

Wh-coordination in Hungarian and Romanian multiple questions

Gabriela Bîlbîie[†] and Anna Gazdik^{†‡*}

[†]LLF – Université Paris Diderot - Paris 7,

[‡]Research Institute for Linguistics, Hungarian Academy of Sciences

1. Introduction

According to one definition, *coordination* is a syntactic operation that usually includes the presence of a conjunction and is based on some repetitive mechanism (Sag et al. 1985). However, the exact nature of this repetitive mechanism has not always been clear. Nowadays, it is generally accepted that the conjuncts must be categorially/functionally similar and they must share a superordinate semantic concept they all can be seen as instances of. The assumption that categorial identity is not always realized is easy to show (based on Sag et al. 1985):

- (1) a. Pat is [a Republican]_{NP} and [proud of it]_{AP}.
b. Pat is [either asleep]_{AP} or [at the office]_{PP}.
c. That was [a rude remark]_{NP} and [in very bad taste]_{PP}.

As these examples illustrate, the syntactic parallelism is weak concerning category and morphosyntactic features like gender, number, mood, etc. However, syntactic parallelism is strong when it comes to the syntactic function of the conjuncts: in the above examples, although the conjuncts are not of the same category, they share the same function (they are all predicative complements of the copula). These observations can be captured by the so-called *Wasow's Generalization*: for a coordination to be well-formed, each conjunct must independently meet the constraints imposed by the shared material. Another constraint concerning coordination constructions applies at the semantic level: the conjuncts must also share the same semantic role. In the following example, although each conjunct would form a grammatical sentence on its own (since the verb is optionally transitive), the sentence is ungrammatical, as they have different semantic roles:

- (2) *John eats [an apple and at midnight].

However, there are some specific structures that apparently violate the strong syntactic and semantic parallelism required in coordination (see examples in Chaves and Paperno 2007). One of these structures is the case of *wh*-phrases,¹ illustrated in (3) for two typologically unrelated

*We would like to thank Anne Abeillé, Olivier Bonami, Jean-Marie Marandin, François Mouret, and the reviewers of EISS 9 for their valuable comments on our previous work on this topic and on the paper itself. Needless to say, only we are to blame for any errors, misconceptions, and unclaritys.

¹For the sake of simplicity, we illustrate this phenomenon with the help of binary examples (in which two conjuncts are coordinated). The coordination of more than two conjuncts works the same way. Throughout the paper, we only talk about 'bare' *wh*-phrases, but the proposed generalizations seem to hold for complex *wh*-phrases as well.

languages: Hungarian (H) and Romanian (R).² Since in these structures the conjuncts apparently do not share the same syntactic function and same semantic role, they are sometimes referred to as *hybrid coordination* (Chaves and Paperno 2007).

- (3) a. **Ki** (és) **mikor** jött? (H)
who (and) when came
b. **Cine** (și) **când** a venit? (R)
who (and) when has come

The paper is structured as follows. In §2 we present some general observations about *wh*-coordination in Hungarian and Romanian multiple questions. In §3 we examine the syntactic structure of these constructions in the two languages. In §4 we move on to our proposed analysis that we formalize within the framework of Head-driven Phrase Structure Grammar (HPSG), followed by a conclusion in §5.

2. The data: generalities

In this section, we introduce some basic properties of *wh*-coordination with respect to the two languages examined. These include the comparison of coordinate structures with other multiple question structures, and, crucially, the problem of *wh*-word order, which creates a lot of confusion in the speakers' acceptability judgments. Although in this paper we mostly used examples that can be found on the internet and in corpora, and we checked their acceptability with native speakers by using questionnaires (presenting each item in an appropriate context), we experienced considerable hesitation in these judgments. The acceptability of these examples thus clearly necessitates further experimental studies, which was beyond the scope of this paper.

2.1. Coordination and other multiple question structures

Multiple questions involving the presence of a conjunction should clearly be distinguished from those without a conjunction, since they do not have the same properties. Those containing a conjunction can further be divided into two groups, based on the possibility for one of the *wh*-phrases to be 'stranded' at the end of the sentence.

2.1.1. Coordinate and 'paratactic' multiple questions

Concerning the examples in (3), we should note that the presence of the conjunction is not obligatory. In the languages used as illustration above, where both 'paratactic' *wh*-phrases (henceforth **paratactic-wh**)³ such as (4), and coordinated *wh*-phrases (henceforth **coord-wh**) such as (5), are possible in multiple questions, the two patterns usually have different interpretational properties.

²In this paper, we concentrate on the coordination of *wh*-phrases in Hungarian (see also Bánréti 2007, and Lipták 2001, 2003) and Romanian (see also Comorovski 1996, and Rațiu 2011), keeping in mind that this structure is possible in other languages as well: Vlach (Merchant 2008), and in the Slavic languages (Chaves and Paperno 2007, and Gribanova 2009 on Russian, Skrabalova 2006 on Czech, and see Kliaschuk 2008, Citko and Gračanin-Yuksek 2012, and Tomaszewicz 2011 for cross-linguistic analyses).

³I.e. The *wh*-phrases are cumulated in the preverbal domain without a conjunction.

- (4) a. **Ki mit** eszik karácsonykor? (H)
 who what eats Christmas-at
 ‘Who eats what at Christmas?’
- b. **Cine ce** mănâncă de Crăciun? (R)
 who what eats at Christmas
- c. Answer: ‘We usually eat fish, our neighbours cabbage, and our colleagues turkey.’
- (5) a. **Ki és mit** tett a levesbe? (H)
 who and what put the soup-into
 ‘Who put something into the soup and what was it?’
- b. **Cine și ce** a pus în supă? (R)
 who and what has put into soup
- c. Answer: ‘It was me and I added some curry to it.’

As can be seen from the examples, paratactic-*wh* as in (4) usually licenses so-called *pair-list answers*, whereas coord-*wh* as in (5) are often argued to license so-called *single pair answers*. Ginzburg and Sag (2000) argue that the pair-list reading is the default reading of multiple questions, and the single pair reading is always related to some additional mechanism, like a *uniqueness presupposition*. Single pair answers are usually given to questions that refer to unique events and inquire about more than one participant of that event. It is, however, also possible that the uniqueness presupposition is cancelled in the answer, since the question has to be divided into subquestions (see Büring 2003), and the answer provides a number of subevents of the main events (Jean-Marie Marandin, László Kálmán, p.c.):

- (6) a. **When** and **where** were the children examined?
 b. Answer 1: On Monday in the school.
 c. Answer 2: *The girls* on Monday at school and *the boys* on Friday at home.

The second distinguishing property between paratactic and coord-*wh* is related to the first. It is usually assumed that *wh*-phrases in paratactic-*wh* obey strict ordering constraints influencing the interpretation of the question, whereas there is no strict syntactic constraint on the order of *wh*-phrases in coord-*wh*, or at least the different orders do not trigger significant interpretational differences. All the orders do not seem to be possible though (see §2.2). Furthermore, in coord-*wh* the first *wh*-phrase can constitute an independent intonational unit, which is not the case with the initial *wh*-phrase in paratactic-*wh*. The above considerations thus motivate the claim that coordinate and paratactic multiple questions are not simple variants of the same structure, but reveal different structures necessitating different analyses.

2.1.2. Coord-*wh* and end-attach-*wh* multiple questions

There is another multiple question structure which has not yet been illustrated. In this type, one *wh*-phrase is in an initial position, whereas the other is ‘stranded’ at the end of the sentence (henceforth **end-attach-*wh***).

- (7) **Ki** jött, és **mikor**? (H) (8) **Cine** a venit, și **când**? (R)
 who came, and when who has come, and when

End-attach-*wh* shares some properties with coord-*wh*. Apart from the fact that the ‘stranded’ *wh*-phrase is marked by a conjunction and it can constitute an independent intonation unit, end-

attach-*wh* usually licenses a single-pair reading. It is also our aim in this paper to examine to what extent end-attach-*wh* can be analyzed in the same way as coord-*wh* in the two languages.

2.2. Asymmetries and acceptability judgments

It is an important observation concerning multiple questions that there is considerable variation in their acceptability among speakers, not independently of the complex interpretational patterns sketched above. Speakers' intuitions are uncertain as regards both (i) the available semantic interpretations (pair-list versus single-pair) and (ii) the ordering variations.⁴

For instance, Comorovski (1996) and Rațiu (2011) consider that, contrary to what happens in paratactic-*wh*, no ordering restriction appears to hold between the *wh*-phrases in coord-*wh* in Romanian. That is, we can switch the order of the *wh*-phrases without any change in acceptability. However, we have found that there are indeed ordering constraints which affect the acceptability.

2.2.1. The main/subordinate clause asymmetry

The first asymmetry that has been observed (Gazdik 2011) is that coordinate multiple questions are sometimes much more natural in embedded clauses than as main clause questions:

- (9) a. Még holnap egyeztetünk, hogy **mikor** és **hol** kéne találkozni. (H)
 still tomorrow agree that when and where should meet
 'We will discuss tomorrow when and where we will meet.'
- b. Te rog să ai grijă **cui** și **ce** spui! (R)
 please have.SUBJ attention who.DAT and what.ACC tell.2
 'Please be careful what you tell and to whom!'

A possible explanation might be found in the fact that these questions usually refer to unique events, and this is why it is more felicitous to ask a question about this event if it has already been introduced into the discourse. One way of signaling this is exactly to introduce it with a main clause and embed the question in the subordinate clause. This problem still needs further investigation though.

It is interesting to note that the main/subordinate clause asymmetry can be observed in the so-called correlative coordination of *wh*-phrases as well. For instance, in Romanian, the use of correlative items *și...și* 'both...and' in main *wh*-coordination, as in (10b), yields an ungrammatical sentence, while it is quite acceptable in embedded clauses such as (10a).⁵

- (10) a. Vreau să știu *și cine, și ce* a spus despre mine.
 want know.SUBJ CORREL who, CORREL what has told about me
 'I want to know both who told and what he told about me.'
- b. **Și cine, și ce* a spus despre mine?
 CORREL who, CORREL what has told about me

⁴The different judgments referred to in the literature are contradictory. For example, Romanian (i) is considered acceptable by Rațiu 2011, while it appears ungrammatical in Kliashuk 2008.

(i) **Ce** și **cine** a cumpărat? (R)
 what and who has bought

⁵Rațiu (2011) ignores grammatical uses of correlative items in subordinate clauses.

2.2.2. The argument/adjunct asymmetry

Apart from the main/subordinate clause differences, the argument/adjunct asymmetry can also play a role. This evokes the insights of Browne (1972) and Lewis et al. (2012), who show that the possibility of coord-*wh* in English is related to the argument/adjunct asymmetry and in particular to the syntactic behaviour of the verbal head (i.e. if it selects for an optional or an obligatory complement). According to experimental studies of Lewis et al. (2012), obligatorily transitive verbs, like *fix* in (11a), are often unacceptable if their subcategorized complement is in the left conjunct, whereas optionally transitive verbs, like *eat* in (11b), are more acceptable in this context. On the other hand, when the left conjunct is an adjunct as in (11c), the transitivity asymmetry seems to disappear.

- (11) a. ***What** and **when** did John fix?
 b. **What** and **when** did John eat?
 c. ...the mechanic decided **when** and **what** {he could eat / he would fix}...

Returning to Hungarian and Romanian, there are four syntactic patterns of coord-*wh*: (a) adjunct–adjunct, (b) adjunct–argument, (c) argument–adjunct, and (d) argument–argument, illustrated in (12) for Hungarian, and in (13) for Romanian.

- | | |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>(12) a. Mikor és miért ment el?
 when and why left.3 PRT
 b. Hol és mit ettél?
 where and what ate.2
 c. Ki és miért ment el?
 who and why left PRT
 d. Ki és mit választott?
 who and what chose</p> | <p>(13) a. Unde și când pleacă?
 where and when go.3?
 b. Unde și ce mâncăm?
 where and what eat.4
 c. Cine și când pleacă?
 who and when go?
 d. Cine și ce a cucerit?
 who and what has conquered</p> |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

The first pattern does not seem to pose any acceptability problem; the other three patterns, however, are subject to more or less variation in acceptability judgments, especially in Romanian, even though all available in both languages. At this stage, we believe that part of this variation is related to the fact that both Hungarian and Romanian are *pro*-drop languages, allowing the dropping of some arguments (subject or complements). This may help us to explain the difference in acceptability between (14b) and (15b) in Romanian: although both verbs *a spune* ‘to tell’ and *a locui* ‘to live’ can have a dropped subject, they do not have the same behaviour with the complement they select for; while the former allows the non-realization of its direct complement, the latter requires an overt locative complement.

- | | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>(14) a. Cine și ce ți-a spus?
 who and what CL.2-has told
 b. Ce și cine ți-a spus?
 what and who CL.2-has told</p> | <p>(15) a. Cine și unde locuiește?
 who and where lives
 b. *Unde și cine locuiește?
 where and who lives</p> |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

The transitivity asymmetry observed in English obviously plays role in end-attach-*wh* in both languages under examination. In these contexts, the syntactic behaviour of the verbal predicate affects the ordering of *wh*-phrases. For example in Romanian (16), the verbal head *a mânca* ‘to eat’ can have an absolute use, without any overt complement, which could explain why inverted orders of *wh*-phrases are acceptable. On the other hand, in (17) the verb *a se baza* ‘to rely

on' requires an overt prepositional complement in the clause, which cannot thus be stranded in an end-attach-*wh* structure.

- | | | | |
|------|--------------------------------------------------------------------------|------|-----------------------------------------------------------------------------|
| (16) | a. Ce ai mâncat, și unde ?
what have.2 eaten, and where | (17) | a. Pe cine te bazezi, și de ce ?
whom rely-on.2, and why |
| | b. Unde ai mâncat, și ce ?
where have.2 eaten, and what | | b. * De ce te bazezi, și pe cine ?
why rely-on.2, and whom |

2.2.3. [+Human] subject first

At the semantic level, in both languages the order of the *wh*-phrases usually follows some universal animacy hierarchy as regards the position of subject-*wh*-phrases. In Hungarian, if the *wh*-phrase *ki* 'who' is present in the question, it has to be the initial *wh*-phrase, otherwise the order is free:

- | | | | |
|------|----------------------------------------------------------------|------|-----------------------------------------------------------------|
| (18) | a. Ki és mit választott?
who and what chose | (19) | a. Ki és miért ment el?
who and why left PRT |
| | b. * Mit és ki választott?
what and who chose | | b. * Miért és ki ment el?
why and who left PRT |

The same preference for human subject in first position is observed in Romanian too, at least in argument-argument combinations; the reverse order is only acceptable in a given context and with a specific prosody, but not in out-of-the-blue utterances.

- (20) a. **Cine** și **ce** a văzut la Victor Ponta pentru a-l alege în funcția unui partid
who and what has seen at Victor Ponta ...
aflat într-o profundă criză?
'Who saw what in Victor Ponta to elect him head of a party in a deep crisis?'
b. ??**Ce** și **cine** a văzut la Victor Ponta ...?
what and who has seen at Victor Ponta ...

A more rigid constraint concerns cases in which both *wh*-phrases bear a [+human] feature: insofar as the order of [+human subject] with respect to [+human complement/adjunct] is concerned, two groups of Romanian speakers have been identified: there is one group who dislikes the presence of a conjunction between these *wh*-phrases, while the other group accepts it. However, neither group accepts the inverted order in these cases.

- (21) a. **Cine** (??și) **pe cine** a văzut?
who (and) whom has seen
b. ***Pe cine** și **cine** a văzut?
whom and who has seen

In both languages, end-attach-*wh* is infelicitous if the sentence-final *wh*-phrase is the subject, as in (22) and (23). Two possible explanations are at hand for the infelicity of such examples. On the one hand, it can be considered reminiscent of the animacy hierarchy referred to above. On the other, cataphoric dependency (the fact that an element is presupposed or referred to in the sentence, before it is actually introduced) is generally not admitted in coordinate structures (unlike in subordination), cf. the 'counter-indefiniteness' effect (Kayne and Pollock 2001).

(22) ??**Miért** hívott és **ki**? (H)
 why called and who

(23) ??**Ce** a mâncat, și **cine**? (R)
 what has eaten and who

Moreover, in both languages, coordinate structures are infelicitous if the verbal predicate involves a *symmetric* or *reversible* semantic relation:

(24) a. ***Ki** és **kivel** játszott? (H)
 who and who.INSTR played
 b. **Cine** (??și) **cu cine** s-a
 who and with whom REFL-has
 întâlnit? (R)
 met

(25) a. ***Ki** és **kit** ütött? (H)
 who and whom hit
 b. **Cine** (??și) **pe cine** a lovit?
 who and whom has hit
 (R)

Interestingly, in these contexts, Romanian speakers prefer the ‘paratactic’ structure. This means that contrary to the general assumption, paratactic-*wh* can license single pair answers in special cases. The type of answer is again related to a unique event, and the question either refers to its participants, as in (26), or even the identity of the participants can be known, and then the question refers to their respective roles in the event, as in (27).

(26) **Cine cu cine** s-a întâlnit?
 who with whom REFL-has met

(27) **Cine pe cine** a lovit?
 who whom has hit

In Hungarian, although paratactic-*wh* containing these question words is grammatical, it does not mean the same, as it exclusively licenses pair-list answers, which presupposes that there is more than one meeting or hitting event. This is why a fourth type of syntactic structure is preferred in the case of symmetric or reversible predicates: in this type, one *wh*-phrase appears sentence-initially, whereas the other remains in its canonical position, as in (28) and (29). This kind of structure is also possible in Romanian, as in (30) and (31).

(28) **Ki** játszott **kivel**?
 who played who.INSTR

(29) **Ki** ütött **kit**? (H)
 who hit whom

(30) **Cine** s-a întâlnit **cu cine**?
 who REFL-has met with whom

(31) **Cine** a lovit **pe cine**? (R)
 who has hit whom

To conclude this section, we have shown that, contrary to what is often assumed in the literature, we do find some ordering constraints in *wh*-coordination, even if the interplay of the constraints governing these orderings is not always clear to us. Furthermore, a gradience can be observed in the acceptability of the data, which means that we can find intermediate examples between fully acceptable and ungrammatical ones. However, cumulativity of violated constraints renders intermediate examples ungrammatical (e.g. the combined violation of semantic and syntactic constraints in Romanian, as shown in (32)).

(32) a. **Ce** și **cine** ți-a spus?
 what and who CL.2-has told
 b. ??**Ce** și **cine** susține?
 what and who defends
 c. ***Ce** și **cine** consumă dimineața?
 what and who eats morning.DEF

Finally, it is also possible that the variation observed with some orderings be explained as a kind

of syntactic priming: the stricter orderings in paratactic-*wh* have an impact on the speakers' preferences as regards the *wh*-combinations in coordination structures, and this influences the acceptability judgments. This may help us to explain why all 10 Romanian subjects participating in a 5-point rating scale task ranked the sentence in (33) above 4 (high degree of acceptability), and the reversed order in (34) below 3 (low degree of acceptability). The issue necessitates further experimental investigations (e.g. experiments with a speeded acceptability judgment task).

- (33) **Cine** și **ce** a văzut? (34) **Ce** și **cine** a văzut?
 who and what has seen what and who has seen

3. Background of the analysis

In this section, we examine the syntactic structure of the already presented constructions. More precisely, our aim is to determine whether they are monoclausal or biclausal.

3.1. Previous analyses

Generally, previous analyses proposed for coordinate multiple questions treat them as either monoclausal or biclausal structures, but they usually assume (except for Citko and Gračanin-Yuksekić 2012, and Haida and Repp 2011) that one type of analysis can be cross-linguistically valid. According to the monoclausal approaches (Kazenin 2002, Lipták 2001, 2003, Skrabalova 2006, and Gribanova 2009), parallel to the analysis of paratactic-*wh*, the *wh*-phrases are fronted by movement to the left periphery of the same clause and then a 'spurious coordinator' is inserted between the *wh*-phrases, as illustrated in (35a). Other analyses assume that coordinate multiple questions are in fact biclausal, as schematized in (35b).

- (35) a. [_{CP} [&P *wh*₁ **conj** *wh*₂] [_{TP} *t*₁ ... *t*₁]]
 b. [&P [_{CP} *wh*₁ [_{TP} *t*₁ ...]] **conj** [_{CP} *wh*₂ [_{TP} ... *t*₂]]]

There are two different biclausal analyses: one involving syntactic ellipsis (Bánréti 2007, Giannakidou and Merchant 1998, Merchant 2008, and Tomaszewicz 2011), which assumes that coordinate multiple questions contain standard clausal coordination accompanied by the ellipsis of all the material except for the *wh*-phrase itself in one of the conjuncts (i.e. *backwards sluicing*). According to the other approach, coordinate multiple questions involve multidominance, characterized by a single element being shared between two mother nodes (Citko and Gračanin-Yuksekić 2012; Rațiu 2011). In what follows, we examine the syntax of these structures in Hungarian and Romanian separately.

3.2. The syntactic structure of Romanian coord-*wh* and end-attach-*wh*

We will argue that both coord-*wh* (36) and end-attach-*wh* (37) are biclausal in Romanian.

- (36) [**Cine** –]_S [și **când** a venit]_S?⁶ (37) [**Cine** a venit]_S [și **când** –]_S?
 who – and when has come who has come and when –

Conclusive evidence in favour of the biclausal analysis comes from the distribution of the

⁶Note that the adjunct in the right conjunct can be analyzed either as an integrated or as an incidental adjunct. In the last case, the structure would be [**Cine** [și **când**–]_S a venit]_S.

interrogative particle *oare* (Rațiu 2011). This particle optionally occurs in interrogative clauses in a relatively free distribution, as in (38a), and usually only appears once per clause, as in (38b). However, as Rațiu (2011) convincingly shows, the interrogative particle *oare* can appear only once per clause with paratactic-*wh*, as in (39), while with coord-*wh* and end-attach-*wh*, it can co-occur with each *wh*-phrase, as in (40a)–(40b) and in (40c), respectively. Thus, each of these two structures behaves like a coordination of two interrogative clauses.

- (38) a. (*Oare*) Cine (*oare*) vine?
 (PRT) who (PRT) comes
 b. *Oare* cine (**oare*) vine?
 PRT who (PRT) comes
- (39) a. *Oare cine* (**oare*) *ce* zice?
 PRT who (PRT) what says
 b. **Cine** (**oare*) **ce** *oare* zice?
 who (PRT) what PRT says
- (40) a. *Oare cine* și *oare ce* va spune?
 PRT who and PRT what will say
 b. **Cine** *oare* și **ce** *oare* va spune?
 who PRT and what PRT will say
 c. *Oare cine* va veni, și *oare când* ?
 PRT who will come, and PRT when

Monoclausal analyses of coord-*wh* usually stipulate that the conjunction is a semantically spurious element (see Merchant 2008). However, in Romanian, coord-*wh* are compatible with other conjunctive items (41) apart from the conjunction *și* ‘and’. If the conjunctive *and* is assumed to have some spurious uses (where it is semantically contentless), it is difficult to extend this analysis to the other coordinators, which seem to always be semantically contentful. More interestingly, the list of available conjunctions in Romanian multiple *wh*-questions contains the conjunction *iar* ‘and’ (41c), which is reserved for contrastive clausal coordination (Bîlbîie 2011).

- (41) a. Nu vreau să mi se spună **când** *sau* **cât** trebuie să mănânc.
 NEG want.1 tell.SUBJ.3 me when or how much must eat.SUBJ.1
 ‘I don’t like being told when or how much I have to eat.’
 b. Mă interesează nu **cine**, *ci* **ce** a făcut.
 me interests not who but what has done
 ‘I’m not interested in who did it, but in what he did.’
 c. Vreau să știu mai întâi **cine**, *iar* apoi **ce** a făcut.
 want.1 know.SUBJ.1 first who, and then what has done
 ‘I want to know first who did it, and then what he did.’

Furthermore, it is possible to coordinate the yes/no question marker *dacă* ‘if’ and a *wh*-phrase, which also shows that the structure is biclausal, since the answer to the second question has to presuppose that the first is already resolved:

- (42) RATB și Metrorex vor anunța vineri **dacă** și **când** intră în grevă generală.
 RATB and Metrorex will announce Friday if and when enter in strike general
 ‘RATB and Metrorex will announce on Friday if and when they enter in general strike.’

Moreover, it is possible to insert sentence-level adverbials (e.g. speech act adverbs) between the *wh*-phrases:

- (43) Nu văd **cum** și, *mai important*, **cine** ar putea să-l dea jos pe Băseșcu.
 not see.1 how and, most importantly, who could overthrow.SUBJ Băseșcu.ACC
 I don't see how, and most importantly, who could overthrow Băseșcu.

Based on the above arguments, we assume that both coord-*wh* and end-attach-*wh* are biclausal in Romanian (*contra* Comorovski 1996).

3.3. The syntactic structure of Hungarian coord-*wh* and end-attach-*wh*

Concerning Hungarian, we claim that like in Romanian, end-attach-*wh* (45) is biclausal, but contrary to Romanian, coord-*wh* (44) is monoclausal.

- (44) [[**Ki** és **mikor**]_{wh-P} jött]_S? (45) [**Ki** jött]_S [és **mikor** –]_S?
 who and when has come who has come and when –

Our first argument to support this claim comes from the distribution of the interrogative particle *vajon*: unlike the possibility of repeating *oare* with each *wh*-word in Romanian coord-*wh* and end-attach-*wh*, its Hungarian equivalent, *vajon*, cannot be repeated in coord-*wh*, as in (46a), but it can in end-attach-*wh*, as in (46b):

- (46) a. *Vajon ki és (*vajon) mikor érkezett?*
 PRT who and (PRT) when arrived
 b. *Vajon ki érkezett és (vajon) mikor?*
 PRT who arrived and PRT when

Secondly, sentence-level adverbials cannot appear between the *wh*-phrases in coord-*wh*, as in (47a), only in a prosodically marked sentence, as in (47b), where the adverbial and the second *wh*-word are incidental constituents. However, they can appear in end-attach-*wh* before the last *wh*-word, as in (47c):

- (47) a. ***Ki** és még fontosabb **mikor** jött be ide?
 who and even more important when came in here
 b. **Ki** – és még fontosabb: **MIKOR** – jött be ide?
 who – and even more important: when – came in here
 c. **Ki** jött be ide, és még fontosabb: **mikor**?
 who came in here and more important when

Thirdly, some auxiliaries can appear between the *wh*-phrases in coordinate structures, which means that the structure is monoclausal (if we assume that the auxiliary and the main verb have to be clausemates):⁷

- (48) **Mit akarunk** és **hol** vacsorázni?
 what want.3 and where eat for dinner
 ‘What do we want to eat for dinner and where?’

The coordination of the clitic *-e* (the interrogative marker used in subordinate clauses) and a *wh*-phrase can only be tested in end-attach constructions. Since it is possible to coordinate them,

⁷It can be argued that *akar* ‘want’ is an auxiliary in Hungarian. Syntactic evidence comes from the fact that it can interrupt the infinitive following it and appear between the verbal particle (if there is one) and the verbal stem.

this supports our claim that end-attach-*wh* is biclausal:

- (49) Léci, léci, jelezzen, aki még nem tette, hogy jön-e és **hányan!!!**
 please please sign.IMP.3, who yet not did, that comes-CL.INTERR and how many
 ‘Please please, tell me if you come and if so, how many of you!’

Finally, conclusive evidence comes from the definite/indefinite conjugation paradigms. Transitive verbs in Hungarian appear in two conjugations: definite and indefinite. They must agree with the definiteness of their object and appear in the definite conjugation if the object is definite, as in (50), whereas they appear in their indefinite form in all other cases, as in (51).

- (50) Olvasok egy könyvet. (51) Olvasom a könyvet.
 read.1.INDEF a book.ACC read.1.DEF the book.ACC

The interrogative word *mit* ‘what’ triggers the indefinite conjugation, since the object is asked about and not yet identified:

- (52) Mit olvasol?
 what read.2.INDEF

Lipták (2001) observed that if coordinate *wh*-structures were in fact biclausal with ellipsis in the first conjunct, one could not explain the definite conjugation on the verb in (53b):

- (53) a. **Mit** készítesz és **hogyan** (készítéd)?
 what prepare.2.INDEF and how (prepare.2.DEF)
 b. **Mit** és **hogyan** {készítesz / *készítéd}?
 what and how {prepare.2.INDEF / prepare.2.DEF}
 ‘What are you preparing and how (are you preparing it)?’

Note that the questions in (53a)–(53b) do not mean the same. In (53a), the part containing the second *wh*-word (and possibly an elliptical verb) is a question separate from the first. The answer to this second question presupposes that we know the answer to the first, which is why the verb form is definite. In (53b), however, the *wh*-phrases are part of one and the same question, hence the indefinite verb form. In coord-*wh*, *mit* cannot be in a separate clause, since it would have to trigger the definite conjugation on the verb in the second clause containing *hogyan*.

Based on the above arguments, we assume (agreeing with Lipták (2001)) that coord-*wh* is monoclausal, whereas end-attach-*wh* is biclausal in Hungarian.

4. A sketch of analysis in HPSG

In the previous section, we have reached the conclusion that while end-attach-*wh* structures are biclausal in both languages, coord-*wh* exhibit different structures: they are monoclausal in Hungarian, whereas they are biclausal in Romanian. In this section, we sketch a formal analysis of coordinate multiple questions within a construction-based version of HPSG that relies on rich inheritance hierarchies of lexical and phrasal types (Ginzburg and Sag 2000).

4.1. General architecture

In HPSG, words and phrases are modeled as feature structures of the type *sign*, where phonology, syntax, and semantics are represented in one description. Structure sharing allows

for certain values to be identical in a feature structure. Words, unlike phrases, have an argument structure feature (ARG-ST) encoding the subcategorization properties of lexical items. The *synsem* objects which occur on the ARG-ST list may be canonical if they correspond to overt linguistic expressions (and in this case they occur in both VALENCE and ARG-ST lists) or non-canonical (and in this case they only occur in the ARG-ST of a word). Non-canonical synsems have as subtypes (i) extracted elements, typed as *gap* (available for extraction dependencies), (ii) ‘empty’ pronouns, typed as *pro*, and (iii) pronominal clitics, typed as *pron-affix*. Phrases are classified along two dimensions: HEADEDNESS and CLAUSALITY. This cross-classification recognizes a distinction between *headed-phrases* versus *non-headed-phrases*, on the one hand, and *clauses* versus *non-clauses*, on the other. In a headed structure, the HEAD features of the mother are identical to the HEAD features of the head daughter.

According to Ginzburg and Sag 2000, the content of a clause is always some subtype of the semantic type *message: proposition* for declarative clauses, *question* for interrogative clauses, etc. The content of a verb specifies a state-of-affairs (SOA), which contributes to the construction of a certain kind of message. Questions are ontologically distinct from propositions, since they lack properties that are characteristic of truth-bearing expressions. Questions are akin to open propositions, and analyzed as *propositional abstracts*, where the set of abstracted elements may be the empty set (0-ary abstracts for polar questions) or a non-empty set (1-ary abstract, 2-ary or *n*-ary abstracts) for *wh*-questions. The *question* type is specified for two features: a feature PARAMS, whose value is a (possibly empty) set of *parameters* corresponding to the set of entities that get abstracted away, and a PROP feature, whose value is a *proposition*. A *wh*-phrase is thus represented as a parameter that introduces a restriction, and multiple *wh*-phrases are accommodated in terms of a non-singleton PARAMS value. The parameter comprises an index and its restriction (the use of the index allows for linking the abstracted parameter to an argument position within the proposition). Following Bonami and Godard 2006, we consider that the *interrogative-clause* has a unary daughter, which is a clause denoting a proposition (type shifting from *proposition* to *question*). Parameters are retrieved from the daughter’s STORE value, as shown by the constraint in (54).

$$(54) \quad \left[\begin{array}{l} \textit{inter-cl} \\ \text{CONT} \left[\begin{array}{l} \textit{question} \\ \text{PARAMS} \{ \boxed{1}, \dots, \boxed{n} \} \\ \text{PROP} \boxed{0} \end{array} \right] \\ \text{STORE} \{ \} \end{array} \right] \rightarrow \left[\begin{array}{l} \text{CONT} \boxed{0} \\ \text{STORE} \{ \boxed{1}, \dots, \boxed{n} \} \end{array} \right]$$

We analyze coordinate phrases as a subtype of *non-headed-phrase*, consisting of (at least) two immediate constituents, which may each be introduced by a conjunction (Abeillé 2005). Inside the coordinate phrase, the conjunct introduced by a conjunction is represented here (for the sake of simplicity) as a *head-marker phrase*. We follow Sag 2005 in considering that lexical entries do not fix the type of their HEAD values, which can be underspecified. This allows us to handle cases in which coordinate elements do not share the same syntactic category. A very simple way to capture Wasow’s Generalization given in the beginning of this paper is to assume that an element in construction with a coordinate structure has access not only to the coordinate structure as a whole, but also to the syntactic property of each conjunct. In order to implement this, we use the feature CONJUNCTS-LIST (CNJ-LST) introduced by Chaves and Paperno 2007 for ‘hybrid’ coordinations in Russian, which allows the coordinate construction to collect the conjoined signs, making them accessible to the head. A simplified version is given in (55).

$$(55) \left[\begin{array}{l} \text{coord-ph} \\ \text{SYNSEM|HEAD|CNJ-LST} \langle \boxed{1}, \dots, \boxed{n} \rangle \\ \text{CONJ nil} \end{array} \right] \rightarrow \left[\begin{array}{l} \text{SYNSEM} \quad \boxed{1} \\ \text{CONJ} \quad \text{nil} \end{array} \right], \dots, \left[\begin{array}{l} \text{head-marker-ph} \\ \text{SYNSEM|HEAD|CNJ-LST} \langle \boxed{n} \rangle \\ \text{CONJ conj} \end{array} \right]$$

HPSG organizes the sign into information which is locally relevant (LOCAL) and information that plays a role in long distance dependencies (NON-LOCAL). In the case of *wh*-questions, the syntactic relation between the fronted *wh*-phrase and the rest of the clause can be accounted for in terms of an extraction phenomenon, which is a non-local dependency, of the type *head-filler-phrase*, as defined by the constraint in (56).

$$(56) \left[\begin{array}{l} \text{head-filler-ph} \\ \text{SLASH} \boxed{S} \\ \text{STORE} \boxed{0} \cup \{ \boxed{1} \} \end{array} \right] \rightarrow \left[\begin{array}{l} \text{LOC} \boxed{2} \\ \text{WH} \{ \boxed{1} \} \end{array} \right], \mathbf{H} \left[\begin{array}{l} \text{SLASH} \boxed{S} \cup \{ \boxed{2} \} \\ \text{STORE} \boxed{0} \end{array} \right]$$

A *head-filler-phrase* requires exactly two immediate constituents: a filler daughter and a sentential head daughter containing a *gap* (i.e. one of the arguments of the verbal head is not locally realized). The presence of a gap is encoded in terms of a non-empty SLASH value, which is amalgamated by the verb, being also the verb's SLASH value. Further, the verb's SLASH value is propagated through the syntactic tree, until a compatible filler constituent occurs. Extraction is thus treated entirely in terms of the inheritance of SLASH specifications. The percolation of the non-local information has as a result the fact that this information is simultaneously present at every node in the extraction path. Interrogative *wh*-phrases which function as fillers bear non-empty specifications for the feature WH whose value is a set containing a parameter, which is retrieved in the STORE value of the mother, contributing thus to the global content.

4.2. Analysis of Hungarian coordinate *wh*-structures

We have concluded above that coordinate *wh*-structures are monoclausal in Hungarian. Lip-ták (2001) comes to a similar conclusion. She claims that the conjoined *wh*-phrases do share a function, which is *focus*. According to her analysis, focus as the common function would account for 'hybrid' coordinations, such as (3), since the conjoined items are pragmatically prominent, or salient in the discourse,⁸ and they bear a pitch accent in prosody as well. Although it is a usual assumption to claim that question words are best analyzed as (a subtype of) foci, this analysis faces a serious problem here: unlike *wh*-phrases, two non-interrogative foci cannot be coordinated in Hungarian:

- (57) a. **Ki** és **mikor** ment el?
 who and when left PRT
 b. *JÁNOS és TEGNAP ment el.
 John and yesterday left PRT

This means that *focus* cannot be the shared function of the conjoined *wh*-phrases. In our approach, the common function of conjoined *wh*-phrases is *filler* (cf. the HPSG ontology), referring to the fact that these constituents do not appear in their canonical position (i.e. they correspond to *gaps* on the ARG-ST list of the verbal head). The only stipulation we need for Hungarian coord-*wh* is to allow the *head-filler-ph* in (58) to have more than one filler, by us-

⁸Chaves and Paperno (2007) observe the same about Russian, although they cannot determine the exact nature of this pragmatic salience.

ing the CNJ-LST proposed for coordination structures. Coordination of *wh*-fillers is reserved to constituents bearing a WH feature, which prevent non-interrogative fillers from being conjoined.

$$(58) \quad \left[\begin{array}{l} \text{head-coord-filler-ph} \\ \text{SLASH } [5] \\ \text{STORE } [0] \cup \{[3], [4]\} \end{array} \right] \rightarrow \left[\text{CNJ-LST} \left\langle \left[\text{LOC } [1] \right], \dots, \left[\text{LOC } [2] \right] \right\rangle, \mathbf{H} \left[\begin{array}{l} \text{SLASH } [5] \cup \{[1], [2]\} \\ \text{STORE } [0] \cup \{[3], [4]\} \end{array} \right] \right]$$

In figure 1, we provide the syntactic tree for the Hungarian coordinate *wh*-structure *Ki és mit látott?* ('Who saw something and what was it?').

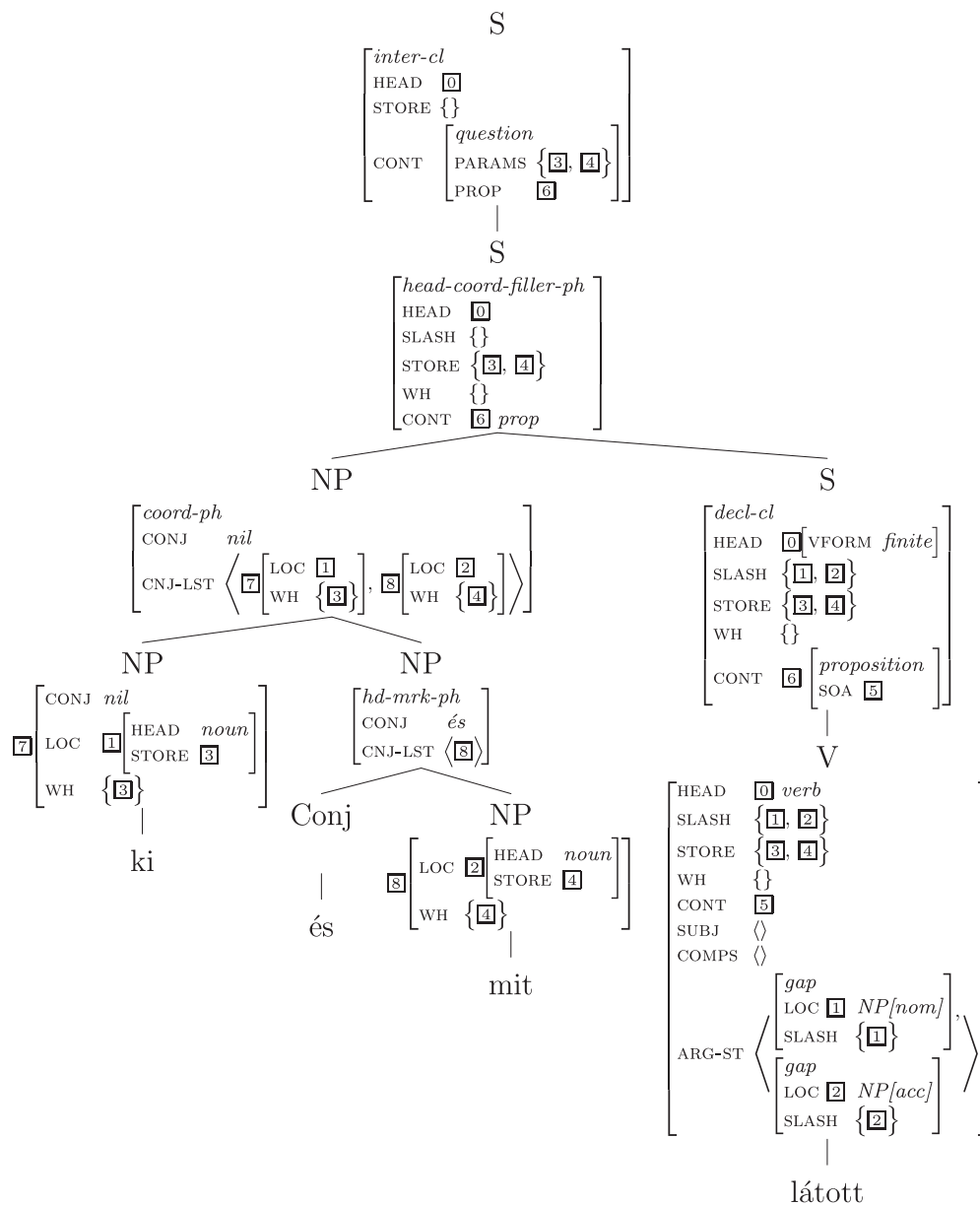


Figure 1: Syntactic tree for Hungarian coordinate *wh*-questions

The ARG-ST of the verbal head contains two *gaps* corresponding to the extracted *wh*-phrases (note the feature-sharing of LOC) and bearing a non-empty value for the feature SLASH. SLASH specifications are amalgamated by the verb and inherited from the verb to the S that it projects. All extracted *wh*-phrases bear a *parameter* value for the feature WH, which is stored by *coordinate-phrase*. In order to be licensed by the grammar, the Hungarian example has then to successively satisfy constraints (55), (58) and (54).

4.3. Analysis of Romanian coordinate *wh*-structures

In the §3.2, we gave some empirical evidence in favour of the biclausal analysis of coordinate multiple *wh*-questions in Romanian. That is, for a coordination such as *Cine și ce a făcut?* ('Who saw something and what was it?'), what we have is a coordination of two interrogative clauses: one which is reduced to the *wh*-phrase and the other one which is a complete clause.

One obvious analysis would be to assume that the first clause undergoes (backward) ellipsis, according to which the elided constituent is structurally represented and interpreted. Thus, ellipsis in the first conjunct is possible under identity with the second conjunct. Such an analysis has to postulate the presence of an empty pronominal in the first conjunct (in order to satisfy the subcategorization requirements of the verbal head). However, cataphoric use of pronominals is usually impossible in coordinate constructions. Moreover, even if we put this problem aside, we observe that not all verbal predicates allow for indefinite null arguments. Therefore, the verb supposedly undergoing ellipsis cannot always be reconstructed in the first conjunct, for instance, because one of its arguments would be missing:

- (59) Polițistul satului îi cunoaște pe toți; știe **cine**
 policeman.DEF village.GEN CL.ACC knows PRT.ACC everyone; knows who
 (*locuiește) și **unde** locuiește.
 (lives) and where (lives)
- (60) **Cine** (*ocupă) și **ce** **loc** ocupă pentru tine?
 who (occupies) and which place occupies for you

The second option is to assume a multidominance analysis (Rațiu 2011). According to this approach, one expects that the 'shared' material could not be realized with each *wh*-phrase and, in particular, could not occur in the first conjunct. However, the 'shared' material can be realized more than once and not necessarily in the second conjunct, as shown in (61b) (*contra* Rațiu 2011).

- (61) a. **Cine** și **ce** a mâncat?
 who and what has eaten
- b. **Cine** a mâncat și **ce** (a mâncat)?
 who has eaten and what (has eaten)

We can avoid these problems if we assume a fragment-based analysis, adapted from Ginzburg and Sag 2000. As shown by the constraint in (62), we use a *wh*-fragment-phrase, which allows for an NP, an PP or an AdvP to be exhaustively dominated by a finite category, which has propositional content. The parameter expressed by the *wh*-expression is inherited via STORE by the mother. Unlike Rațiu 2011 who has to posit two different analyses in order to account for coord-*wh* with co-arguments and co-adjuncts respectively, the analysis we propose can handle

both of these cases, as well as the mixed cases (argument–adjunct or adjunct–argument).

$$(62) \quad \left[\begin{array}{l} wh\text{-fragment-ph} \\ \text{HEAD } [VFORM \textit{finite}] \\ \text{STORE } \{\boxed{1}param\} \\ \text{CONT } \textit{proposition} \end{array} \right] \rightarrow \left[\begin{array}{l} \text{HEAD } \textit{noun} \vee \textit{prep} \vee \textit{adv} \\ \text{STORE } \{\boxed{1}\} \end{array} \right]$$

The only stipulation we must make is to posit a special coordination rule which combines a (full) clause of type *head-filler-ph* and a ‘fragmentary’ phrase sharing the same properties (in particular, the same semantic content) to build the equivalent of a multiple question. The right semantic description is obtained via the content identity imposed by the coordination construction, as shown in (63).

$$(63) \quad \left[\begin{array}{l} \textit{special-coord-ph} \\ \text{STORE } \boxed{1} \cup \boxed{2} \\ \text{CONT } \boxed{3} \end{array} \right] \rightarrow \left[\begin{array}{l} \text{STORE } \boxed{1} \\ \text{CONT } \boxed{3} \end{array} \right], \left[\begin{array}{l} \text{STORE } \boxed{2} \\ \text{CONT } \boxed{3} \end{array} \right]$$

With all these ingredients, we are now able to derive in the figure 2 the syntactic tree for the Romanian example *Cine și ce a făcut?* (‘Who saw something and what was it?’).⁹

The analysis we proposed for Romanian *coord-wh* can be easily extended to the end-attach-*wh* in both languages. For the lack of space, we cannot provide an illustration of this.

5. Conclusion

This paper addressed the problem of coordinate *wh*-phrase structures in Hungarian and Romanian, concentrating on cases where the conjoined *wh*-phrases are in the preverbal domain. We showed that the significant hesitation and variation in the acceptability judgments render the analysis of such structures very difficult. Instead of neglecting it, we attributed the acceptability problem to the interplay of various syntactic and semantic factors. However, experimental studies must be conducted in order to obtain more clear-cut data, which is left for further research.

We rejected those analyses that aim to assign a unique universal structure to *coord-wh* construction cross-linguistically, since we argued that *coord-wh* is monoclausal in Hungarian but biclausal in Romanian, and thus necessitates different analyses. The universalism, if it exists, concerns rather end-attach-*wh*, which receives the same treatment in both languages.

Finally, as far as the Law of Coordination of Likes is concerned, *wh*-coordination is not problematic at all in either language, since the identity constraints imposed by the coordination are always satisfied: in Hungarian, *coord-wh* structures share the *filler* function, while in Romanian, *coord-wh* (as well as end-attach-*wh* in both languages) share the same propositional content.

⁹We represent the ARG-ST of the verbal head as containing two non-canonical arguments: the first one, typed as *pro*, corresponds to an unexpressed subject, while the second one, typed as *gap*, corresponds to the extracted direct object.

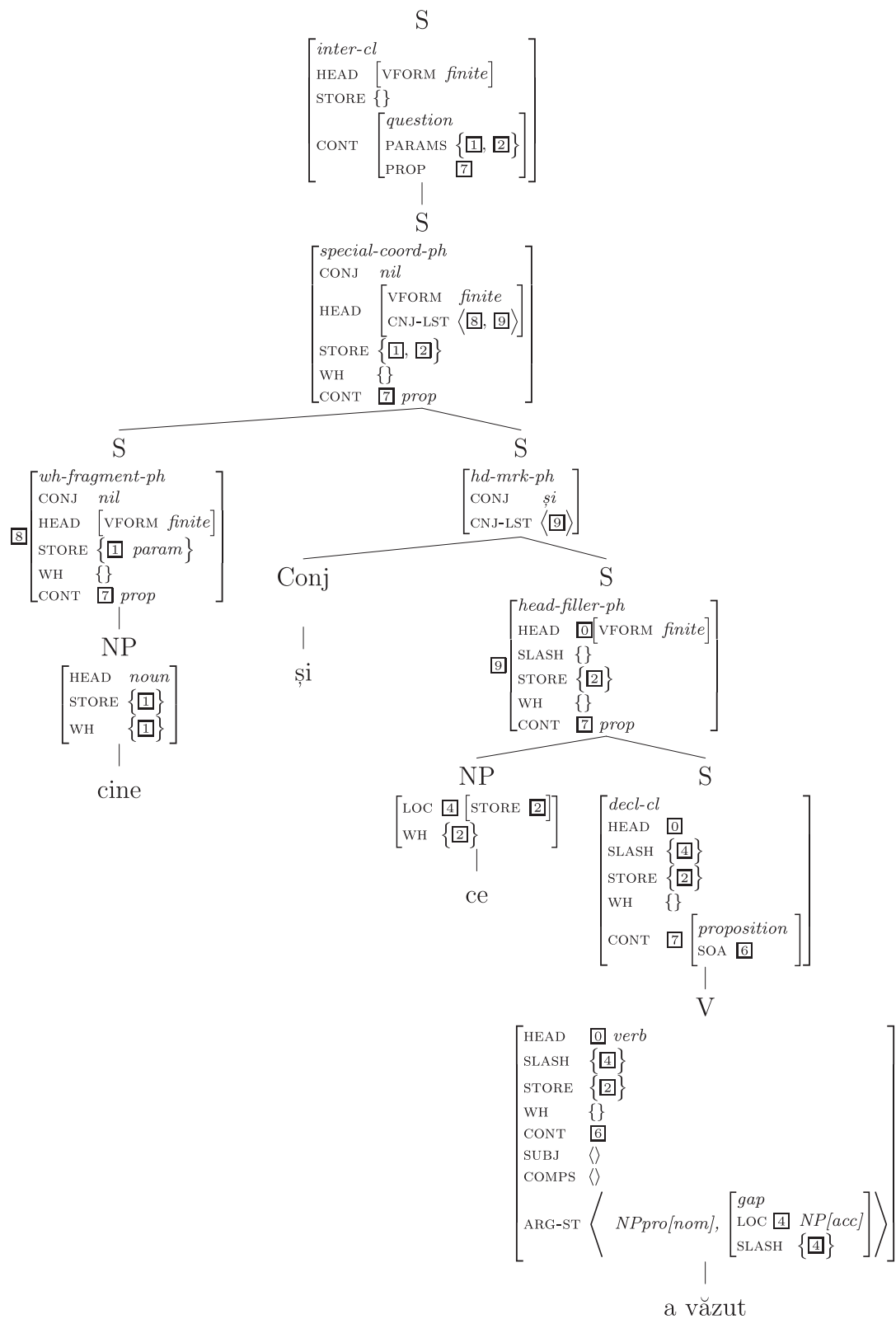


Figure 2: Syntactic tree for Romanian coordinate *wh*-questions

References

- Abeillé, Anne. 2005. Les syntagmes conjoints et leurs fonctions syntaxiques. *Langages* 160.42–66.
- Bánréti, Zoltán. 2007. *A mellérendelés és az ellipsis nyelvtana a magyarban* [The Grammar of Coordination and Ellipsis in Hungarian]. Budapest: Tinta könyvkiadó.
- Bîlbîie, Gabriela. 2011. *Grammaire des constructions elliptiques: Une étude comparative des phrases sans verbe en roumain et en français*. Université Paris Diderot-Paris 7 dissertation.
- Bonami, Olivier, and Danièle Godard. 2006. Sentence initial adverbials and clause types in French. Presentation at Alliance Workshop, Colchester.
- Browne, Wayles. 1972. Conjoined question words and a limitation on English surface structures. *Linguistic Inquiry* 3.223–226.
- Büring, Daniel. 2003. On D-trees, beans, and B-accent. *Linguistics and Philosophy* 26.511–545.
- Chaves, Rui Pedro, and Denis Paperno. 2007. On the Russian Hybrid Coordination Construction. *Proceedings of the 14th International Conference on HPSG*, 46–64. Stanford: CSLI Publications.
- Citko, Barbara, and Martina Gračanin-Yuksek. 2012. Towards a new typology of coordinated Wh-questions. *Journal of Linguistics*. 1–32.
- Comorovski, Ileana. 1996. *Interrogative phrases and the syntax-semantics interface*. Dordrecht: Kluwer Academic Publishers.
- Gazdik, Anna. 2011. *Multiple questions in French and in Hungarian: A Lexical-Functional analysis with special emphasis on the syntax-discourse interface*. Université Paris Diderot-Paris 7; ELTE Budapest dissertation.
- Giannakidou, Anastasia, and Jason Merchant. 1998. Reverse sluicing in English and Greek. *The Linguistic Review* 15.233–256.
- Ginzburg, Jonathan, and Ivan Sag. 2000. *Interrogative investigations: The form, meaning and use of English interrogatives*. Stanford: CSLI Publications.
- Gribanova, Vera. 2009. Structural adjacency and the typology of interrogative interpretations. *Linguistic Inquiry* 40.133–154.
- Haida, Andreas, and Sophie Repp. 2011. Monoclausal question word coordination accross languages. *Proceedings of NELS 39*, ed. by S. Lima, K. Mullin, and B. Smith. To appear.
- Kayne, Richard S., and Jean-Yves Pollock. 2001. New thoughts on stylistic inversion. *Subject Inversion in Romance and the Theory of Universal Grammar*, ed. by A. Hulk and J.-Y. Pollock, 595–621. Oxford: Oxford University Press.
- Kazenin, Konstantin I. 2002. On Coordination of Wh-Phrases in Russian. Ms., Universität Tübingen.
- Kliaschuk, Mikalai. 2008. L'interrogation coordonnée dans les langues slaves. *Proceedings of the 2007 Annual Conference of the Canadian Linguistic Association*.
- Lewis, Shevaun, Bradley Larson, and Dave Kush. 2012. What and when can you fill a gap with something? 25th Annual CUNY Conference on Human Sentence Processing. New York.
- Lipták, Anikó. 2001. *On the syntax of wh-items in Hungarian*. Utrecht University dissertation.
- Lipták, Anikó. 2003. Conjoined questions in Hungarian. *Multiple wh-fronting*, ed. by C. Boeckx and K. Grohmann, 141–160. Amsterdam: John Benjamins Publishing Company.
- Merchant, Jason. 2008. Spurious coordination in Vlach multiple wh-questions. Ms. University of Chicago.
- Rațiu, Dafina. 2011. *De la syntaxe à la sémantique des propositions interrogatives: Étude comparative des questions multiples en roumain*. Université de Nantes dissertation.
- Sag, Ivan. 2005. La coordination et l'identité syntaxique des termes. *Langages* 160.110–127.
- Sag, Ivan, Gerald Gazdar, Thomas Wasow, and Steven Weisler. 1985. Coordination and how to distinguish categories. *Natural Language and Linguistic Theory* 3.117–171.
- Skrabalova, Hana. 2006. Parataxe apparente et coordination des interrogatifs en tchèque. *Faits de Langues* 28.231–242.
- Tomaszewicz, Barbara. 2011. Against Spurious Coordination in multiple wh-questions. *Proceedings of the 28th West Coast Conference on Formal Linguistics*, ed. by M. B. Washburn, K. McKinney-Bock, E. Varis, A. Sawyer, and B. Tomaszewicz, 186–195. Cascadilla Proceedings Project.