Ellipsis in sequences of multiple wh-words
Hana Gruet-Skrabalova

To cite this version:
Hana Gruet-Skrabalova. Ellipsis in sequences of multiple wh-words. Linguística Brunensia, Masarykova univerzita, Brno, 2016, 64 (1), pp.57-69. <hal-01679872>

HAL Id: hal-01679872
https://hal.archives-ouvertes.fr/hal-01679872
Submitted on 10 Jan 2018
Abstract:
This article deals with elliptical sequences of multiple wh-words in Czech. It is argued that such sequences involve two different constructions, elliptical multiple questions and elliptical assertions. Wh-words in elliptical multiple questions share the properties of wh-words in non-elliptical questions; they have the same syntactic distribution and semantic interpretation, they do not show superiority effects, and their combinability depends on their reading. Moreover, elliptical questions can be reconstructed into syntactically complete questions. In contrast, elliptical assertions differ from elliptical questions in several aspects. They always contain the word jak in initial position, they don’t have question interpretation and cannot be reconstructed into full clauses. I propose that elliptical multiple questions result from TP-deletion after movement of the wh-words into CP domain (so-called multiple sluicing) and the TP-ellipsis is driven by the Focus feature. This explains that it can also apply to referential focused DPs. As for elliptical assertions, I claim that they are base-generated and involve semantic ellipsis, i.e. a null category, which gets its interpretation from a TP in the previous context.

Key-words:
Ellipsis; Wh-words; Multiple Question; Sluicing; Czech

1. Introduction
This paper deals with elliptical sequences of multiple wh-words, as in (1) and (2). These sequences are elliptical in that they have sentential interpretation despite the absence of a syntactically full clause. We argue however that they cannot receive a uniform analysis. In particular, only sequences in (1) can be derived from full multiple questions by TP-deletion (Merchant 2001). In contrast, sequences in (2) are base-generated and contain a null category e that receives its clausal content by a semantic reconstruction (Chung et al. 1995).

(1) Přenáší informace, ale nevím od koho komu.
(he) brings information but (I) don’t know from who-gen who-dat
‘He brings information but I don’t know to whom he brings information, and from whom.’

(2) Prý jsou hotely u pláže už plné? Jak kde.
they say hotels on the beach are full how where
‘It seems that the hotels on the beach are full. Some are and some are not.’
The paper is organized as follows. In section 2, we present the properties of ellipsis in questions with a single wh-word, known as sluicing. In section 3, we compare the properties of wh-words in multiple questions and in elliptical sequences. We show that sequences in (1) share most properties with non-elliptical questions, while those in (2) behave differently, in particular with respect to ordering and interpretation. Section 4 deals with the type of ellipsis in these sequences: deletion of a syntactic structure driven by the focus feature in (1), and semantic reconstruction in (2). Section 5 sums up our findings.

2. Sluicing
Sluicing refers to ellipsis of an interrogative clause in which everything except for the interrogative word was elided (Ross 1969). Sluicing appears both in embedded contexts and in independent clauses related to a previous utterance, as shown in (3). The antecedent of the elided clause can be found in independent utterances as in (3b), and in coordinate and subordinate clauses, see (3a) and (4a) respectively. It usually precedes the wh-remnant, but it can follow it in an appropriate context, see (4b).

(3) a. John bought something, but I don't know what. (= what John bought)
   b. John bought something. What? (= what did John buy?)

(4) a. Pokud chceš někoho pozvat, dobre si rozmysli koho.
   if (you) want someone invite, bethink well who
   b. (Ukázal jsem jí mapu): Nevím ještě kam, ale někam si určitě vyjedeme.
   (I) showed her the map: I don’t know where but we will go out somewhere

The sentence containing the antecedent of sluicing also contains an explicit or an implicit correlate of the wh-word, corresponding to an indefinite DP or AdvP, like někdo, něco, někde, etc.:

(5) Domluvil si (s ně kým) práci v Tel Avivu, ale nevzpomínám si s kým. (ČNK)
   (he) arranged (with someone) a job in Tel Aviv, but (I) don’t know with whom

In languages with morphological case-marking, the wh-remnant must bear the same case as its antecedent or as the wh-word in non-elliptical questions1, see (6). The wh-remnant must also be accompanied by the same preposition as its correlate, in contrast to langages like English that allow preposition-stranding, see (7).

(6) a. Jan koupil někomu dárek, ale nevim komu / *koho (komu / *koho koupil dárek)
   John bought someone a gift but (I) don’t know who / who
   b. Jan chce někoho pozvat, ale nevi ještě koho / *komu. (koho / *komu chce pozvat)
   John wants invite someone but (he) doesn’t know who / who

(7) a. Musel dostat od někoho informaci. A já se ptám, *(od) koho.
   (he) must have received information from someone and I ask (from) whom
   b. He was talking to somebody, and I want to know (to) whom.

---

1 Contrary to cleft-sentences where the wh-word is in nominative case (cf. Grebenyova 2006).
The case-marking and the possibility (or not) of preposition stranding constitute two principal pieces of evidence for deriving sluicing by TP-deletion from non elliptical questions (Ross 1969, Lasnik 2001, Merchant 2001), as illustrated in (8):

(8)  

a. John bought something but I don't know \([\text{CP } \text{what } [\text{TP } \text{John bought } \underline{\text{t}}]]\).

b. Jan něco koupil, ale nevim \([\text{CP } \text{co } [\text{TP } \text{Jan koupil } \underline{\text{t}}]]\).

Assuming the analysis in (8), we make hypothesis that the elliptical sequences we are dealing with are derived from questions with more than one wh-word, so-called multiple questions (cf. Stjepanović 2003 for Serbo-Croatian, Grebenyova 2006 for Russian). This hypothesis predicts (i) that the wh-remnants have the same properties as the wh-words in non-elliptical multiple questions, and (ii) that they can be syntactically reconstructed into full questions. In the following section, we show that these predictions are only born out for wh-sequences like in (1).

3. Properties of wh-words in multiple questions and in elliptical sequences

There exists substantial literature about the properties of multiple wh-words in Slavic languages. In this section, we compare wh-words in multiple questions and in elliptical sequences, focusing on their adjacency, ordering, interpretation and combinability. We show that elliptical wh-sequences fall into two groups: sequences that behave like multiple questions, and sequences with an initial jak ('how') that behave differently, in particular with respect to their ordering and interpretation.

3.1 Adjacency

All wh-words in informative multiple questions in Czech must be fronted. They can however be separated by second position clitics (auxiliary verbs and pronouns), as in (9). The clitics’ position is related to interpretation, see 3.3.

(9) 

a. Co komu jsi koupil? / Co jsi komu koupil?  
   what who-DAT cl.2SG bought / what cl.2SG who-DAT bought  
   ‘What did you buy, and for whom?’

b. Rád bych věděl, kdo se kde schovává / ?kdo kde se schovává.  
   I wonder who cl.REFL where hides / who where cl.REFL hides  
   ‘I wonder who is hiding, and where.’

In contrast, wh-remnants in elliptical sequences must always be adjacent, see (10). Assuming that clitics constitute a barrier between TP and CP domains, the ellipsis site in wh-sequences must involve TP and clitics.

(10) 

a. Prý jsi každému něco koupil. Řekni mi co komu / *co jsi komu.  
   say (you) cl.2SG bought something to everyone tell me what to-whom/cl.2SG to-  
   whom  
   ‘It seems that you bought someone something. Tell me what you bought, and to  
   whom.’

---

3 With some differences that will be explained in section 4.
3.2 Ordering and correlates

We observe no superiority effects in multiple questions in Czech, whether they are embedded or not (Meyer 2004):

(11) a. Co komu koupil? / Komu co koupil?
    what who-DAT (he) bought / what who-DAT (he) bought
b. Řekni mi kdo koho / koho kdo pozval k tanci.
    tell me who who-ACC / who-ACC who invited to dance

Although both orders are possible in elliptical sequences, parallel ordering of wh-remnants with respect to their correlates is often preferred, see (12b). The correlates of multiple wh-remnants correspond to quantified and indefinite XPs as in (12), or to plural referential DPs, as in (13), and the wh-remnants always bear the case of their correlates.

(12) a. Každému něco koupil, ale komu co / co komu, to nevim.
    to-everyone something (he) bought but what who-DAT / what who-DAT I don’t know
b. Každý pozval někoho k tanci, ale nevzpomínám si kdo koho / ??koho kdo.
    everyone invited someone to dance but I don’t remember who who-ACC / who-ACC who
c. Spojka [...] přenáší informace, ale nevim od koho komu. (ČNK)
    (the liaison officer) transmits information but (I) don’t know from who-GEN who-DAT

(13) a. Domluvili se, že je budou střílet do srdce. A rozdělili si, kdo koho. (ČNK)
    (they) arranged that they will shoot them in heart and they decided who who-ACC
b. Moji kamarádi se na tu oslavu (nějak) přestrojili, ale nevzpomínám si už, jak kdo.
    my friends for the party (somehow) dressed up, but I don’t remember how who

On the contrary, sequences introduced by jak (‘how’) as in (14) behave differently because (i) the order of the wh-remnants is strict, and (ii) the wh-remnants do not have indefinite nor quantified correlates. So, contrary to (13b) where jak has an indefinite correlate nějak, jak in (14) is linked to specific adverbials (of different semantic types) in the antecedent clause, i.e. full, nicely, main, very. The wh-word following jak always has a plural referential correlate, generally implicit, which bears the same case, see (14b,c).

(14) a. Prý jsou hotely u pláže už plné? Jak kde. / *Kde jak. (ČNK)
    they say hotels on the beach are full how where / where how
b. Je pravda, že se tady (lidem) přijemně tráví čas? Jak komu. / *Komu jak.
    is (it) true that (people-DAT) passes nicely time here how who-DAT / who-DAT how
c. Bylo vaší hlavní inspirací Chile? Záleží jak v čem / *v čem jak.
    was you main inspiration Chile it depends how in what / in what how

---

4 In our corpus (ČNK, subcorpus SYN2010), jak appears much more frequently in the initial position.
5 According to Grebenyova (2006), the parallelism is obligatory in Russian.
3.3 Distribution

Multiple questions can be both independent and embedded, see (11) above. In comparison, elliptical sequences mostly appear embedded under predicates that subcategorize an interrogative clause, except for jak-sequences in (14). These latter mostly constitute independent utterances, or appear embedded under the predicate záleží (‘it depends on’), which is due to their interpretation (see 3.4).

Furthermore, elliptical sequences usually follow their antecedent clause. However, we can find sequences nevím jak kdo (‘I don’t how who’) preceding their antecedent, as in (15a), in which the wh-word kdo contrasts with an explicit referential correlate and can be itself replaced by another referential DP. When the antecedent precede the wh-remnants, kdo refers to each individual included in the correlate, thus I don’t know which one of us was satisfied in (15b). This order is therefore impossible when the correlate denotes a single individual, compare (16a) and (16b). In section 4, we show that these sequences can receive the same syntactic analysis as other sequences embedded under ‘to know’.

\[(15)\] a. Nevím, jak kdo / jak vaši kolegové, ale my jsme byli velmi spokojeni. \((ČNK)\)
\[\text{(I) don’t know who / how your colleagues but we were very pleased}\]
\[\text{b. My jsme byli velmi spokojeni, ale nevím, jak kdo.}\]

\[(16)\] a. Nevím, jak koho / jak Pavla, ale mě docela rozčílí, když... \((ČNK)\)
\[\text{(I) don’t know how who-ACC / how Paul-ACC but (it) quite irritates me when...}\]
\[\text{b. *Mě docela rozčílí, ale nevim jak koho, když....}\]

3.4 Interpretation

Multiple questions can have pair-list (PL) reading and specific-pair (SP) reading. In Czech, non-adjacent wh-words require pair-list (or distributive) reading and conjoined wh-words require specific reading, see (15). In absence of clitics, the question is compatible with both readings (Gruet-Skrabalova 2011). Adjacent and conjoined wh-items can also have contrastive reading if the participants are known and the question only asks to identify their respective roles.

\[(17)\] a. Kdo se komu omluvil? Jan se omluvil Petrovi, Pavel Tomášovi a Martin Ondroví. \((PL)\)
\[\text{who CL.REFL to-whom apologized J. CL.REFL apologised to P., P. to T. and M. to O.}\]
\[\text{b. Kdo komu se omluvil? Jan se omluvil Petrovi (a ne Petr Janovi). \((SP)\)}
\[\text{who to-whom CL.REFL apologized J. CL.REFL apologised to P (and not P. to J.)}\]
\[\text{c. Kdo a komu se omluvil? Jan se omluvil Petrovi (a ne Petr Janovi). \((SP)\)}
\[\text{who to-whom CL.REFL apologized J. CL.REFL apologised to P (and not P. to J.)}\]

Elliptical sequences as in (12) and (13) can be reconstructed into full multiple questions, as shown in (18), and are interpreted like that. Their specific interpretation depends on the correlates, see (19): quantified correlates require distributive reading and therefore adjacent wh-remnants, while indefinite correlates require specific reading and occur more frequently with conjoined6 wh-remnants.

---

6 Cf. Tak se mi zdá, že se tady někdo někому líbí. No to by mě teda vážně zajímalo, kdo a komu. \((ČNK)\)
\[\text{it seems to me that here someone to-someone pleases I really wonder who and who-DAT}\]
In contrast, elliptical *jak*-sequences as in (14) are interpreted as assertions. They imply that a previous assertion, given or presupposed in the context, is true for some *x* (individuals, objects, moments, etc.), but not for all, as shown by the paraphrase and the context in (20). Thus, they always have distributive reading. The type of *x* is given by the wh-word following *jak*: *x* is an individual from the set *kdo* (*‘who’*) in (20a), and a moment from the set *kdy* (*‘when’*) in (20b). *Jak* behaves here therefore not as an interrogative word, but as an adverbial quantifier that distributes the affirmative or the negative value of the preceding assertion over the set of *x* given by the subsequent wh-word. The syntactic reconstruction is impossible, as shown in (21), because the deleted TP should be both affirmative and negative.

(20)   a. Je pravda, že se tady příjemně tráví čas? *Jak komu,* = Někomu ano, někomu ne. is (it) true that (one) has a nice time here how who-DAT = to-someone yes, to-someone no
b. Za co utrácíš kapesné? *Jak kdy,* někdy to utratím, někdy šetřím. for what (you) spend money how when sometimes (I) spend it, sometimes (I) spare it


### 3.5 Combinability
Although all wh-words can appear in multiple questions, two adjacent adjuncts are usually considered as less felicitous in comparison to conjoined adjuncts. This is however due to the possibility or not to obtain a distributive reading (see above): in (22), distributive reading of *kdy kde* (*‘when where’*) is possible only in the example (b): ‘for each period of holidays, where will you be?’. In (23), distributive reading of *jak kdy* (*‘how when’*) also better obtains in (b): ‘he doesn’t know for each party how his girlfriends were dressed’.

(22)   a. *Kdy kde / Kdy a kde* se máte setkat?
when where / when and where should you meet
b. Řekni mi *kdy kde / kdy a kde* budete během prázdnin.
tell me when where / when and where (you) will be during holidays

(23)   a. *??Jak kdy / Jak a kdy* hodnotíte studenty na univerzitě?
how when / how and when (you) evaluate students at the university
b. Neví vůbec, *jak kdy* byly oblečeny jeho přítelkyně. (ČNK)

(he) doesn’t know how when were dressed his girlfriends
In our corpus study of 16 combinations\(^7\), we have found almost 1800 occurrences of multiple questions, the five most frequent combinations being co kdo (253), co kdy (223), jak co (188), kdo kdo (163), and kdo co (137). We haven’t found any occurrence of sequences kdy jak, kde kam / kam kde, kde jak, and kam jak / jak kam. In contrast, we have found less than 100 elliptical sequences\(^8\), the most frequent being independent jak-sequences (88): jak kdo (67), jak co and jak kdy / jak kde. We have only found 5\(^9\) elliptical sequences with other wh-words than jak. In one of them, the wh-words have rather indefinite interpretation (see section 4.1):

(24) Líp ti to neřeknu a nevím, zda kdy kdo.
(1) cannot say it better to you and I don’t know of when who

### 3.5 Summary

The properties discussed in this section show that elliptical wh-sequences fall into two groups: those that have the properties of multiple questions (that we call \(E\)-sequences) and those that do not (jak-sequences):

<table>
<thead>
<tr>
<th>Multiple questions</th>
<th>Adjacency not necessarily</th>
<th>Interpretation question distributive / specific + contrastive</th>
<th>Free order yes</th>
<th>Contexts main or embedded</th>
<th>Correlates indefinites, quantifiers, referential XPs</th>
<th>Combinations all but less acceptable</th>
<th>Reconstruction yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>(E)-sequences ex. (1)</td>
<td>yes</td>
<td>question distributive / specific + contrastive</td>
<td>yes</td>
<td>mostly embedded under ‘to know’</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>jak-sequences ex. (2)</td>
<td>yes</td>
<td>assertion distributive</td>
<td>no</td>
<td>mostly main or embedded under ‘to depend’</td>
<td>specific adverbal + referential XPs</td>
<td>jak + wh-word</td>
<td>no</td>
</tr>
</tbody>
</table>

Table 1: Properties of multiple questions and elliptical sequences

### 4. Two types of ellipsis

In this section, we propose that elliptical sequences involve two different types of ellipsis: \(E\)-sequences involve syntactic deletion in multiple questions, while jak-sequence contain a base-generated null category \(e\).

#### 4.1 Deletion in multiple questions

Wh-remnants in \(E\)-sequences behave as wh-words in multiple questions. Consequently, \(E\)-sequences can receive the syntactic analysis of sluicing constructions discussed in the section 2: they contain a clausal structure, the wh-words move to the left-periphery, and TP is deleted under identity with the TP in the antecedent clause (Merchant 2001)\(^{10}\):

---

\(^7\) Our study is limited to the following wh-words in both orders and all cases: kdo (who), co (what), kdy (when), kde (where-LOC), kam (where-DIR), jak (how).

\(^8\) Sequences of conjoined adjuncts are also very frequent.

\(^9\) We don’t consider the lexicalized sequence kdo s koho (‘who of who GEN’) meaning ‘who of them two will prevail over’. Den konečného zúčtování se přiblížil. Nyní se ukáže, kdo s koho, (ČNK) day of final judgement came close now it will reveal who will have the upper hand.

\(^{10}\) The identity of ellipsis does not concern agreement features (Merchant 2001, 2006).
(25) Každý někoho pozval, ale nevím [CP kdo₁ koho₂ [TP t₂ pozval t₂]]

In Merchant (2001)’s analysis, TP-deletion is licensed by the head C which bears the features [+Q] and [+Wh]. It has been however argued (Bošković 1998, 2002, Stepanov 1998, Stepanović 2003) that movement of multiple wh-words in Slavic is driven by the feature Focus, and that multiple sluicing is licensed by every head bearing the feature Focus (Grebenyova 2006). We propose that a focus-driven analysis of sluicing is appropriate for multiple sluicing in Czech, because it allows to explain adjacency and specific or at least contrastive interpretation of wh-remnants (see section 3). For that, we assume that only adjacent wh-words in multiple questions appear both in the CP domain, i.e. before clitics, and occupy the specifiers of IntP and FocP respectively (Rizzi 1997, Lenertová 2001, Gruet-Skrabalova 2011), see (26). Consequently, sluicing licensed by the head Focus will result in sequences of adjacent wh-remnants with specific or contrastive interpretation. The elided structure corresponds to TP and clitics, thus FinP. Semantically, the contrast obtains most easily with the wh-words of the same type, which explains the preference for elliptical sequences kdo kdo₁.¹¹

(26) a. [ForceP [IntP Kdo [FinP se [TP komu [vP t omluvil t]]]]]
   => *..., ale nevím [ForceP [IntP kdo [FinP se [TP komu [vP t omluvil t]]]]]
   b. [ForceP [IntP Kdo [FocP komu [roc] [FinP se [TP [vP t omluvil t]]]]]
   => ale nevím [ForceP [IntP Kdo [FocP komu [roc] [se [vP t omluvil t]]]]]

Finally, focus-driven analysis of sluicing predicts that sluicing should be also possible after focused referential DPs (Grebenyova 2006). This is actually true for both a single DP and a sequence of two DPs, as shown in (27). Consequently, our analysis can also account for sequences combining wh-words, referential DPs and indefinite pronouns, as long as they are contrastively focused, see (28).

(27) a. Mohl to někdo udělat sekerou? Spiš bych řekl, že mačetou. (ČNK)
   could someone do it axe-INSTR I would say that truncheon-INSTR
   b. Že Češi spolupracuji s Francouzi je jisté, ale jestli Francouzi s Čechy už tak jisté
      není.
      that Czechs cooperate with French is sure but that French with Czechs is not so sure

(28) a. Každý s někým tancoval, ale kdo s Evou (tancoval), na to si už nevzpomínám.
   everyone with someone danced but who with Eva (danced) I don’t remember
   b. Nevím, jak koho / jak Pavla (to rozčíluje), ale mě to teda rozčíluje.
      (I) don’t know how who-ACC / how Paul-ACC (it irritates) but (it) certainly irritates me
   c. Líp ti to neřeknu a nevím, zda kdy kdo (ti to řekne líp).¹²
      (I) cannot say it better to you and I don’t know whether when who (will say it better)

¹¹ This also explains that three and more wh-remnants are always conjoined or separated by a comma.
¹² We assume that the polarity feature of the clause is checked in the CP (Laka 1990), thus not relevant for TP-identity.
4.2 Semantic reconstruction

Jak-sequences do not allow syntactic reconstruction, although they have clausal interpretation. These non-sentential assertions confirm the value of a previous (presupposed) assertion for some but not all $x$ in a set given by the wh-word following jak. The variable $x$ has an implicit correlate in the presupposed assertion, see (29a). The syntactic reconstruction is impossible because these sequences imply both an affirmative and a negative clause, see (29b) and section 3.4.

(29)  
a. Je pravda, že se tady (lidem) příjemně tráví čas? Jak komu.
    is (it) true that (people) have a nice time here \text{how who-DAT} 

b. Někому se tady tráví čas příjemně, někому ne.
    someone has a nice time here, someone not

We suggest that jak-sequences are base-generated and contain a null category [e] that is semantically reconstructed at LF as other anaphors (Chung et al. 1995, Lobeck 1995). The fact that the interrogative DP following jak must bear the same case as its correlate is necessary to ensure its syntactic and semantic integration into the ellipsis site (Chung 2013:29). More detailed semantic analysis of these constructions is however beyond the scope of this paper.

(30)  
a. Je pravda, že se tady příjemně tráví čas? [CP Jak komu [TP e]].

b. LF: [CP Jak komu [TP lidem, se tady příjemně tráví čas]]

5. Conclusion

In this paper, I dealt with elliptical sequences of multiple wh-words in Czec. I argued that they involve two different constructions, elliptical multiple questions and elliptical assertions. I first showed that wh-words in elliptical multiple questions share the properties of wh-words in non-elliptical questions: they have the same syntactic distribution and semantic interpretation, they do not show superiority effects, and their combinability depends on their reading. Moreover, elliptical questions can be reconstructed into syntactically complete questions. In contrast, elliptical assertions differ from elliptical questions in several aspects. They always contain the word jak in initial position (the order of wh-words is therefore not free), they don’t have question interpretation and cannot be reconstructed into full clauses. I then proposed that elliptical multiple questions result from TP-deletion after movement of the wh-words into CP domain (so-called multiple sluicing) and that the TP-ellipsis is driven by the focus feature. Thus explains that it can also apply to referential focused DPs. As for elliptical assertions, I claimed that they are base-generated and involve semantic ellipsis, i.e. a null category, which is interpreted anaphorically with respect to a TP in the previous context.
References