

# The curious case of the two-headed phrase: A generative approach to co-compounds<sup>1</sup>

Lena Borise<sup>1</sup> & Tamás Halm<sup>2,3</sup>

<sup>1</sup>University of the Basque Country (EHU/UPV); <sup>2</sup>Pázmány Péter Catholic University; <sup>3</sup>Hungarian Research Centre for Linguistics

International Conference on the Structure of Hungarian 16  
University of Graz  
June 28<sup>th</sup>, 2023

## 1. Introduction

- Co-compounds (**CCs**): morphosyntactically parallel and semantically related pairs of elements.
- CCs are well-described in the typological literature and are usually analysed as a type of (asyndetic) **coordination** (Wälchli 2005).
- CCs are amply attested in various languages, including **Hungarian**<sup>2</sup> (1) and **Khanty** (2) (both: Ugric, Uralic).<sup>3</sup>

- (1) a. *János adta-vette a használt autókat.* Hungarian  
John sold-bought the used cars  
'John was trading (lit. selling-buying) used cars.'
- b. *Anti fel-alá-sétált.*  
Tony up-down-walked  
'Tony was walking around (lit. up-down).'
- c. *János megosztotta velem ügyét-baját.*  
John shared me.with affair.3SG.ACC-problem.3SG.ACC  
'John shared all his goings-on (lit. affair-problem) with me.'
- d. *A volt barátok jobbjára a maguk rohanós-sietős életével vannak elfoglalva.*  
the ex friends mostly the own running-hurrying life.with are busy  
'Ex-friends are preoccupied with their own busy lives.'
- (2) a. *je:ji-γən man'i-γən li:k-kən jin't-γən.* Khanty  
older\_brother-DU younger\_brother-DU eat-PST.3DU drink-PST.3DU  
'Older brother and younger brother ate and drank.'
- b. *nuy-kən mi:n-γən*  
2SG-DU 1SG-DU  
'you and me'

---

<sup>1</sup> We would like to thank Márta Csepregi, Katalin É. Kiss, Marcel den Dikken and all the members of our project team and two ICSH reviewers for helpful comments and advice. Our research reported here has been supported by Grant 129921 of the National Research, Development and Innovation Fund of Hungary, the Bolyai scholarship of the Hungarian Academy of Sciences and the Bolyai Plus scholarship of the New National Excellence Programme of the Ministry of Innovation and Technology of Hungary.

<sup>2</sup> A terminological note. Traditionally, what we regard as co-compounds here have been divided into 3 categories in Hungarian descriptive grammars:

(i) *álikerszók* 'fake echo words': CCs that are phonologically similar, such as *ázik-fázik*;  
(ii) *lazga szerkezetű mellérendelések* ' coordinations of a loose structure': CCs that are phonologically dissimilar such as *süt-főz*;  
(iii) *valódi mellérendelések* 'real coordinations' such as *rúgkapál*: CCs that are phonologically dissimilar and where only the last element receives inflectional suffixation.

We argue that these distinctions are irrelevant and analyze all CCs within categories (i-iii) simply as CCs.

<sup>3</sup> The Hungarian data was collected from native speakers, either via elicitations or drawn from corpora or the internet. The Khanty data comes from elicitations with native speakers of Surgut Khanty, unless otherwise noted; examples from textual sources are accompanied by references.

- We take co-compound formation to be a **syntactic operation**, in contrast to much of the descriptive literature on Hungarian, where co-compounds are taken to be a lexical category (*álikerszók* = fake echo words)
- CCs have only received sporadic attention from generative syntacticians, and have mostly been analysed as a **subtype of exocentric compounds** (Scalise, Fábregas & Forza 2009), a classification which we would like to challenge.
- In contrast, according to a tentative analysis proposed in Borise & É. Kiss (2022), in Khanty “the members of a co-compound are **juxtaposed lexical heads** (nouns, adjectives, numerals, or verbs) rather than juxtaposed phrasal projections”.

In this paper, we build on this suggestion and argue that co-compounding is an instance of **two heads undergoing Merge** and being dominated by a **shared layer of functional projections**. We support this analysis with Hungarian and Khanty data.

- While never overtly spelled out (to our knowledge), an implicit assumption of endocentricity (Chomsky 1970) and the projection principle (Chomsky 1993) has been that it is **exactly one head** that heads and projects a phrase (cf. Lichte 2021 for a recent overview).
- It might therefore appear problematic that phrases can in fact be **two-headed** – however, we will argue that as long as a few sensible conditions (that CCs are subject to) are met, the existence of a two-headed phrase is **unproblematic** for the standard understanding of **endocentricity and the projection principle**.

## 2. Data

### 2.1 Similarities between Hungarian and Khanty

- In both Hungarian and Khanty, CCs are made up of juxtaposed lexical elements with **no overt coordinator**:

- (3) a. *ügy-é-t* (\*és) *baj-á-t* Hungarian  
 affair-3SG-ACC and problem-3SG-ACC  
 ‘his goings-on (lit. affair-problem)’
- b. *xɛ:nti xo* (\*pɛ:nə) *xɛ:nti ne:* Khanty  
 Khanty man and Khanty woman  
 ‘(Khanty) people (lit. Khanty men-Khanty women)’

- The two members of a CC are **obligatorily adjacent and inseparable**: when subject to movement (such as focus-induced movement of the verb), they move as a unit:

- (4) a. *János fúrta-faragta szét fúrta-faragta a szekrényt, nem pedig Mari.*  
 John drilled-carved PRT the cupboard not although Mary.  
 ‘It was John who shred the cupboard into pieces, not Mary.’

- b. *Öt osztály saját maga sütötte-főzte meg sütötte-főzte ebédjét bográcsban vagy grillen.*  
 5 classes own self fried-boiled PRT lunch cauldron.in or grill.on  
 ‘Five classes prepared their own lunch in a cauldron or as a barbecue.’

- Their elements are **closely related semantically**: they are synonyms (1d), taxonomic sisters (1c), antonyms (1b), or reverses (1a).
- In both Modern Hungarian and Khanty, **strict morphological parallelism** between the two members of a CC is required. This means, e.g., that bound inflectional morphemes appear on both elements and must be fully matching, as illustrated for Hungarian in (5) and for Khanty in (6):

(5) a. *ügy-é-t* - *baj-á-t* Hungarian  
 affair-3SG-ACC problem-3SG-ACC  
 ‘his goings-on (lit. affair-problem)’

b. *\*ügy-é-t* - *baj-a-í-t*  
 affair-3SG-ACC problem-3SG-PL-ACC  
 (‘his goings-on (lit. affair-problems)’)

(6) *\*i:mp-əm* *ke:fkə-γət-əm*<sup>4</sup> Khanty  
 dog-1SG cat-DU-1SG  
 (‘my dog and my two cats’)

- Moreover, in the presence of possessive marking, **strict parallelism with respect to possessor identity** is required (i.e., morphological parallelism in the absence of indexing parallelism is not allowed). This is shown for Khanty in (7).

(7) *i:mp-ət* *ke:fkə-γət*<sup>5</sup> Khanty  
 dog-3SG cat-3SG  
 ‘his/her<sub>i</sub> dog & his/her<sub>i/\*j</sub> cat’

## 2.2 Differences between Hungarian and Khanty:

- ‘**Accidental**’/spontaneous co-compounds, licensed by context, are possible in Khanty -- but even then, they must have a **degree of semantic relatedness**, (8a). If the two members do not have a semantic relationship, they are not felicitous as a co-compound even in the presence of a unifying context, (8b).

(8) a. *v:tji-γən* *tje:tji-γən* Khanty  
 father-DU grandmother-DU  
 ‘father & grandmother’  
 (context: *Father and grandmother (on father’s side) are relatives.*)

b. *??Kənjikə-γən* *sv:rt-γən* *pəsən* *aftu-nə* *βot-γən*.  
 book-DU pike-DU table top-LOC be-PST.3DU  
 ‘A book and a pike were on the table.’

- To our knowledge, ‘accidental’ co-compounds **do not arise in Hungarian**, even for closely related semantic concepts.

(9) a. *\*A róka-farkas elért* *az erdő szélére*.  
 the fox-wolf PRT.reached the forest edge  
 b. *A róka és a farkas elért* *az erdő szélére*.  
 the fox and the wolf PRT.reached the forest edge  
 ‘The fox and the wolf reached the edge of the forest.’

- Some instances of **inflectional morphology on the second element only** were attested in earlier stages of both Hungarian (10) and Khanty (11), but this pattern is no longer productive in either language. There are some possible fossils of this stage in Modern Hungarian though, (12).

<sup>4</sup> The dual suffix *-γən* is realized as *-γət-* in the context of the following possessive suffix.

<sup>5</sup> Note that /γ/ in the possessive suffix in *ke:fkə-γət* is used solely for vowel hiatus resolution.

(10) a. †*A Muraszombati járásbírótság majd eligazítja az atyafiak ügybaját.* (1900)<sup>6</sup>  
(instead of Modern Hungarian: *ügyét-baját*)

b. †*Lót-futott mindenik amerre futbatott.* (1774)<sup>7</sup>  
(instead of Modern Hungarian: *lótott-futott*)

(11) *kur-en      uč-en-a      [...]      kerjenti-tajen*  
foot-2SG    clothes-2SG-LAT    fall-IMP.3DU  
‘They shall fall to your feet [and] clothes’ (Lewy 1911: 21)

(12) a. *rüg-kapál-t*  
kick-flail-PST.3SG

b. *\*rüg-ott-kapál-t*  
kick-PST.3SG-flail-PST.3SG  
‘She threw her limbs around.’

- For the purposes of this talk, we limit our attention to the synchronically productive pattern of affixation on both members.
- In Khanty, CCs that consist of countable nouns commonly carry a **DU suffix** on each member (there is a certain degree of variability between speakers and contexts with respect to the obligatoriness of dual marking).
  - This does not mean that each member refers to two items – instead, DU-marking indicates the close semantic connection between the two members of the CC co-compound.
  - If used as a subject or topical object, a co-compound elicits **dual agreement on the verb** – see e.g., (2), (8a), and (13).

(13) *Me:    qut-əm-v                    βoqi-yən    v.mp-yən    jv.βət-yən.*  
1SG house-1SG-LAT    fox-DU    dog-DU    come-PST.3DU  
‘A fox and a dog came to my house.’

- Functional load & productivity:
  - In (contemporary) Khanty, **overt coordination of words/phrases is possible** (as a relatively recent development; Borise & É. Kiss 2022).
  - Nonetheless, **co-compounding is still productive in Khanty** and, especially in colloquial speech and story-telling contexts, may be preferred to overt coordination (subject to semantic and morphological restrictions detailed above).
  - In Modern Hungarian, overt **coordination is the dominant strategy**.
  - Co-compounding is a **secondary** strategy -- but still a fully productive one. Appearances may be deceptive here: while many of the most frequently used CCs are quasi-idiomatic (due to the archaic nature of one of the elements: *sír-rí* cry-weep ‘sob’, *lót-fut* race-run ‘hurry around’), it is clear that the co-compounding remains productive:

(14) a. *A mai iskolások között már nem számít kirívónak,*  
the current schoolchildren among already not counts striking.as  
*ha valaki kódol, programoz a gépen.* (Internet)  
if someone codes-programmes the computer.on  
‘These days it is no longer surprising if a schoolchild is doing some coding on her own computer.’

<sup>6</sup> Muraszombat és vidéke 16:39, 1900, September 30th; note also the relevant entry in The Explanatory Dictionary of the Hungarian Language (Bárczi & Országh 1959).

<sup>7</sup> Dugonics András: Trója veszedelme. Landerer, Pozsony, 1774.; note also the relevant entry in The Explanatory Dictionary of the Hungarian Language (Bárczi & Országh 1959).

- b. *Szorgos lakatosok kalapálták-domborították a lemezeket.* (HNC<sup>8</sup>)  
 busy metal workers hammered-pressed the shields  
 ‘The metal workers were busily shaping the shields into form.’
- c. *Órákon át lakodalmas zenére mulatott-táncolt a násznép.* (HNC)  
 hours through wedding tunes.to had.fun-danced the wedding party  
 ‘The guests kept partying for hours to faux-traditional wedding tunes.’

### 3. In favour of symmetric syntax

- We propose that the two elements of a CC are combined via **two syntactic heads undergoing Merge**.
- We support this analysis by demonstrating that:
  - in the presence of a **complement**, the two members of a CC necessarily share it.
  - any **higher functional projections** necessarily apply to/modify both members of a CC.
- We also address **the issue of labelling** that emerges when two heads undergo Merge in a symmetric fashion.
- Finally, we provide arguments **against**:
  - treating CCs as **exocentric compounds** and
  - treating CCs as instances of **asyndetic coordination**.

#### 3.1 Shared complements

- In Hungarian, verbal particles (and verb modifiers in general) are standardly analysed as **phrasal complements** to the verbal head (Piñón 1995; É. Kiss 2002; Den Dikken 2004, a.o.).
- A verbal CC selects for **a single verbal particle**, which shows that the elements of a verbal CC in Hungarian cannot have independent complements:

(15) *János el-tett - (\*el-)vett a konyhában.*  
 John PRT-put.down PRT-pick.up the kitchen.in  
 ‘John whiled away the time by moving (lit. putting down - picking up) stuff around in the kitchen.’

(16) *A bizottság tagjai meg-hányták - (\*meg-)vetették a kérdést.*  
 the committee members PRT-threw PRT-cast the problem  
 ‘The members of the committee thrashed out (lit. threw-cast) the problem.’

(17) *A szocialista realizmus egyik legkifejezőbb alkotását  
 a tömeg darabokra törte - (\*darabokra) zúzta.*  
 the socialist realism one most.expressive artefact  
 the crowd pieces.into broke pieces-into crushed  
 ‘The crowd smashed into pieces the most expressive artefact of socialist realism.’

#### 3.2 Shared functional projections

- In Hungarian, a nominal CC can only associate with **a single possessor**, (18), and an adjectival CC can only be modified by **a single adverbial**, (19).
- (18) a. *Összegyűlt a falu fiatal-ja-öreg-je.*  
 assembled the village young-3SG-old-3SG  
 ‘The whole population of the village assembled.’

<sup>8</sup> The Hungarian National Corpus (Oravecz, Váradi & Sass 2014).

- b. \**Összegyűlt a falu fiatal-ja - a város öreg-je.*  
 assembled the village young-3SG the town old-3SG  
 intended: ‘The young people from the village and the old people from the town assembled.’

- (19) *Mari kicsit bús - (\*kicsit) komor volt.*  
 Mary slightly gloomy slightly sorrowful was.  
 ‘Mary was a bit sad.’ (lit. ‘Mary was a bit gloomy-sorrowful.’)

- Similarly, a modifying **adjective** necessarily applies **to both elements of a CC**, as the example (20) from Khanty shows:

- (20) *Mə: ənət sv:rt-γən v:γər-γən qv:tt-əm.*  
 1SG big pike-DU ide-DU catch-PST.1SG  
 ‘I caught a big pike and [a big] ide.’ (NOT: ‘I caught a big pike and an ide.’)

→ Combined, (15)-(20) show that the members of a CC share their complements as well as modifiers and higher functional projections.

### 3.3 Merge & labelling

We propose that the derivation of a phrase containing a CC (e.g., a verbal one) proceeds as follows:

- Both elements of the CC are part of the numeration as **separate elements**.
- They are combined in syntax via **symmetric Head-Head Merge** ( $[_\alpha H_1 H_2]$ ).
- The two heads are **equidistant from  $\alpha$** , but since they contribute the same category, this unresolved competition does not constitute a problem:  $H_1$  and  $H_2$  together contribute the category  $V$  to  $\alpha$  (e.g.  $[_V adv+veszV]$ ).
  - This is in line with Chomsky (2013: 43), who shows that **the labelling problem does not arise** if the two heads are (non-accidentally) identical in a relevant respect, providing the same label.
- In case the two heads have identical subcategorization properties and theta-grids, they, **together, project the (extended) VP** (the elements of a CC indeed always do). Otherwise, the derivation crashes.
- Further on,  **$\alpha$  functions as a single head** for the purposes of e.g., movement, with the potential exception of post-syntactic suffixation via M-Merger, for which both heads are visible.

\*\*\*

Note that we assume that the members of a CC form a symmetric structure.

- Regular **coordinated expressions** in generative theory are assumed to have **an asymmetric structure**; for empirical and theoretical arguments, see Munn (1993).
- In the case of Khanty and Hungarian CCs, however, neither the empirical evidence of asymmetry, nor the most powerful theoretical arguments for it are valid:
  - There is **no conjunction** between the co-compounded elements that would form a closer unit with one of them, thereby establishing an asymmetric c-command relation between them or a hierarchical prosodic unit.
  - Theoretically, symmetric structures are claimed to be avoided by natural languages because the resulting constituent **cannot be labelled** (e.g., Kayne 1994; Di Sciullo 2002; 2005): the labelling algorithm is looking for a lexical head, but in a symmetric structure minimal search is ambiguous, locating two heads.
  - However, as Chomsky (2013: 43) argues, this problem does not arise if **the two heads** are (non-accidentally – Marcel den Dikken, p.c.) **identical** in a relevant respect,

providing **the same label**. Members of a CC necessarily share the same category and are also strictly identical featurally, so that the labelling problem does not arise.<sup>9</sup>

### 3.4 Linearization

A further problem of **symmetric structures** is that they cannot be **linearized**.

- We tentatively assume that the linearization of a co-compound is **free in syntax**.
- It may be determined **semantically**, based on importance or primacy -- e.g., in the case of *eat-drink* in (2a).
- Alternatively, it can be determined **phonologically**.
  - In Hungarian co-compounds, for example, **the front-vowel member precedes the back-vowel one**: *üt-vág* ‘beat-chop’, *tejbe-vajba fürdet* ‘in milk – in butter bathe = to provide for luxuriously’,
  - and **the shorter member precedes the longer one**: *boldog-boldogtalan* ‘happy-unhappy = everybody’ (Sóskuthy 2012; Patay 2018; Borise & É. Kiss 2022)
  - Note that this pattern is relatively **easy to override** (e.g., in poetry, for reasons of rhyme or metre), which supports the idea of a syntactically free, phonologically conditioned linearization:

- (21) *Fürödtünk vajba-tejbe, de én nem akartam, hogy én legyek az /*  
 bathed butter.in-milk.in but I not wanted that I be that  
 ‘We were having the time of our lives (bathing in butter-milk) but I did not want to be the one  
*Akit a kalapácsos tegnap este először vág fejbe.*  
 whom the person.with.hammer yesterday evening first hit head.in  
 whom the guy with the hammer would hit in the head first yesterday evening.’

(lyrics of the song *Ki csinál rendet?* [Who shall create order?] by the punk rock band Tankcsapda)

### 3.5 Not exocentric compounds

- **An exocentric compound**: one that is not a hyponym of one of its elements, e.g., in English: *redhead, killjoy, pickpocket*, etc.
- *Pace* Scalise et al. (2009), we argue that CCs are **not exocentric compounds** as they are endocentric along all three dimensions identified by Scalise et al. (2009):
  - they are **categorically endocentric**: the constituents in head position impose their categorial features on the whole construction;
  - they are **morphologically endocentric**: the morphological features of the whole construction are identical to the morphological features of its internal constituents;
  - they are **semantically endocentric**: their meaning/semantic type can be compositionally derived from the type (and meaning) of their constituents.
- Scalise et al. (2009) consider (nominal) CCs morphologically exocentric because, in many languages, the gender of the CC is **neuter** when the constituents have **non-identical gender**. Since Uralic languages have no grammatical gender, this switch is not attested in Hungarian or Khanty; accordingly, there is no reason to assume that CCs are morphologically exocentric.<sup>10</sup>

<sup>9</sup> Chomsky (2013) posits the requirement of *non-accidental* identity in order to explain why two DPs merged in an XP-YP structure (e.g. in a small clause with a predicate nominal) cannot be labelled as such (and as a consequence, there is need for movement to break up the symmetrical construction). A moot point here is what exactly counts as non-accidental identity. We contend that members of a CC do qualify as non-accidentally identical since they are strictly required to be categorially and featurally fully identical (even though this identity is not conditioned by agreement).

<sup>10</sup> Even in languages where such a switch occurs, it appears to be rule-based: switch to neuter iff the gender of the constituents is non-identical – arguably, this could be handled as a resolution mechanism of conflicting feature values and does not necessitate the stipulation of an external morphological head.

#### 4. Against coordination

We argue against approaching CCs as instances of **asyndetic coordination**, based on several pieces of evidence:

- Overt **coordinators** are prohibited in CCs, as was shown in (3);
- Stipulating a silent conjunctive or disjunctive coordinator between CCs would lead to **interpretations** that are empirically unavailable and would fail to produce interpretations that are empirically available;
- The observed pattern of **dual marking** in Khanty is incompatible with conjunctive (or disjunctive) coordination.

##### 4.1 Interpretation without coordination

- In order for the sentence containing a verbal CC to be interpretable as **a single proposition at LF**, the two heads need be reconciled in terms of their semantic contribution.
- As we will argue below, this reconciliation happens not via coordination (disjunction or conjunction), but rather, through **substitution** by an immediate shared superordinate term.
- This is most obvious if one considers so-called collective co-compounds (a term introduced by Wälchli 2005): CCs where the two terms are taxonomic sisters such as in *süt-főz* fry-cook/boil ‘prepare dish(es)’.
- Consider (22), as an answer to the question *What is John doing right now?*

(22) PF: *János a konyhában süt-főz.*  
John the kitchen.in fries-cooks.  
‘John is preparing dishes in the kitchen.’

- Crucially, (22) can be uttered truthfully if John is in the kitchen preparing some dish (e.g., a cold salad) without either frying it (boiling it in oil or fat) or cooking it (boiling it in water). This means that LF<sub>1</sub> and LF<sub>2</sub> are not correct representations of the meaning of (22), however, LF<sub>3</sub> is:

(23) PF: *János a konyhában süt-főz.*  
John the kitchen.in fries-cooks.  
✗ LF<sub>1</sub>: FRY(**John**) ∧ COOK(**John**)  
‘John is frying and cooking some dishes in the kitchen.’  
✗ LF<sub>2</sub>: FRY(**John**) ∨ COOK(**John**)  
‘John is frying or cooking some dishes in the kitchen.’  
✓ LF<sub>3</sub>: DISH-PREPARE(**John**)  
‘John is preparing dishes in the kitchen.’

- We propose that the **semantic reconciliation** happens by way of a strictly local operation which searches for the common immediate **superordinate concept** of the two elements.
- In hyponym-hypernym graph terms: the operation checks whether there is a node which immediately dominates the two elements – if so, at LF, the two-word sequence is replaced with the semantic feature bundle associated with that node:
  - *fiú-lány* ‘boy-girl’ [human, non-adult, male]-[human, non-adult, female] is replaced with the feature bundle [human, non-adult]
  - *süt-főz* ‘fry-cook’ [prepare dish, by boiling in fat/oil]-[prepare dish, by boiling in water] is replaced with: [prepare dish]
- Sometimes there is an actual word that corresponds the superordinate feature bundle: [human, non-adult] has the lexicalization *gyerek* ‘child’, but this is not necessary: [prepare dish] has no lexicalization in Hungarian
- Note also that domination does not need to be exhaustive (cf. again *süt-főz*).



- If the search fails, **the derivation crashes at LF**. The strict locality of the search explains why only synonyms, antonyms, taxonomic sisters and reverses are allowed in CCs.<sup>11</sup>

\*\*\*

- Similar kind of evidence is available in Khanty, where the concept of mankind/humanity/people in general can be expressed as *ruť-qǎntəy* ‘Russian-Khanty’. It is clear, though, that *ruť-qǎntəy* is not a coordination of the two terms, either on a conjunctive or disjunctive reading, since, in (26), the most salient representatives of the humanity that can kill the protagonist are Selkups. This is schematized in (29):

- (24) (Context: Khanty are at war with Selkups. The Khanty protagonist is surrounded by Selkup warriors and thinks to himself):
- nām-a ruť-nə qǎntəy-nə wāl-təyə ja kič-əm əntə wǎl-a.*  
 in.vain Russian-LOC Khanty-LOC kill-INF PCL desire-1SG NEG be-PRS.3SG  
 ‘It’s of no use, being killed by other people.’ (Lit.: ‘My desire to be killed by other people in vain doesn’t exist.’) (Márta Csepregi, personal archive)

- (25) ✗ LF<sub>1</sub>: RUSSIANS Λ KHANTY  
 ‘It’s of no use, being killed by Russians and Khanty.’  
 ✗ LF<sub>2</sub>: RUSSIANS V KHANTY  
 ‘It’s of no use, being killed by Russians or Khanty.’  
 ✓ LF<sub>3</sub>: HUMANITY/PEOPLE  
 ‘It’s of no use, being killed by other people.’

\*\*\*

- There is evidence that in the case of nominal CCs too, we have a substitution with the **superordinate terms** instead of coordination (**either disjunction or conjunction**). Consider:

- (26) *Mə:fə v:ləŋ kv:pə-fə:j ji:nʰ?*  
 Masha morning coffee-tea drink.PST.3SG  
 ‘Did Masha drink something?’

- This sentence cannot be interpreted as a constituent question (Was it coffee or tea that Masha drank?), **only as a yes-no question**.
- Also, the question can be answered in the affirmative if Masha drank only coffee or only tea (or indeed some other morning drink)
- The (un)available readings are the following:
  - ✗ ‘Did Masha drink coffee or tea in the morning?’
  - ✗ ‘Did Masha drink coffee and tea in the morning?’
  - ✓ ‘Did Masha drink something (some drink) in the morning?’

#### 4.2 The interpretation of dual marking in Khanty

- **Dual marking**, which is common on countable nouns in CCs in Khanty, is not ordinary number marking, in that a dual-marked nominal does not refer to two items. Instead, each of the dual-marked members of a CC refers to **one item**, as in (27), repeated from (13):

<sup>11</sup> The idea that some CCs (so-called collective CCs in the terminology of Wälchli 2005) denote a superordinate term of the elements is not in itself a novelty: this has been discussed by Wälchli (2005) and also by Szabolcsi (1976) in the context of Hungarian and Khanty. What is novel in our approach is (i) the extension of this idea to other types of CCs as well and (ii) the integration of this insight into a generative model of syntax and semantics.

(27) *Mv: qut-əm-v βoqi-yən v:mp-yən jv:βət-yən.*  
 1SG house-1SG-LAT fox-DU dog-DU come-PST.3DU  
 ‘A fox and a dog came to my house.’

- In this, CCs are markedly different from coordination structures, in which **dual is used to refer to two items**:

(28) *Ma qut-əm-v (kə:t) βoqi-yən pə:nə (kə:t) v:mp-yən jv:βət-yən.*  
 1SG house-1SG-LAT two fox-DU and two dog-DU come-PST.3DU  
 ‘Two foxes and two dogs came to my house.’

- Similarly, the **strict morphological parallelism** between the members of CC, as shown in section 2.1, is not easily explainable by the coordination analysis.<sup>12</sup>

## 5 Morphology: in favour of lowering

- We propose that the morphological properties of CCs (e.g., strict morphological parallelism) are a by-product of **agreement** of both members of a CC with a c-commanding head.
- Adopting Borise & É. Kiss’s (2022) analysis of CCs, we assume that the parallel morphology result from **post-syntactic M(orphological)-Merger** (Halle & Marantz 1993), whereby the suffixes are lowered to the heads post-syntactically, prior to lexical insertion.
- An interesting piece of data from Hungarian concerns verbal CCs falling into **two different paradigms** in terms of the 2SG subject agreement suffix (-sz vs. -l)
- A case in point is *oszt-szoro* (divide-multiply ‘make calculations, consider one’s options’). In isolation, the 2SG forms for the two verbs are the following (frequencies are from the Hungarian WebCorpus 2.0, Nemeskey 2020):

(29) a. *oszt-asz* standard form 98% (386 hits)  
 b. <sup>†</sup>*oszt-ol* dialectal 2% (8 hits)

(30) a. *\*szoro-asz* unattested 0% (0 hits)  
 b. *szoro-ol* standard form 100% (41 hits)

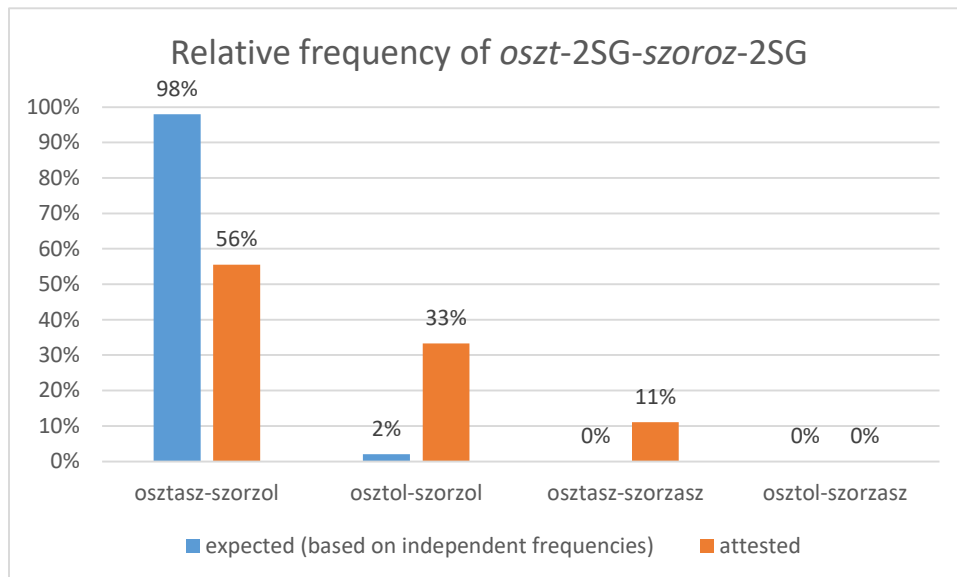
- In a CCs, the following forms are attested:

(31) a. *osztasz-szoro* 15 hits  
 b. *osztol-szoro* 9 hits  
 c. *osztasz-szoroasz* 3 hits  
 d. *osztol-szoroasz* 0 hits

- This distribution suggests that there is a strong preference for the suffixation to be **identical** not only featurally (2SG) but also **in terms of phonological realization**: almost half of the attestations involve forms where the suffix matches phonologically.<sup>13</sup>

<sup>12</sup> A reviewer raised the following point here: if co-compounding is indeed as straightforward as we argue that it is, why is not more widespread cross-linguistically? We have two tentative answers. Firstly, co-compounding is a very limited strategy: the members need to be semantically closely related, need to have identical suffixes and cannot be modified separately (so *the car and the dog, your dog and my cat* or *one dog and three cats* are ineffable using CCs). Syndetic coordination is vastly more expressive, and, once a language has developed syndetic coordination, there may be a tendency for co-compounding to recede and even disappear. Secondly, it is not obviously the case the co-compounding is easy computationally: the search for a superordinate term may be more a demanding task than calculating the conjunction (or disjunction) of two expressions.

<sup>13</sup> This pattern lends support to our analysis which relies on m-merger, and in particular the \*lowering\* of inflectional suffixes post-syntactically. An alternative analysis (raised by one of our reviewers) would be to assume that the co-heads of the co-compound have their inflections on them at the point of Merge, with the (identical) feature bundles of the inflectional suffixes coalescing on the co-compound and undergoing checking/valuation en bloc against an outside probe.



- This preference is strong enough to lead to the production of forms that are otherwise ungrammatical: \**szorzasz* (for all speakers) and %*osztol* (for vast majority of speakers).
- In Optimality Theoretical terms, this can be analysed as the interplay of two constraints: STEM-SUFFIX MATCH and SUFFIX-SUFFIX MATCH.
- Within the *osztasz*-dialect:

<i>oszt-2SG-szoros-2SG</i>	STEM-SUFFIX MATCH	SUFFIX-SUFFIX MATCH
a. <i>osztasz-szorzol</i>		*
b. <i>osztol-szorzol</i>	*!	
c. <i>osztasz-szorzasz</i>	*!	
d. <i>osztol-szorzasz</i>	*!*	*

<i>oszt-2SG-szoros-2SG</i>	SUFFIX-SUFFIX MATCH	STEM-SUFFIX MATCH
a. <i>osztasz-szorzol</i>	*!	
b. <i>osztol-szorzol</i>		*
c. <i>osztasz-szorzasz</i>		*
d. <i>osztol-szorzasz</i>	*!	**

- Within the *osztol*-dialect:

<i>oszt-2SG-szoros-2SG</i>	STEM-SUFFIX MATCH	SUFFIX-SUFFIX MATCH
a. <i>osztasz-szorzol</i>	*!	*
b. <i>osztol-szorzol</i>		
c. <i>osztasz-szorzasz</i>	*!*	
d. <i>osztol-szorzasz</i>	*!	*

However, a very strict morphological parallelism between the terms of a co-compound to the extent of phonologically identity (modulo vowel harmony) is more compatible with post-syntactic lowering.

<i>oszt-2SG-szöröz-2SG</i>	SUFFIX-SUFFIX MATCH	STEM-SUFFIX MATCH
a. <i>osztasz-szörözöl</i>	*!	*
b. <i>osztol-szörözöl</i>		
c. <i>osztasz-szörözasz</i>		*!*
d. <i>osztol-szörözasz</i>	*!	*

- Note that *osztasz-szörözöl* can only be derived assuming STEM-SUFFIX MATCH >> SUFFIX-SUFFIX MATCH, and conversely, *osztasz-szörözasz* can only be derived assuming SUFFIX-SUFFIX MATCH >> STEM-SUFFIX MATCH
- Similar patterns are attested with *búzol-balaszta* and *sütsz-főzől* (where *búzol-balaszta* and *sütöl-főzől*) also strikingly frequent) and, to a lesser extent, with *adsz-veszel* (*adol-veszel*).

## 6. Further research directions

- Can this analysis be extended to so-called **echo words**?
- These behave syntactically like CCs, but crucially, the first element is (i) a non-word (a slightly modified version of the second element in terms of its phonological makeup, cf. Sóskuthy (2012)), (ii) its meaning is, if any, identical to that of the second element, and (iii) echo words are perceived to be diminutive versions of the second element:

- (32) a. *riheg-röhög*  
nonword-laugh  
'giggles'
- b. *ici-pici*  
non.word-small  
'tiny'

- Can this analysis be extended to **reduplication**?

- (33) a. *cukor-cukor* sugar-sugar 'real sugar' as opposed to some artificial sweetener which might also count as *cukor* 'sugar' under a laxer precision standard
- b. *most-most* now-now 'precisely now' as opposed to *most* 'now, now-ish'
- c. *búzta-búzta* draw.3SG.OBJ-draw.3SG.OBJ 'kept drawing it intensively/for an extended time' as opposed to *búzta* 'was drawing it'

## 7. Conclusion

Based on evidence from Khanty and Hungarian, we found that (certainly in these languages):

- The term coordinative compounds is something of a misnomer, as CCs, in fact, lack coordination and they are not compound words stored as such in the Lexicon;
- Co-compounding is a syntactic operation involving the Merge of two heads;
- Co-compounding (the existence of two-headed phrases) is compatible with standard generative assumptions as long as the two heads meet certain requirements (identical category and thematic and selectional requirements, close semantic relationship): exactly the ones that CCs are, in fact, subject to.

## References

- Bárczi, Géza & László Országh. 1959. *A magyar nyelv értelmező szótára I–VII*. [The Explanatory Dictionary of the Hungarian Language]. Budapest: Akadémiai Kiadó.
- Borise, Lena & Katalin É. Kiss. 2022. The emergence of conjunctions and phrasal coordination in Khanty. *Journal of Historical Linguistics* Online First. <https://doi.org/10.1075/jhl.21016.kis>.
- Chomsky, Noam. 1970. Remarks on Nominalization. In Roderick A. Jacobs & Peter S. Rosenbaum (eds.), *Readings in English Transformational Grammar*, 184–221. Waltham, MA; Toronto; London: Ginn and Company.
- Chomsky, Noam. 1993. *Lectures on government and binding*. Walter de Gruyter.
- Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49.
- Den Dikken, Marcel. 2004. Agreement and “clause union.” In Katalin É. Kiss & Henk van Riemsdijk (eds.), *Verb Clusters: A study of Hungarian, German and Dutch*. John Benjamins Publishing.
- Di Sciullo, Anna Maria. 2002. *Asymmetry in Grammar*. Amsterdam ; Philadelphia: John Benjamins Pub.
- Di Sciullo, Anna Maria. 2005. *Asymmetry in Morphology*. Cambridge, MA: MIT Press.
- É. Kiss, Katalin. 2002. *The Syntax of Hungarian*. Cambridge: Cambridge University Press.
- Halle, Morris & Alec Marantz. 1993. Distributed Morphology and the Pieces of Inflection. In Kenneth Hale & S. Jay Keyser (eds.), *The View from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111–176. Cambridge MA: MIT Press.
- Kayne, Richard S. 1994. *The antisymmetry of syntax*. Vol. 25. Cambridge MA: MIT Press.
- Lewy, Ernst. 1911. *Zur finno-ugrischen Wort- und Satzverbindung*. Göttingen: Vandenhoeck.
- Lichte, Timm. 2021. Against strict headedness in syntax. *Journal of Language Modelling* 9(2). 291–348.
- Munn, Alan Boag. 1993. *Topics in the syntax and semantics of coordinate structures*. University of Maryland at College Park PhD Thesis.
- Oravecz, Csaba, Tamás Váradi & Bálint Sass. 2014. The Hungarian Gigaword Corpus. In *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC'14), Reykjavik, Iceland, 1719–1723*. European Language Resources Association (ELRA).
- Patay, Fanni. 2018. A magyar ikerszók optimalitáselméleti elemzése. In György Scheibl (ed.), *LingDok 17. Nyelvészdoktoranduszok dolgozatai*, 67–84. Szeged: Szegedi Tudományegyetem Nyelvtudományi Doktori Iskola.
- Piñón, Christopher. 1995. Around the progressive in Hungarian. *Approaches to Hungarian* 5. 153–190.
- Scalise, Sergio, Antonio Fábregas & Francesca Forza. 2009. Exocentricity in compounding. *Gengo Kenkyu (Journal of the Linguistic Society of Japan)* 135. 49–84.
- Sóskuthy, Márton. 2012. Morphology in the extreme: echo-words in Hungarian. In Ferenc Kiefer & Zoltán Bánréti (eds.), *Twenty Years of Theoretical Linguistics*, 123–143. Budapest: Tinta Kiadó.
- Szabolcsi, Anna. 1976. On a grammatical aspect of Ostyak parallelism. *Language Sciences* 41. 17–19.
- Wälchli, Bernhard. 2005. *Co-Compounds and Natural Coordination*. Oxford: Oxford University Press.