for the original suggestion that Hungarian uses a focus projection.

The rule given in (28) allows *either* the embedded subect to be identified with the matrix focus (in which case the functional uncertainty path is instantiated as: COMP SUBJ), *or* the matrix object (in which case the functional uncertainty path is instantiated as: OBJ). If we link the embedded SUBJ to the matrix FOC, another equation is necessary for linking the matrix OBJ into the functional "chain" if you will, either by equating it with the matrix FOC or by equating it with the embedded SUBJ. To allow this, I propose the additional annotation on the focus position in (29).

(29) FP → (NP) F′
(↓CASE)=ACC⇒ (↑OBJ)=↓

In other words, if any constituent is accusative-marked in the focus position of a clause, then it is interpreted as the object of that clause, whether or not it is also interpreted as the object of another clause. This will link the matrix OBJ into the functional chain, even if the focus rule is used to link the embedded subject with the matrix focus, instead of the matrix object.

Independent evidence for this solution comes from the fact that definiteness agreement in focus-raising is not limited to embedded-clause subjects. Any focus-raised phrase with accusative case triggers definiteness agreement with the matrix verb. Thus, phrases corresponding to embedded-clause objects, which retain their definiteness marking from the embedded clause, also trigger definiteness agreement in the matrix clause, as shown in the contrast between (30a) and (30b).

(30) a. Csak ez-t akar-om/*-ok hogy el-mond-j-ad (Kenesei, 1994)
only this-ACC want-1SG.DEF/1SG.INDEF that out-say-SBJ-2SG.DEF
'It's only this that I want you to say.'

b. Csak két dolg-ot akar-ok/*-om hogy el-mond-j-ál
only two thing-ACC want-1SG.INDEF/1SG.DEF that out-say-SUBJUNCTIVE-2SG.INDEF
'There's only two things that I want you to say.'

In (30a), the focus-raised element *ezt* 'this' is definite, so the verb *akarom* 'want' is as well. In (30b), the numerically-quantified NP *két dolgot* 'two things' is indefinite, and so is the matrix verb *akarok* 'want'. If definiteness agreement is sufficient evidence for objecthood, then it appears that embedded clause objects become matrix clause objects when they focus raise as well as embedded clause subjects.

1.3 Expletive construction

If the matrix verb has an athematic object, then that object position should be able to be filled with an expletive. Indeed, that is what we find in examples like (31), I would like to claim.

(31) Azt lát-om, hogy Péter jön
it.ACC see-1.SG.DEF that Peter.NOM come.3.SG
'I see that Peter is coming.'

I will argue that in this construction, the accusative pronoun *azt* is an expletive, filling the athematic OBJ position.

Following Berman (2003) in her analysis of similar constructions in German, I will refer to the construction instantiated in (31) as the "correlative" construction. For such constructions, there are two potential analyses, corresponding to two different types of lexical entries for bridge verbs.

(32) a. Extraposition analysis
The accusative pronoun is the head of the object selected by the verb, and the clause

is contained within the OBJ as a complement to the pronoun, or related to it through apposition or adjunct modification. This analysis may be schematized thus:

$$\begin{bmatrix} \text{PRED} & \text{`say}\langle \text{SUBJ OBJ} \rangle\text{'} \\ \text{OBJ} & \begin{bmatrix} \text{“azt”} & \\ \text{GF} & \left[\text{“hogy Péter jön”}\right] \end{bmatrix} \end{bmatrix}$$

b. Expletive analysis

The accusative pronoun is an athematic object, and the clause is the direct complement of the verb. It is schematized thus:

$$\begin{bmatrix} \text{PRED} & \text{`say}\langle \text{SUBJ COMP} \rangle\text{OBJ'} \\ \text{OBJ} & \left[\text{“azt”}\right] \\ \text{COMP} & \left[\text{“hogy Péter jön”}\right] \end{bmatrix}$$

(The extraposition analysis is the LFG equivalent of the analysis of correlative constructions given by É. Kiss (1990), wherein the clause is extraposed, leaving a trace within the object NP; the expletive analysis is more like the analysis given by Kenesei (1994), in which the accusative pronoun is an expletive and the clause is a direct complement of the matrix verb.) For a summary of the possibilities, either the accusative pronoun (32a) or the clause (32b) is semantically selected by the matrix verb. I offer two arguments for the expletive analysis, wherein the clause is semantically selected.

1.3.1 Possibility of extraction

Kenesei (1994) argues for the expletive analysis on the basis of the focus-raising construction itself, which clearly shows that there must be at least some lexical entry for bridge verbs in which the complement is directly selected by the verb. His argument is as follows. Extraction is not possible from complex noun phrases, or adjunct phrases, and yet it is possible in examples like (2) and (10). If extraction is possible only from argument clauses, then the CP in (10) must be an argument of the verb, not a right-dislocated subconstituent of a complex noun phrase object.

Further evidence against the analysis involving extraposition from a complex noun phrase comes from the complementary distribution of focus-raising constructions with the accusative pronoun. Under the extraposition analysis of (31), the CP *[hogy érkezik]* is a subordinate part of a complex noun phrase even in focus-raising constructions, so an example like (2) has the analysis in (33), with a silent pronoun heading the noun phrase complement of the verb.

(33) János holnap mond-t-a $[_{\text{NP}}\ [_{\text{CP}}$ hogy érkezik]]
John tomorrow say-PAST-3.SGDEF that arrives
‘It was tomorrow that John said it that he would arrive.’

But the silent head of the putative complex noun phrase cannot be spelled out:[11]

[11] É. Kiss (1990) maintains that the extraposition analysis is correct even for focus-raising cases, and proposes to deal with the ungrammaticality of *azt* in (34) in the following way.

> Owing to the Visibility Condition [Chomsky (1981)], sentential arguments are assigned case. A CP category, however, cannot bear case – therefore, it must be formally subordinated to an NP. It is the complex NP that functions as the argument of the V, picking up its case and theta-role. The case will be borne by the pronominal head of the

(34) *János holnap mond-t-a azt hogy érkezik
John tomorrow say-PAST-3SG.DEF it-ACC that arrives
'It was tomorrow that John said it that he would arrive.'

This evidence shows that there is a structural difference between extraposed clauses and clauses from which focus-raising may take place, and therefore suggests that bridge verbs directly select for their clausal complements, rather than being indirectly related to them as in the extraposition analysis.

1.3.2 Lack of direct object NPs

An empirical problem for the extraposition analysis of correlative constructions such as (31) is that it overgenerates direct objects. Because OBJ is a thematic argument of the verb in that analysis, the verb is predicted to be able to take a simple direct object, but normally, this is impossible:

(35) *A hír-t mond-t-a
the news-ACC say-PAST-3SG.DEF
'He said the news.'

The only elements that can appear as the sole noun phrase complement to bridge verbs are the pronouns *az-t* 'that-ACC' and *mi-t* 'what-ACC'.

(36) a. Azt mondta. 'He said it.'
b. Mit mondott? 'What did he say?'

In fact, proper names may appear as the sole overt argument to a bridge verb, as in (37),

(37) János-t mond-t-a.
John-ACC say-PAST-3SGDEF
'He said John.'

but this example has an "elliptical character" and can only be used when the context specifies the content of the elided clausal complement (Katalin Kiss, p.c.).[12]

We can explain the grammaticality of the examples in (36) with the following assumptions: the pronouns *az-t* 'that-ACC' and *mi-t* 'what-ACC' operate as expletives in (36), and their associated clauses undergo CP deletion. The idea that *mit* can operate as an expletive is supported by the "scope-marking" construction, illustrated in (39).

NP. Since the pronominal head is a semantically empty dummy NP, the theta-role will be borne by the clause. The clause inside the NP is formally an adjunct to the pronominal head; consequently, its CP boundary is a blocking category and a barrier for its constituents. The NP dominating the CP inherits barrierhood; thus extraction from the CP is prevented. *If the pronominal head is phonologically empty, it becomes transparent, presumably because* pro, *recoverable from the verbal suffi x, is deleted, and the NP projection, having lost its head, is pruned. This way the CP is directly governed by the V.* [Italics mine]

Kenesei (1994) argues against this analysis on the basis that it violates the Projection Principle: "if the verb is subcategorized for a noun phrase at one level of structure, it cannot have a clausal complement at another" (Kenesei, 1994, pp. 312–313). In LFG, the Visibility Condition is built into the very framework; a verb cannot subcategorize for one grammatical function at one level of representation and a different grammatical function at another, because there is only one level of representation at which grammtical functions are represented, namely f-structure. Thus, this explanation is out of the question for an LFG analysis.

[12]There are other genuine examples of noun phrase complements to bridge verbs, such as (38), where the direct object is a complex noun phrase headed by a pronoun, containing a clause.

(38) Mond-t-´al [$_{NP}$ az-t [$_{CP}$ hogy J´anos j¨on]]
say-PAST-2SG.DEF it-ACC that John come.3SG
'I said it that John is coming.'

Assuming that indeed this example contains a complex noun phrase, we may be forced to assume that matrix verbs allow OBJ complements. I am not able to determine whether or not we are forced to do so at this point.

(39) Mi-t mond-ott, hogy mikor jön?
what-ACC say-3SG.INDEF that when come.3SG
'When did he/she say that he/she would come?'

See Horvath (1995) for an extensive set of arguments that *mit* is behaving as an expletive in the scope marking construction.

Furthermore, although the pronoun *azt* cannot be present in the matrix clause when focus-raising or question formation takes place, it is possible to relativize out of the complement to a bridge verb when the expletive is present. (These were found using Google, but a native Hungarian speaker informant confirms the grammaticality of these examples.)

(40) Az egyik dolog, **amiben**$_i$ azt gondolom, hogy jobbak vagyunk t_i, mint mások ...
the primary thing in which it-ACC think-1SG that better be.1PL than others
'The primary thing in which I think (it) that we are better than others...'
`www.hte.hu/kiadvanyok/hirlev2002/ hirlevel_2002jan01.html`

(41) egy optimális cél, **amivel**$_i$ azt hiszem, hogy mindenki egyetért t_i, ...
an optimistic goal with which it-ACC think-1SG that everyone agrees
'an optimistic goal, with which I think (it) that everyone agrees'
`www.mkogy.hu/naplo34/143/1430020.html`

Similarly, a dative possessor in the embedded clause may also raise past the expletive into the higher clause, as shown in (42a). This is not be possible with complex noun phrases, for example as in (42b).

(42) a. **En-nek**$_i$ is az-t hiszem, hogy [t_i a legfontosabb tulajdonsága] a folytonosság.
this-DAT FOC it-ACC think-1SG that the most_important property the continutity
'I think that the most important property of this is continuity.'

b. ***Ennek**$_i$ is hallot-t-am az ötlet-et, hogy [t_i a legfontosabb tulajdonsága] a folytonosság.
this-DAT FOC hear-PAST-1SG the idea-ACC that the most_important property the continuity
'I heard the idea that the most important property of this is continuity.'

If the correlative construction had the same structure as a complex noun phrase, as the extraposition analysis claims, then (42a) should be just as bad as (42b). The difference in grammaticality suggests a structural difference between the two constructions. Only the expletive analysis contains such a distinction. Hence, the construction in (31) reflects the expletive pronoun that we would expect under the athematic object analysis of *Pétert* in (10).

2 Functional and anaphoric control

In the previous section, I concluded that accusative-marked focus-raised subjects such as *Pétert* in (10) are athematic objects, based on both the long-distance and island-sensitive nature of focus-raising and the the expletive construction. In this respect, the focus-raising construction is like raising rather than equi. As laid out by Bresnan (1982), raising constructions involve functional control, while equi involves anaphoric control. So it makes sense as a default assumption at this point that the focus-raising construction under consideration involves functional control. In fact, I have already implicitly proposed that the identification is functional.

However, Kroeger (1993) documents a type of equi with functional control, in Tagalog, showing that the distinction between raising and equi is orthogonal to the distinction between functional and anaphoric control. This carves out the typological possibility of a raising construction with anaphoric control. In this section, I will argue that the predicted but unattested raising with anaphoric control is in fact attested in one dialect of Hungarian.

2.1 Gervain's observation

As noticed by Gervain (2002), there is some disagreement in the literature on focus-raising as to the obligatoriness of the accusative marker *-t* on the focussed constituent in examples like (10). For some speakers, *Péter* can appear in nominative case with the same interpretation as in (10), shown in (43).

(43) Péter mond-t-am, hogy jön
Péter.NOM say-PAST-1SG that come.3SG
'It is Peter that I said is coming.'

The speakers who allow (43) also allow (10). Yet, the speakers who allow only (10) are more liberal than the others in a different respect: they don't require that the embedded verb (*jön* 'come') agree in number with the focussed constituent in certain cases, whereas the others do. This difference can be observed when a semantically plural, but syntactically singular phrase such as *az összes lány* 'every girl' in (44) is raised. That *az összes lány* is grammatically singular is evident from its nominal morphology, which is singular, and the fact that it agrees with a singular verb, as shown in (44).

(44) Az összes lány jön/*jön-nek.
the all girl come.3SG/come-3PL
'All the girls are coming.'

Az összes lány obviously picks out a plurality of girls and is in that sense semantically plural; for further evidence, the plural personal pronoun *ők* 'they' could later be used to refer back to it, as the subject of the next sentence, for example.

When *az összes lány* is focus-raised, there are, in principle, four different possible outcomes, depending on two binary choices. The raised phrase may have nominative case, as in (45), or accusative case, as in (46). Next, the embedded verb 'come' may be either plural, as in the (b) examples of (45) and (46), or singular, as in the (a) examples. As Gervain (2002) shows, not all possibilities are grammatical, and there is some disagreement among Hungarian speakers as to which of them are. In her study, Hungarian speakers fell into two groups (A and B); their judgements for each sentence are represented in the righthand columns in (45)–(46).

(45) NOMinative-marking

a.	Az összes **lány** mond-t-ad, hogy **jön** the all girl.NOM say-PAST-2SG that come.3SG	✓A	*B
b.	Az összes **lány** mond-t-ad, hogy **jön-nek** the all girl.NOM say-PAST-2SG that come-3PL	*A	*B

(46) ACCusative-marking

a.	Az összes **lány-t** mond-t-ad, hogy **jön** the all girl-ACC say-PAST-2SG that come-3SG	✓A	✓B
b.	Az összes **lány-t** mond-t-ad, hogy **jön-nek** the all girl-ACC say-PAST-2SG that come-3PL	*A	✓B

For Group A, both nominative and accusative case are allowed on the focus-raised element, and number agreement between the focus-raised element and the embedded verb is required. For Group B, only accusative case is allowed on the focus-raised element, but number agreement is not required; Group B accepts

(46b), in which the embedded verb is plural but the focus-raised element *az összes lány* is grammatically, though not semantically, singular.[13] The pattern is summarized below.

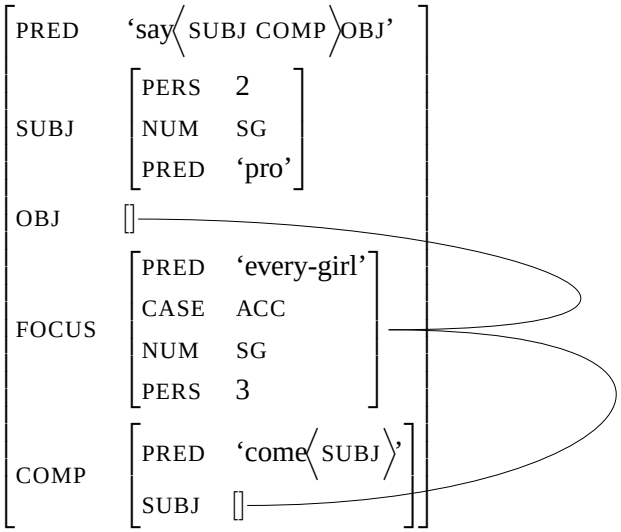
(47)

	accusative-marking	agreement (embedded verb is singular)
Group A	optional	required
Group B	required	optional

2.2 Functional and anaphoric control

This pattern suggests that while Group A uses functional identification between the matrix focus and the embedded subject as discussed in the previous section, Group B uses anaphoric identification. The Group A pattern can be accounted for under the assumption of functional identification as follows. The examples with number disagreement, (45b) and (46b) are ruled out because the matrix focus is functionally identified with the embedded subject, so the number features will clash. (46a) has an analysis very similar to the one given for (10), ignoring the representation of the quantifier.

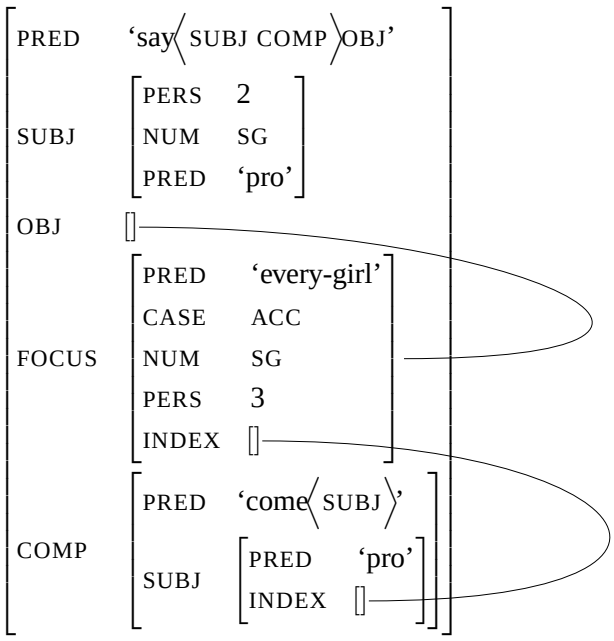
(48)



For Group B, the possibility of non-agreement between the raised accusative element and the embedded subject in (46b) suggests that the identification is, by contrast, anaphoric, for that group.

Under these assumptions, here is an analysis of (46a):

[13] In fact, the judgments were given on a 5-point scale; I have simplifi ed '?' to unmarked, and '???' to '*'. The average judgments for each group were not all stars or "perfect" marks; it remains an open question how to account for the more subtle aspects of the pattern. See Gervain (2002, 2003) for discussion.

(49)


Gervain (2002) proposes an analogous analysis in Minimalist style. She argues that Group B uses a "resumptive strategy", that is, they use a resumptive pronoun in the embedded clause, whereas Group A uses regular *wh*-movement. This explains the fact that both plural and singular agreement are possible for Group B. Anaphors with semantically plural antecedents may be either plural or singular, giving rise to a collective or distributive interpretation respectively. Thus the singular pronoun *őt* in (50a) has a distributive interpretation, and the plural pronoun *őket* in (50b) has a plural interpretation.

(50) a. Két fiú$_i$ hiszi azt, hogy Mária szereti (őt$_i$)
two boy.SG think.3SG it that Mary love.3SG him
'There are two boys each of whom thinks that Mary loves him.'

b. Két fiú$_i$ hiszi azt, hogy Mária szereti őket$_i$
two boy.SG think.3SG it that Mary love.3SG them
'Two boys think that Mary loves them.'

As Gervain points out, the idea that Group B uses a pronoun in the embedded clause is also supported by the fact that Group B does not rule out focus-raising out of complex noun phrases:

(51) Az elnököt mondtad, hogy hallottad a hírt, hogy megérkezett (ok for Group B only)
the president.ACC said.2SG that heard.2SG the news that arrived.3SG
'You said you heard the news that *the president* had arrived.'

whereas Group A does. This is in line with the fact that movement out of complex noun phrases is impossible, but anaphoric binding into them is possible, as shown in (52).

(52) No man$_i$ asked whether or not he$_i$ would be fired.
(cf. **Who$_i$ asked whether or not t_i would be fired?*)

In LFG terms, the movement/resumption distinction may be interpreted using the functional/anaphoric identification distinction.

However, supposing that Group B uses anaphoric identification does not explain why Group B, unlike Group A, does not tolerate nominative case on the focus-raised element, as shown in (45a). To explain this,

I would like to propose that the complementizer *hogy*, for Group B, has a specification similar to that which Falk (2001) gives of English *that*, in order to explain the *that*-trace effect:[14]

(55) *hogy* C $(\uparrow \text{SUBJ}) \neq ((\text{GF}^{+} \uparrow)\ \text{GF})$

This solution echoes the explanation given by Gervain (2002) for the impossibility of nominative case: she writes, "I assume that it is the nature of the complementizer *hogy* 'that' that is parametrically different in the two dialects. In the movement dialect, the complementizer allows the raised operator to properly govern its trace, thus complying with the ECP. In the other dialect, however, the complementizer is 'opaque' and does not allow proper government. Thus resumption is needed to save the sentence from the resulting ECP violation" (Gervain, 2002, p. 77). The difference here is that the idea of the "Last Resort" is being modelled with an OT constraint violation.

This constraint prevents subject extraction, thereby preventing nominative focus-marked constituents. At the same time, it motivates the possibility of anaphoric identification between the raised element and the embedded subject.

Notice that in (49), there is an athematic object that is not functionally integrated, so the semantic Coherence condition is violated: there is a PRED inside a governable function that is not semantically selected. This particular violation of the semantic Coherence condition does not make us want to suppose that the semantic Coherence condition does not hold; it is still important for ruling out the Hungarian equivalent of examples like **John rained*. Rather, the semantic Coherence condition can be violated only in a limited range of cases. In particular, the semantic coherence condition can be violated when a violation of the constraint in (55) would be incurred to satisfy it. Thus, the constraint in (55) outranks the semantic Coherence condition, in the OT sense.

Let us label the semantic Coherence condition SEMANTIC-COHERENCE, and the constraint in (55), *SUBJECT-EXTRACTION. With these constraints, it is possible succinctly to describe the pattern associated with both groups as follows: Group A uses the constraint subhierarchy (56); Group B uses (57).

(56) SEMANTIC-COHERENCE $\gg$ *SUBJECT-EXTRACTION

(57) *SUBJECT-EXTRACTION $\gg$ SEMANTIC-COHERENCE

Hypothesizing these constraints thus allows us to explain both of the differences between Group A and Group B as consequences of a single difference in constraint ranking.[15]

[14]The presence of this constraint in Hungarian is ironic, because as É. Kiss (1987) points out, on the surface Hungarian has the opposite of the English *that*-trace effect:

(53) a. Kit mondtam, hogy j¨on?
b. *Kit mondtam, j¨on?

(54) a. *Who did you say that is coming?
b. Who did you say is coming?

[15]These are not all the constraints that are relevant. Now that the semantic Coherence condition is violable, a constraint is needed to rule out (22), which is ungrammatical in both dialects. To account for this, I might propose that the Extended Coherence condition, a more specifi c condition than semantic Coherence, is highly-ranked for both Group A and Group B.

(58) Extended Coherence: FOCUS and TOPIC must be linked to the semantic predicate argument structure of the sentence in which they occur, either by functionally or by anaphorically binding an argument. Bresnan and Mchombo (1987)

Alternatively, the Extended Coherence Condition may be a by-product of GEN; if the OT-LFG Input is semantically well-formed, it is even diffi cult to imagine a function producing candidate f-structures from the input that violate the Extended Coherence Condition.

Also, as was pointed out by Peter Sells at the LFG03 conference, another constraint to the effect, FOCUS-RAISE!, needs to be among those relevant for evaluation.

Conclusions

I have argued that accusative-marked focus raised subject such as *Pétert* in (10) occupy an athematic object function. Further, depending on dialect, this function may host either functional or anaphoric binders for the embedded subject. If this analysis is correct, then the semantic Coherence condition is violable. This violability may be exploitable in the explanation of other intrusive pronoun phenomena, such as resumptive pronouns.[16]

The analysis has certain typological implications as well. First, if it is true that the accusative-marked focus-raised phrases are objects, then they represent long-distance (A') movement to an argument (A) position, which is not usually assumed to exist. Second, if the analysis is correct, then Hungarian fills out a typological paradigm slot that is predicted to exist, but not yet attested: raising with anaphoric control.

References

Asudeh, A. (2002). Richard III. In Mary Andronis, Erin Debenport, A. P. and Yoshimura, K., editors, *CLS 38: The main session. Papers from the 38th meeting of the Chicago Linguistic Society*, volume 1, Chicago, IL. Chicago Linguistic Society.

Bartos, H. (1997). On "subjective" and "objective" agreement in Hungarian. *Acta Linguistica Hungarica*, 44:363–384.

Berman, J. (2003). *Clausal Syntax of German*. Studies in Constraint-Based Lexicalism. CSLI Publications.

Bresnan, J. (1982). Control and complementation. In Bresnan, J., editor, *The Mental Representation of Grammatical Relations*, pages 282–390. MIT Press.

Bresnan, J. and Mchombo, S. (1987). Topic, pronoun, and agreement in Chicheŵa. *Language*, 63:741–82.

Brody, M. (1990). Remarks on the order of elements in the Hungarian focus field. In Kenesei, I., editor, *Approaches to Hungarian*, volume 3, Szeged. JATE.

Brody, M. (1995). Focus and checking theory. In Kenesei, I., editor, *Approaches to Hungarian*, pages 31–43, Szeged. JATE.

Chomsky, N. (1981). *Barriers*. MIT Press, Cambridge.

Dalrymple, M. (2001). *Lexical Functional Grammar*. Academic Press, San Diego, CA.

É. Kiss, K. (1987). *Configurationality in Hungarian*. Akadémiai Kiado.

É. Kiss, K. (1990). Why noun-complement clauses are barriers. In Mascaró, J. and Nespor, M., editors, *Grammar in Progress*, pages 265–277. Foris, Dordrecht.

É. Kiss, K. (2002). *The Syntax of Hungarian*. Cambridge University Press.

Falk, Y. (2001). *Lexical-Functional Grammar: An introduction to Parallel Constraint-Based Syntax*. CSLI Publications.

[16] Copy-raising constructions like *Richard seems like he's in trouble* (Potsdam and Runner, 2001) seem like good candidates, because the controller appears to be in an argument slot that is not thematic. However, Asudeh (2002) analyzes these nicely as involving assignment of a thematic role to the controller by the preposition *like*. If we take his analysis, then the semantic Coherence condition is not violated in copy-raising constructions.

Gervain, J. (2002). Linguistic methodology and microvariation in language: the case of operator-raising in Hungarian. Master's thesis, University of Szeged, Hungary.

Gervain, J. (2003). Újra a focuszemelésról: A rézumptív névmások térmésete. Master's thesis, University of Szeged.

Horvath, J. (1986). *Focus in the theory of grammar and the syntax of Hungarian*. Foris, Dordrecht.

Horvath, J. (1995). Partial *wh*-movement and *wh* "scope-markers". In Kenesei, I., editor, *Approaches to Hungarian*, volume 5.

Horvath, J. (1998). Multiple wh-phrases and the wh-scope-marker strategy in Hungarian interrogatives. *Acta Linguistica Hungarica*, 45:31–60.

Kenesei, I. (1994). Subordinate clauses. In Kiefer, F. and É. Kiss, K., editors, *The Syntactic Structure of Hungarian*, volume 27 of *Syntax and Semantics*. Academic Press.

Kroeger, P. (1993). *Phrase Structure and Grammatical Relations in Tagalog*. PhD thesis, Stanford University.

Lipták, A. (2001). *On the Syntax of* Wh-*Items in Hungarian*. PhD thesis, Leiden University.

Marácz, L. (1989). *Asymmetries in Hungarian*. PhD thesis, University of Groningen. §7.2–§7.4, pp. 297–325.

Potsdam, E. and Runner, J. T. (2001). Richard returns: Copy raising and its implications. In Mary Androis, Chris Ball, H. E. and Neuvel, S., editors, *CLS 37: The main session. Papers from the 37th meeting of the Chicago Lingusitic Society*, volume 1, Chicago, IL. Chicago Linguistic Society.

Szendrői, K. (2001). *Focus and the syntax-phonology interface*. PhD thesis, University College London.

Zolnay, G. (1926). Mondatátszövődés. *Értekezések a Magyar Tudomanyos Akadémia Nyelv- és Széptudományi Osztálya Köréből*, 23.