

The emergence of conjunctions and phrasal coordination in Khanty

Abstract

Prior to widespread contact with Russian, Khanty (Uralic; Finno-Ugric) did not have overt conjunctions or phrasal coordination. Instead, Khanty texts from late 19th-early 20th centuries only include examples of conjunction-less clausal juxtaposition, which was used for both clausal and phrasal coordination. By comparing Khanty texts over the 20th century, we trace the emergence of overt conjunctions and coordination of phrasal constituents. We show that overt conjunctions first appear in the context of clausal coordination; then, coordination of smaller phrases with overt conjunctions became possible. Based on novel elicitation data, we demonstrate that, in contemporary Khanty, (i) overt conjunctions are commonplace, and (ii) coordination of sub-clausal phrases is derived via phrasal coordination, as opposed to clausal coordination followed by conjunction reduction. We also show that ellipsis in Khanty is generally quite restricted. Based on this diachronic picture, we conclude that coordination of phrasal constituents only emerged in Khanty once overt conjunctions became available. We derive this correlation from the Maximize On-line Processing principle (Hawkins 2004), and show that this maxim, usually invoked in the context of speech planning and production, can be successfully applied to modelling language change.

Keywords: Khanty, Uralic, Finno-Ugric, coordination, conjunctions, conjunction reduction, ellipsis, co-compounds, on-line processing.

1. Introduction

1.1 Background and goals

This paper demonstrates that phrasal coordination and conjunction reduction appeared in Khanty upon the emergence of conjunctions in the course of the 20th century, and proposes an account for this correlation.

The existence of languages without syndetic coordination is well-known. Mithun (1988) claims that conjunctions in many languages have been grammaticized quite recently, and argues – based on Chafe's (1985) analysis of spoken English, and on evidence from African languages (Welmers 1973), Native American languages (Craig 1977; Cole 1982; Suárez 1983), West-Siberian Chukchee (Bogoras 1922), and Kamchadal (Worth 1961) – that syndetic coordination has arisen parallel with literacy, or parallel with the emergence of bilingualism involving a literary language. This is what we attest in the recent history of Khanty (Uralic; Finno-Ugric) as well.

In Khanty texts recorded at the beginning of the 20th century, prior to general Russian-Khanty bilingualism, coordinated clauses are merely juxtaposed. What is remarkable about earliest attested Khanty – and differentiates it from the languages in Mithun's (1988) sample – is that phrasal coordination is practically non-existent; clausal juxtaposition is used in its stead as well. Lack of phrasal coordination also means lack of ellipsis in coordinated sentences that are partially identical. Additionally, noun phrases denoting closely related concepts could be combined in compound-like constructions, called co-compounds by Wälchli (2005). Co-compounding was used as a substitute for overt coordination in early 20th century Khanty (and is still used as a supplementary strategy today).

In later texts reflecting the growing influence of Russian, overt conjunctions and disjunctions appear, first between clauses. With some delay, phrasal coordination emerges: we attest conjunctions and disjunctions in the contexts of coordinated VPs, arguments, and clausal adjuncts. Coordinated adjuncts may be non-adjacent, which indicates that conjunction reduction has also become part of Khanty grammar. However, NP-internal ellipsis and coordination of nominal heads are still illicit. This suggests that overt coordination is introduced into the syntactic structure in a top-down fashion, first targeting larger constituents and then smaller ones.

After coordination of smaller constituents has become possible, we show that, in the default cases in today's Khanty, it does not arise from coordination of clauses followed by ellipsis, but instead is derived by coordinating phrasal constituents. Additionally, we show that ellipsis in Khanty is generally quite restricted. Taken together, we interpret the facts outlined above as pointing to a correlation between the emergence of overt conjunctive particles and the emergence of true phrasal coordination. We derive this correlation from a processing maxim, the principle of Maximize On-line Processing proposed by Hawkins (2004). While this principle is usually taken to govern speech planning and production, we demonstrate that it can be successfully applied to modelling language change by applying it to the emergence of overt conjunctions and phrasal coordination in Khanty.

1.2 Maximize On-line Processing

Hawkins (2004:51) proposed that the architecture of grammar is governed by the following rules: 'Express the most with the least' and 'Express it earliest'. He formulated the latter as the principle of Maximize On-line Processing, provided in (1). 'On-line Property to Ultimate Property ratios' here refer to increasing processing efficiency by "selecting and arranging linguistic forms so as to provide the earliest possible access to as much of the ultimate syntactic and semantic representation as possible" (Hawkins 2004:9), and the intuition behind (1) is that "many preferences [in grammar] appear to be correlated with the earlier assignment of common properties, in one ordering or structural variant, v. their later assignment in another." (Hawkins 2004:50).

(1) Maximize On-line Processing

The human processor prefers to maximize the set of properties that are assignable to each item X as X is processed, thereby increasing On-line Property to Ultimate Property (OP/UP) ratios. The maximization difference between competing orders and structures will be a function of the number of properties that are unassigned or misassigned to X in a structure/ sequence S, compared to the number in an alternative.

Hawkins claims that the failures of assigning syntactic properties and structural relations to constituents in the course of the online parsing of a sentence can be quantized, counted, and compared. This is done by adding up the following factors:

(2) Unassignment factors

- a. the number of words and phrases that undergo some temporary unassignment of properties on-line, compared with an alternative structure/sequence in which the relevant properties are immediately assignable;
- b. the number of any mother–daughter attachments that are temporarily unassignable to the words and phrases in (a);
- c. the number of any relations of combination or dependency that are temporarily unassignable to the words and phrases in (a).

(3) Misassignment factors (selected)

- a. the number of words and phrases that undergo some temporary misassignment of properties on-line;
- b. the number of any additional dominating nodes that must be introduced into the syntactic tree when correcting the misassignments in (a);
- d. the number of any relations of combination or dependency that are temporarily misassigned to the words and phrases in (a);

The principle of Maximize On-line Processing is used to formalize and quantify the preference for structures like (4a) over those like (4b), given earlier recognition and introduction into the structure of the second PP (i.e., assignment of its properties):

- (4) a. *John* [_{VP} *went* [*to London*] [*in the late afternoon*]].
b. *John* [_{VP} *went* [*in the late afternoon*] [*to London*]].

The principle of Maximize On-line Processing was proposed as a rule that governs speech planning and production. Hawkins' idea was that the maxims that shape speech planning, in the long term, also shape the architecture of grammar itself. To the best of our knowledge, however, they have not yet been applied to the diachronic development of coordination. In the remainder of the paper, after laying out the diachronic picture, we demonstrate that phrasal coordination in earliest attested Khanty was unavailable due to the (i) absence of overt conjunctions, (ii) impoverished nominal morphology (no case or possession markers), and (iii) verb-finality. In this morphosyntactic context, asyndetic coordination of phrasal (e.g., nominal) constituents would potentially result in strings of nominals with low OP/UP ratios. Introduction of overt conjunctions increased the OP/UP ratios of sequences of noun phrases, which lifted the restriction on phrasal coordination.

1.3 Sources and methodology

The data from older stages of Khanty come from four corpora.¹ The oldest variety of Khanty considered here is described on the basis of a corpus of 4000 words, including four tales collected by Heikki Paasonen in the eastern Yugan area in 1901 (published by Vértes in 2001, annotated in the Ob-Ugric Database (OUIDB)). We also cite some Khanty examples recorded in the same period from Lewy's (1911) analysis of Finno-Ugric word and clause conjunction.

Next, the 6000-word autobiographical notes of K. M. Marenjanin, a speaker of the northern Sherkaly dialect, written in 1936 and published by Steinitz in 1989, were consulted. Marenjanin's notes are of special significance because, owing to their content, they are presumably closer to everyday speech, and exempt from conventionalized constructions to a greater degree than the folklore texts recorded earlier.

The next stage of Khanty is represented by a 3740-word corpus of texts collected by Rédei during a fieldwork in the northern Kazym area in 1964, annotated in the OUIDB. Finally, we considered a 4200-word corpus of Eastern Khanty texts collected in the 1990s by Márta Csepregi (Csepregi 1998; 2002), also annotated in the OUIDB.

Contemporary Khanty data come from elicitations with three female speakers of Eastern (Surgut) Khanty (age range: 21-69 y.o.). They were asked to rate pre-constructed sentences in Surgut Khanty and/or suggest their own paraphrases. Additionally, a small spoken corpus, collected from two female speakers (age: 29 and 68) in 2017 by Márta Csepregi and Katalin Gugán was consulted for examples of coordination.

The paper is organized as follows: Section 2 describes coordination in Khanty texts recorded in 1901, displaying no conjunctions and no phrasal coordination except for nominal complexes called co-compounds. Section 3 demonstrates the gradual evolution of syndetic coordination, i.e., the emergence of conjunctions (3.1), their spreading followed by the emergence of phrasal coordination and conjunction reduction (3.2), and the generalization of these constructions (3.3) by analyzing texts from 1936, from 1964 and from the 1990s. Section 4 discusses coordination in contemporary Khanty. Section 5 proposes the final account of the observed phenomena. Section 6 concludes.

¹ There do not seem to be any differences with respect to the processes described here among the different Khanty dialects. Therefore, 'Khanty' is used to refer collectively to the Khanty varieties that data are available for. The relevant aspects of individual Khanty varieties are described in the context of particular examples.

2. Coordination in Old Khanty

2.1 Asyndetic clausal coordination

In Khanty texts recorded around the beginning of the 20th century, prior to obligatory Russian schooling and Russian-dominant bilingualism, we do not attest syndetic coordination. In the Heikki Paasonen corpus, consecutive or simultaneous events are described by juxtaposed clauses without overt conjunctions. The commas linking these clauses presumably indicate that the juxtaposed clauses represented a single Intonational-Phrase-like prosodic unit in the oral presentation.

- (5) a. [βv:jəx qu:t kəntf-təyə mənn-əs i:ki], [i:mi jaqqən qu:t].²
game fish search-INF go-PST.3SG man woman at.home stay.PST.3SG
'The man went hunting-fishing, the woman stayed at home.' (OUDB 1316)
- b. [ru:t βer-t-ətəy, βer-t-ətəy].³ [i:ttən torrəm pəttəylə-s].
boat do-PRS-SG<3SG do-PRS-SG<3SG at_night heaven darken-PST[3SG]
'He built [and] built the boat, in the evening the sky darkened.' (OUDB 1315)
- c. [mənn-əs], [pon noq tɐ:t-s-i], [qu:t tɐ:t-s-i],
go-PST.3SG fish_basket up pull-PST-PASS.3SG fish catch-PST-PASS.3SG
[tɐ:rəy tɐ:t-s-i], [sv:rt tɐ:t-s-i], [jɐ:β tɐ:t-s-i].
ruff catch-PST-PASS.3SG pike catch-PST-PASS.3SG perch catch-PST-PASS.3SG
'He left, the fish-basket was pulled up, fish was caught, ruff was caught, pike was caught, perch was caught.' (OUDB 1316)

Since Khanty is a pro-drop language, juxtaposed clauses with a shared null subject are non-distinct from juxtaposed predicate phrases. Such examples behave in the same way as full juxtaposed clauses: they do not have overt conjunctions either:

- (6) [qo:təm sv:t tʃv:ət mə-ss-əm], [kɐ:rkəm i:mi βə-ss-əm].
three hundred ruble give-PST-1SG hard-working woman take-PST-1SG
'I paid 300 rubles, I took a hard-working wife.' (OUDB 1316)

Uncertain quantities are expressed by the asyndetic coordination of complete clauses to be interpreted disjunctively:⁴

- (7) a. [qo:əpti βat-t-ət] [βv:nəpti βat-t-ət].
long live-PRS-3PL short live-PRS-3PL
'They live for a long time, [or] they live for a short time.' (OUDB 1313)
- b. [mənn-əs], [mənn-əs], [əj tɐ:ə-nə qv:t-s], [kɐ:t tɐ:ə-nə qv:t-s].
walk-PST.3SG walk-PST.3SG one place-LOC sleep-PST.3SG two place-LOC
sleep-PST.3SG
'He walked, he walked, he slept at one place, [or] he slept at two places.' (OUDB 1313)

² We adopt the spelling of the sources of the examples, in most cases the Ob-Ugric Database (OUDB), which uses IPA transcription. Examples from other sources are provided unaltered.

³ Morpheme complexes of the type SG<3SG cross-reference both an object (in this case, singular) and a subject (in this case, 3rd person singular). Only topical (specific and referential) objects elicit verbal agreement. Since Khanty is a pro-drop language, the second *βerlətəy* in (5b) could also be analyzed as a full clause with a pro subject and a pro object.

⁴ The coordination of clauses of this kind used to perplex scholars studying traditional Khanty texts, e.g. Steinitz (1941). Szabolcsi (1990) argues that such seemingly contradictory statements, which are typical in Khanty poetry, are coordinate structures covering a semantic space by describing its two subfields.

The use of adverbs or particles as connectives is extremely rare. We attest a few occurrences of *pə:nə/pa:ne*, the locative form of *pa* ‘other’, presumably meaning something like ‘otherwise’, ‘on the other hand’. By now, *pa:ne* has developed into a conjunction corresponding to *and*, but these early texts still show no sign of its grammaticalization into a regular linking element:

- (8) *βεli βεl, pə:nə niob βεl.*
 reindeer kill-PST.3SG on_the_other_hand/and elk kill-PST.3SG
 ‘He killed reindeer and he killed elk.’ (OUBD 1316)

In disjunctive coordinate clauses expressing approximate quantities, the quantity expression is sometimes preceded by *məβ(ə)/müw*, originally an indefinite pronoun, the equivalent of ‘(some)what’. (9b), with the 2nd *məβə* following a pro-dropped subject and object, represents a context where the indefinite pronoun modifying quantity expressions could be reinterpreted as a disjunctive particle linking clauses.

- (9) a. *tot məβ əj qatl βat-s-əyən məβə kə:t qatlyən βat-s-əyən.*
 there about one day live-PST-3DU about two day-DU live-PST-3DU
 ‘There, they stayed for about one day, they stayed for about two days.’ (OUBD 1314)
- b. *əj lə:pət li:-s-tən, məβə kə:t lə:pət-yən li:-s-tən.*
 one week eat-PST-SG<3DU about two week-DU eat-PST-SG<3DU
 ‘They at it for one week, perhaps they ate it for two weeks.’ (OUBD 1316)

Additionally, instead of clausal juxtaposition, propositions describing simultaneous or consecutive actions are often connected by the subordination of one of the propositions in the form of a non-finite verbal projection: a converb (10a), or a present or past participle (10b) (see also Sipos 2015).

- (10) a. *jəŋk qv:nəŋ-nə n'e:βər li:β-min su:tfəytə-l-yən.*
 water bank-LOC foam eat-CVB walk.around-PRES-3DU
 ‘They walk around, eating foam on the river bank.’ (OUBD 1314)
- b. *i:mi βv:n-yə joβt-əm lə:t-nə ju:β toj-v qu:ŋt-əs.*
 woman near-TRNS come-PTCP.PST time-LOC tree top-LAT climb-PST.3SG
 ‘The woman having come close, he climbed to the tree top.’⁵ (OUBD 1315)

An early description of Khanty coordination by Lewy (1911: 12) is based on fieldwork carried out by Sergiy Patkanov in the southern dialect area, which was most exposed to Russian influence, and which was the first to disappear in the middle of the 20th century. Lewy also mentions the conjunctions *i* ‘and’ and *ali* ‘or’, borrowings of the Russian *i* and *ili*, that may occasionally occur in cases of sentence coordination. In contrast, the Paasonen-corpus we examined contains no instance of them, which is expected, given that they come from Eastern (Yugan) dialects, which did not yet have as much Russian influence.

2.2 Lack of phrasal coordination/conjunction reduction

Coordination via juxtaposition is only common on the sentence level; the juxtaposition of clausal sub-constituents is very rare. Apparently, in the typical traditional Khanty clause, each grammatical function is represented by an expression denoting a single entity. If we described the event in (5c) in English, the catching of different kinds of fish in the same fishbasket at the same time would be described by a single clause involving coordinate noun phrases; in Khanty, however, the same predicate is predicated about each kind of fish in a separate clause. Situations where multiple participants have the same property or perform the same action (11a), or

⁵ The disjoint reference of the *pro* subject of the main clause and the lexical subject of the participle is made clear by the context.

multiple objects are affected in a similar way by the same agent (11b) are expressed by multiple juxtaposed clauses. The lack of ellipsis and the resulting repetitiveness exemplified here may, at first sight, seem to be a rhetorical device used in folk tales to give the text a certain rhythm, and perhaps to help its memorization. However, the data in Sections 3.1 and 3.2 will show that the lack of phrasal coordination and conjunction reduction is a general phenomenon in pre-Russification Khanty; it is not restricted to a particular genre or register.

- (11) a. *torrəm ji:r βer-təyə mɔ:st-l, məγ ji:r βer-təyə*
 sky animal_sacrifice do-INF need-PRS.3SG earth animal_sacrifice do-INF
mɔ:st-l.
 need-PRS.3SG
 ‘A sky animal sacrifice needs to be made, an earth animal sacrifice needs to be made.’ (OUDB 1313)
- b. *pro pu:pi toβə qu:j-s-ətəγ, por βv:jəγ toβə qu:j-s-ətəγ, βoqu*
 bear there leave-PST-SG<3SG wolf there leave-PST-SG<3SG fox
toβə qu:j-s-ətəγ, tʃe:βər toβə qu:j-s-ətəγ, tʃu:t pu:rnə
 there leave-PST-SG<3SG hare there leave-PST-SG<3SG that after
kəmləγ qu:j-s-ətəγ.
 wolverine leave-PST-SG<3SG
 ‘He left behind the bear, he left behind the wolf, he left behind the fox, he left behind the hare, he left behind the wolverine.’ (OUDB 1315)

When a subject performs multiple actions, clause juxtaposition is usually indistinguishable from vP-juxtaposition, as a repeated subject tends to be represented by a silent *pro*; nevertheless, there are also examples with a repeated overt subject:

- (12) *v:ləŋ tɛ:tnə o:pi-l li:-s-i. tʃu:t pu:rnə*
 first time older_sister-3SG eat-PST-PASS.3SG that after
pi:tfənyəli kɛ:t-l-s-i, o:s pi:tfiŋyali noq li:-s-i jəppəγ-nə.
 little.bird catch-PST-PASS.3SG also little.bird up eat-PST-PASS.3SG owl-LOC
 ‘First her older sister was eaten. After that the little bird was caught, also the little bird was eaten by the owl.’ (OUDB 1314)

When a free choice indefinite subject is involved in disjunction, it is spelled out with each of the disjunctive predicates, as shown in (13). (13) has a complex structure: the juxtaposed clauses form closer units pairwise, with the indefinite subject of the first member of the pair co-referring with the *pro* subject of the second member:

- (13) *qo: əj βəti toj, u:lə βel-təγ, qo: əj tɛ:β taj-əs,*
 someone reindeer have.PST.3SG kill-PST-SG<3SG someone horse have-PST.3SG
u:lə βeltəγ, qo: əj mɛs taj-əs, u:lə βeltəγ, qo: əj v:tf
 kill-PST-SG<3SG someone cow have-PST.3SG kill-PST-SG<3SG someone sheep
taj-əs, u:lə βeltəγ pɔ:ri βer-təyə.
 have-PST.3SG kill-PST-SG<3SG feast do-INF
 ‘Someone had a reindeer, he killed it, someone had a horse, he killed it, someone had a cow, he killed it, someone had a sheep, he killed it to have a feast.’ (OUDB 1313)

A strategy of circumventing the coordination of subject or object NPs participating in the same event is to supply all but one of them with a comitative suffix.⁶

⁶ Stassen (2003) calls this the comitative strategy of NP-coordination, and claims that it is just as wide-spread crosslinguistically as the *and*-strategy. He argues that its use correlates with other grammatical properties. Namely,

- (14) *tʃu: qo: i:mi-l-nɛt nʲe:βrem-əl-nɛt tot v:məs-l-ət.*
 that man wife-3SG-COM child-3SG-COM there sit-PRS-3PL
 ‘That man is sitting there with his wife, with his children.’ (OUDB 1313)

The corpus of 4000 words does not contain any instance of adjective or adverb coordination. If two modifiers modify the same constituent, the modified constituent is repeated. Example (15a) involves repeated VPs with adverbial modifiers of the same type. As illustrated by (15b), the juxtaposition of two different kinds of adverbial adjuncts (an approximative and a manner adverbial) is also avoided. These types of low adverbials usually represent new information, and as Chafe (1987) observed, speakers typically introduce only one major piece of information at a time in spoken language.

- (15) a. *i:ttən li:tot əntem, v:ləŋ li:tot əntem.*
 in.the.evening food NEG.exist in.the.morning food NEG.exist
 ‘There is no food in the evenings, there is no food in the mornings.’ (OUDB 1313)
- b. *nɛ:j-nɛm v:məst-l səjlək v:məst-l.*
 fire-APPR sit-PRS.3SG silently sit-PRS.3SG
 ‘He sits next to the fire, he sits silently.’ (OUDB 1313)

2.3 Co-compounding

Khanty has its own way of combining expressions that have the same grammatical function in a single clause, as illustrated by *βv:jəx qu:l*, ‘game fish’, i.e., ‘animals hunted on land and in water’ in example (5). Following Wälchli (2005), who described this construction type in Mordvin (Uralic) and some (East-)Asian languages, we call it ‘co-compounding’. This strategy is more restricted than regular phrasal coordination, and the members of a co-compound display a semantically, structurally, and prosodically stronger bond than that attested between coordinated expressions. Semantically, the members of a co-compound denote closely associated concepts. According to Wälchli (2005), they may form an additive co-compound (an exhaustively listed set: *father-mother, hand-foot*), a generalizing one (extreme opposites: *day-night*), a collective one (a whole that comprises all referents that have the properties exemplified by the two words: e.g., *milk butter* in Chuvash means ‘dairy products’), or a synonymic one.

Most co-compounds in Khanty represent the additive relation, denoting either “natural” two-member sets such as *hands-feet* (16a), *rain-wind* (16b), *woman-man* (17b), *house-barn* (17c), or temporary pairs determined by the given situation such as a fox and a hare acting jointly in a fairy tale (17b). The co-compound *game-fish* in (16c) has a collective meaning.

- (16) a. *kət-il kər-il i:lə mo:ritə-min təβt-v pan-tɛyə mɔ:st-l*
 hand-3SG foot-3SG off break-CVB fire-LAT put-INF need-PRS.3SG
nʲe:βrem-əl-nɛt.
 child-3SG-COM
 ‘With his hands [and] feet broken, he needs to be put into the fire together with his child.’ (OUDB 1313)
- b. *torrəm jom-γə βv:t-γə jə-s.*
 weather rain-TRNS wind-TRNS become-PST.3SG
 ‘The weather turned into rain [and] wind.’ (OUDB 1313)
- c. *tʃi βɛlm-ɛl βv:jəx qu:l məŋɛti pi:tʃ*
 so kill-PTCP.PST-3SG game fish we.DAT fall-PRS.3SG
 ‘So the game [and] fish killed by him will fall on us.’ (OUDB 1313)

if a language employs comitative rather than *and*-type coordination, it is likely to be non-cased and non-tensed. Khanty, and the Ob-Ugric languages, in general, do not support this generalization.

- d. *ju:β po:m əj tʃe:mə li:-l-təγ*
 wood hay completely eat-PRS-SG<3SG
 ‘It consumes wood [and] hay completely.’ (OUDB 1313)

As regards their form, co-compounds consist of exactly two morphologically parallel juxtaposed nominals. If the two nominals refer to individuals or unique countable referents, both bear a dual suffix, and so does the verb agreeing with them, as in (17a–c).⁷ Note that in (17b), the event involves three referents with the same function. Only two of them can form a co-compounds, which is signalled by dual marking; the third one figures in a separate clause.

- (17) a. *i:mi-γən i:ki-γən pay taj-s-əγən.*
 woman-DU man-DU son have-PST-DU
 ‘The woman [and] the man had a son.’ (OUDB 1315)
- b. *βoqu-γən tʃe:βar-γən təβ jot-ɐl jə-s-γən, kəmləγ təβ*
 fox-DU hare-DU he with-3SG come-PST-3DU wolverine he
jot-ɐl jə-s.
 with-3SG come-PST.3SG
 ‘The fox [and] the hare came with him, the wolverine came with him.’ (OUDB 1315)
- c. *id’at ɔməs-ta xɔt-eyen, tabas-eyen-a jūxtot*
 opposite sit-PTCP.PRS house-DU barn-DU-LAT come.PST.3SG
 ‘He came to the house [and] barn sitting opposite.’ (Lewy 1911: 13)

When the members of a co-compound have no dual suffixes, they receive a kind reading. Such co-compounds – e.g., the subject in (16c) and (18), and the object in (16d) – elicit singular verbal agreement.

- (18) *ru:tʃ qantəγ qoʎnə sə:rnəm βat-lʃ*
 Russian Khanty how further live-PRS.3SG
 ‘How will the Russians [and] the Khanty live on?’ (OUDB 1315)

The dual suffix has an individuating function in co-compounds. Note, however, that a co-compound with dual marking on both members denotes two referents rather than four. Both of these facts are predicted if the co-compound is assumed to be dominated by a single NumP, whose [+dual] head agrees with both members of a co-compound. It also follows that the members of a co-compound always have the same number specification. Indeed, when the participants playing the same role in a situation have different cardinalities, they cannot form a co-compound; the comitative strategy is used instead, as in (14) above.

A co-compound may be also dominated by a PossP, with possessive agreement spelled out on both nominals – see (16a) and (19b). The members of a co-compound share the same case, and the case suffix is usually present on both – see (16b) and (19c). In (19a), however, case marking is only spelled out on the second conjunct. According to Lewy (1911), this is possible if the two nominals have another identical (dual or possessive) suffix repeated on both of them, as in (19a,b). When there is no other suffix to signal their morphological parallelism, the case morpheme must be spelled out on both nouns, as in (19c).⁸

⁷ In the Obdorsk dialect of Khanty, according to Nikolaeva (1999: 44), two juxtaposed NPs bear dual marking only if they denote animate participants related by a close (typically family) relationship. We have found that the dual marking is also present on the members of inanimate co-compounds if they have unique singular referents.

⁸ Examples (17c) and (19a,b) are from the extinct southern (Irtish) dialect of Khanty, but co-compounds of this type are attested in the northern dialects as well, including our Northern Khanty corpus from 1964, discussed in Section 3.2.

- (19) a. *īme-ŋen* *ige-ŋen-na* *ent* *tēvā-i*
 old_woman-DU old_man-DU-LOC not eat-PASS.PST.3SG
 ‘It wasn’t eaten by the old woman [and] old man.’ (Lewy 1911: 21)
- b. *kur-en* *uč-en-a...* *kerŋemtī-taŋen*
 foot-2SG clothes-2SG-LAT fall-IMP.3DU
 ‘They shall fall to your feet [and] clothes.’ (Lewy 1911: 21)
- c. *kur-a* *uč-a* *kerŋentīdāi-ŋen*
 foot-LAT clothes-LAT fall-PST.DU
 ‘They fell on feet, on clothes.’ (Lewy 1911: 21)

Examples (16c) and (17c) each contain a modifier. Although it is only spelled out once, it applies to both members of the co-compound. Structurally, therefore, it must be left-adjoined to a projection that dominates both nominals, such as NumP in (17c). A bare co-compound, e.g. one in (16c), is presumably dominated by an NP node. A question is what grammatical category the members of these co-compounds represent.

Wälchli (2005) claims that co-compounding is not restricted to nominals. Although the Paasonen tales only contain nominal co-compounds, our Khanty consultants have also produced adjectival, numeral and verbal co-compounds, for example:

- (20) a. *Mv: ənət ʃojən sv:rt qv:tł-əm.*
 1SG big fatty pike catch-PST.1SG
 ‘I caught a big-fatty pike.’
- b. *Mi:ʃe ki:t qaləm sv:rt tu:β.*
 Misha two three pike bring. PST.3SG
 ‘Misha brought two-three pikes.’
- c. *je:ji-γən man'i-γən li:k-kən jin't-γən*
 elder_brother-DU younger_brother-DU eat-PST.3DU drink-PST.3DU
 ‘The two brothers ate [and] drank.’

Crucially, the members of a co-compound – whether nominal, adjectival or verbal – never have complements of their own. Modifiers can only modify a co-compound as a whole.⁹ These facts suggest that the members of a co-compound are juxtaposed lexical heads (nouns, adjectives, numerals, or verbs) rather than juxtaposed phrasal projections. The bound inflectional morphemes (number, possessive agreement, and case suffixes on nominals, and tense and agreement morphemes on verbs), licensed by the functional projections dominating the co-compound, appear on the heads as a result of M(orphological)-Merger (Halle & Marantz 1993). This operation lowers the appropriate inflectional morphemes to the N⁰ or V⁰ heads post-syntactically, before lexical insertion. The inflectional morphemes are realized on both juxtaposed elements.

A further question is what structural relation the two members of a co-compound bear to one another. 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 co-compounds, 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,

⁹ The members of a nominal co-compound occasionally have modifiers of their own such as *Khanty man* – *Khanty woman* (= the (Khanty) people), *Red folk* – *White folk*, but these modifier + noun complexes form a non-productive class of lexicalized/ “frozen” expressions; they do not arise in discourse spontaneously.

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. In Khanty, the members of a co-compound necessarily share the same category, so that the labelling problem does not arise. 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 *house-barn* and *eat-drink*), or it can be determined phonologically. (In Hungarian co-compounds, for example, the front-vowel member precedes the back-vowel member – e.g., *üt-vág* ‘beat-chop’, *tejbe-vajba fürdet* ‘in milk – in butter bathe = to provide for luxuriously’, and the shorter member precedes the longer one, e.g., *boldog-boldogtalan* ‘happy-unhappy = everybody’.) Because of these considerations, we do not exclude the possibility that a co-compound is a symmetric structure consisting of two lexical heads, but leave definitive conclusions for further research.

2.4 Interim summary

Based on the texts recorded in 1901, prior to Russification, Khanty only had asyndetic clausal coordination, i.e., clauses were juxtaposed without any grammaticalized conjunctions. The use of linking adverbs was also extremely rare. The connectedness of juxtaposed clauses (marked using commas in the transcripts) was presumably indicated by prosody. Alternatively, propositions describing subsequent or simultaneous events could be connected asymmetrically, with one of them formulated as a participial phrase. The coordination of phrasal constituents was avoided – except for pairs of noun phrases representing a single concept or a pair of closely associated concepts, which could be combined into co-compounds. One of two noun phrases fulfilling the same (subject or object) function could also be marked by a comitative suffix. Apart from these cases, when an eventuality had multiple distinct participants fulfilling the same role, the description of the eventuality involved a separate proposition for each participant. In other words, if multiple distinct individuals had the same property or were involved in the same action, it was predicated separately about each of them. If an individual was associated with multiple properties or multiple actions, they were predicated about the individual in separate propositions forming separate clauses.

3. The emergence of syndetic coordination

3.1 Marenjanin's autobiographical notes from 1936: the first conjunctions

In the autobiographical notes of K. M. Marenjanin, written in 1936, the typical way of coordinating clauses is juxtaposition, too – see (21), where the two juxtaposed clauses describe consecutive events:

- (21) *Wɔš-na kāt mǔj χutəm χāt! ut-əs, kara pelək peta*
 city-LOC two or three day be-PST.3SG Kara side in.the.direction.of
kīt-s-a.
 send-PST-PASS.3SG

‘He was two or three days in the city, he was sent in the area of Karinsk.’

(Steinitz 1989: 140)

Maremjanin also links sentences that express consecutive or simultaneous events by turning one of them into a past or present participial phrase. The participial phrase often repeats the previous finite clause, as in (22).¹⁰

- (22) *Tǔjt-em jǎjk-a nǒp-əs. Tǔjt-em jǎjk-a nǒp-əm*
 sledge-1SG water-LAT drift-PST.3SG sledge-1SG water-LAT drift-PTCP.PST
tǎχə-na nǒwə jɔχ...
 place/time-LOC white folk
 ‘My sledge drifted into the water. At the time [of] my sledge having drifted into the water, the Whites...’
 (Steinitz 1989: 134)

Occasionally, we also attest coordinated clauses linked syndetically, by conjunctions *i/ij* ‘and’ or *a* ‘but’, borrowed from Russian:

- (23) a. *Jaj-em tɔw-ŋ-ət kǐr-əs ǐ manət teśat-s-ətte*
 brother-1SG horses-DU-3SG harness-PST.3SG and 1SG.ACC prepare-PST-SG<3SG
wɔš-a.
 city-LAT
 ‘My brother harnessed his two horses, **and** he prepared me for the city.’
- b. *tet-ǒt-na tusa tapət-s-əte, a tumət-sǒχ-na ǎnt*
 food-LOC well feed-PST-SG<3SG but clothes-overcoat-LOC not
tumpəptə-s-te.
 dress-PST-SG>3SG
 ‘He fed me well with food, **but** he didn't dress me in clothes [and] overcoat.’
 (Steinitz 1989: 153)

Whether the coordinated clauses are simply juxtaposed or linked by a conjunctive particle, they still display no conjunction reduction. The juxtaposed clauses of (24), for example, contain the same verb iterated four times; the coordination of the objects is avoided:

- (24) *Jǒnttə tɔw-ət wer-s-əm, jǒnttə uχt-ət wer-s-əm, jǒnttə*
 playing horse-PL make-PST-1SG, playing sledge-PL make-PST-1SG, playing
sese-t wer-s-əm, jǒnttə śɔrkan-ət wer-s-əm.
 looptrap-PL make-PST-1SG, playing bowtrap-PL make-PST-1SG
 ‘I made toy horses, I made toy sledges, I made toy looptraps, I made toy bowtraps.’

The clauses of (25), linked by the Russian conjunction *ǐ*, have different subjects and a shared VP. The VPs are spelled out without ellipsis; they are iterated three times:

- (25) *Tǎm zawod-ət fabrikaj-ət ũw-t-ət ǐ tǔtəŋ-tǔjt-ət ũw-t-ət*
 this works-PL factory-PL roar-MULT-3PL and fiery-sledge-PL roar-MULT-3PL
ǐ awtomobil-ət ũw-t-ət.
 and car-PL roar-MULT
 ‘These works-factories make noise, and railways make noise, and cars make noise.’
 (Steinitz 1989: 145)

The conjunction *i* appears sporadically between coordinated noun phrases, as well:

- (26) *Men jǎχ-s-amŋ sǒta-jǒχan-a ǐ muχtəŋ-jǒχan-a χǔt kǎš-ta.*
 1DU go-PST-1DU Sǒta-river-LAT and Muχtəŋ-river-LAT fish look.for-INF
 ‘We went to Sǒta-river **and** to Muχtəŋ-river to catch fish.’
 (Steinitz 1989: 139)

The few instances of nominals linked by *i* adhere to the same restrictions as co-compounds: they denote closely related concepts and are morphologically parallel: *unt mǎγ-na ǐ χǎr mǎγ-na*

¹⁰ This strategy of clause-chaining has also been pointed out in various Oceanic languages – see Ohori (2004) and Terrill (2004).

‘forest-LOC and tundra-LOC’ (Steinitz 1989: 185), *aše-m ǐ šatsaše-m* ‘father-1SG and grandfather-1SG’ (Steinitz 1989: 175). Nevertheless, co-compounds are still common and productive. Though some of them may be conventionalized or lexicalized (*town sūšn* ‘in spring - in autumn’, or *mōn tōwtaw mistaw* ‘our horses - cows’ in (27a) below), there are also spontaneous expressions such as *zawodat-fabrikajət* ‘factories-works’ in (25).

- (27) a. *mōŋ tōwtaw mīstaw šōras χu-na arə pūš*
 we horse-PL-1PL cow-PL-1PL merchant man-LOC many time
χorjat-ij-s-aj-ət.
 seize-FREQ-PST-PASS-3PL
 ‘Our horses [and] cows were many times seized by the merchantman.’
 (Steinitz 1989: 189)
- b. *nōwə jχχ wūrta jχχ t’at’as-ta pīt-s-ət.*
 white folk red folk fight-INF begin-PST-3PL
 ‘The Whites (and) the Reds began to fight.’ (Steinitz 1989: 136)

The members of the co-compound *mōn tōwtaw mistaw* ‘our horses - cows’ in (27a) share an overt possessor, which elicits agreement on both of them, and the members of the co-compound *tumət-sōχ-na* ‘clothes-overcoat-LOC’ in (23) share a case suffix – which supports the assumption that a co-compound consists of two juxtaposed NPs dominated by a single PossP and a single KP projection. The possessive agreement suffix is present on both NPs; the case suffix, however, sometimes appears only on the second NP. Unlike in the Irtish dialect illustrated in (16), the case suffix in Marenjanin’s text can be spelled out only once, whether or not the members of the co-compound bear identical agreement suffixes.

Above, (21) contains coordinated numerals connected by *mīj*, the Khanty equivalent of the indefinite/approximative pronoun *məβə* (*kāt mīj χutəm χāt!* ‘two or three days’). In example (9a) from 1901, a similar approximate timespan (*one or two days*) is still expressed by the juxtaposition of complete clauses (*There, they stayed for one day, they stayed for two days*). It is unclear whether the expression *kāt mīj χutəm χāt!* ‘two or three days’ is an instance of phrasal disjunction and conjunction reduction, with *mīj* functioning as a disjunctive particle, or the pattern ‘numeral *mīj* numeral’ grammaticalized as a means of expressing approximate numbers. Since other types of DP-internal coordination and DP-internal ellipsis are rejected or only marginally accepted in Khanty to this day, the latter assumption seems more likely.

In sum, the Marenjanin text shows the same basic characteristics as the Paasonen tales: it abounds in juxtaposed clauses and avoids ellipsis, and only coordinates NPs if they denote closely related concepts. At the same time, we also attest the occasional use of the conjunction *i/ij* between clauses, and sporadically, between noun phrases, as well. In this respect, Marenjanin’s language use is likely to be ahead of the general course of evolution in Khanty because he became balanced Khanty-Russian bilingual before the majority of Khanties did. As we learn from his autobiography, he worked for Russian merchants since the age of 9 and received several years of Russian schooling as a Komsomol cadre in Soviet times.

3.2 Northern Khanty texts from 1964: spread of conjunctions, emergence of phrasal coordination

In the 3740-word corpus of Khanty texts collected by Rédei during a fieldwork in the northern Kazym area in 1964 (Rédei 1968; OUDB 878, 883, 1022, 1117, 1228), consisting of three fairy tales, as well as an account of the religious beliefs of the Khanty and a brief account of the bear cult, the juxtaposition of clauses without an overt conjunction is still common:

- (28) *βuli sox jeməŋ taxi-ja ixət-t-a, s’ata xaj-t-a.*
 reindeer skin sacred place-LAT hang-PRS-PASS.3sg there leave-PRS-PASS.3SG
 ‘The reindeer hide is hung up at the sacred place, it is left there.’ (OUDB 878)

At the same time, we attest the spreading of a native conjunction, *pa:*, a cognate of the adverb *pə:nə* ‘otherwise’, a few instances of which were pointed out in the Paasonen tales. Since the Rédei corpus of 1964 is similar in size to the Paasonen corpus of 1901, quantitative comparisons can be made. Whereas the texts collected by Paasonen in 1901 display 4 occurrences of *pə:nə*, and a single occurrence of *məβə* functioning as a linking adverb, the Rédei texts contain 56 instances of *pa:*, and 5 instances of *muj*, the northern equivalent of *məβə*, most of which clearly function as conjunctions. Note that the Rédei corpus represents the same register as the Paasonen corpus (fairy tales and descriptions of religious practices), but, nonetheless, there was a marked increase in the use of overt conjunctions in the Rédei corpus. This means that the lack of overt conjunctions in the Paasonen corpus is not attributable to the register (alone) – if that was the case, we would expect to see few overt conjunctions in the Rédei corpus as well.

Pa: appears both between clauses (29a), and in the initial or post-topic position of independent sentences (29b), i.e., it functions as an additive particle with an adjacent or distant (or implicit) first associate. Because of pro-drop, clausal coordination is often indistinguishable from VP-coordination (29c):

- (29) a. *nɔ:-t juβtəsət pa: mə:jpər xɔ:j-t-a.*
 arrow shoot-PRES.3SG and bear hit-PRS-PASS.3SG
 ‘The arrow shoots **and** the bear is hit.’ (OUDB 1022)
- b. *ikil-el it ol-əs. imile-t letot-lat pasan-a*
 man-3SG down lie-PST.3SG woman-3sg food-PL<3SG table-LAT
faβi-s-ətle. luβ pa: iti ol-əs.¹¹
 arrange-PST-PL<3SG she also down lie-PST.3SG
 ‘The man lay down. The woman arranged the food on the table. She **also** lay down.’
 (OUDB 1117)
- c. [*ʃepan [sʲart-ət]*] *pa:* [pro [*lop-ət*, *muj kanʃə-t jij*
 shaman shamanize-PRS.3SG and tell-PRS.3SG what search-PRS.3SG water
βø:rt]]
 spirit
 ‘The shaman shamanizes **and** tells what the water spirit is searching for.’
 (OUDB 878)

Adversative parallel clauses are linked by *a*, borrowed from Russian:

- (30) *je:tn-a ji-t, pasan-ən isiti letoti xaj-t-əm,*
 evening-LAT become-PRS.3SG table-LOC same.way full-of-food leave-PRS-SG<1SG
a: min ant ol-t-əmən la:βət-ti pit-t-əmən.
 but 1DU NEG lie-PRS-1DU wait-INF will-PRS.1DU
 ‘Evening is coming, I leave food on the table again, **but** we won't sleep, we will wait
 awake.’ (OUDB 1117)

We also attest a disjunctive connective, grammaticalized from *muj* ‘what’ (the equivalent of the Eastern Khanty *məβ*), originally an indefinite used as an approximator. A similar pair of sentences in the Paasonen tales, cited under (7a), are still juxtaposed without a connective.

- (31) pro *xuβ man-əs muj βa:n man-əs*
 long go-PST.3SG or short go-PST.3SG
 ‘He went for a long time, **or/perhaps** he went for a short time.’ (OUDB 1117)

¹¹ The 3rd person singular possessive suffixes can cross-reference a possessor, or they can function as markers of definiteness (see Nikolaeva 2003, É. Kiss & Tánčzos 2018).

Whereas phrasal coordination was barely present in the Maremjanin notes, the texts recorded in 1964 already contain different types of constituents conjoined by *pa:*, among them modified NPs (32a) and predicative adjectives (32b).

- (32) a. *luβ sorm-a ji-te-l jupijən s'ar-lat me:t a:j*
 he death-LAT become-PTCP.PRES-3SG after shaman-PL.3SG most small
pox-lat-a pa: met a:j e:βi-lat-a pit-l-ət.
 son-PL.3SG-LAT and most small daughter-PL.3SG-LAT pass-PRS-3PL
 ‘After he dies, his shamanic skills go to his youngest sons **and** youngest daughters.’
 (OUDB 878)
- b. *jə:xət-lat mə:jpər βoxəltə-ti pata βerən-s-aj-ət [βə:na-fək]*
 bow-PL<3SG bear overcome-PTCP.PRES for make-PST-PASS-3PL big-COMP
pa: [ta:ka-fək].
 and fast-COMP
 ‘His bows for shooting a bear were made bigger **and** faster.’ (OUDB 1022)

Pa: still does not occur NP-internally, e.g., between attributive adjectives. If a nominal is modified by two adjectives of the same semantic type, two NPs are projected (33):

- (33) *s'i βe:s'əŋ neŋ-ət, xorasəŋ neŋ-ət pila aŋk-eł a:s'ie-l*
 this pretty woman-3SG beautiful woman-3SG with mother-3SG father-3SG
xos'a joxi man-əs.
 to home go-PST.3SG
 ‘With this pretty woman, this beautiful woman, he went home to his mother and father.’ (OUDB 1117)

The corpus contains several instances of disjunctive phrasal coordination by means of *muj*. In most of these cases, *muj* has already lost its approximator function. *Muj* is sometimes strengthened by *pa:*. If *muj* precedes both disjuncts as in (34b), only the one between the two conjuncts is combined with *pa:*.

- (34) a. *əlmənti ki s'i βuti is-ət [semsajot-ət-a] muj pa: [tə:rəm-a]*
 as if that reindeer soul-3SG spirit-PL-LAT or else god-LAT
man-ət.
 go-PRS.3SG
 ‘supposedly that reindeer's soul goes to the spirits **or else** to god.’ (OUDB 878)
- b. *liβ pro fəpan-lat-ən muj [ta:jəm ixət-man] muj pa:*
 they them shamanhood-3PL-LOC either axe hang-CVB or else
[pe:nis'ar-ən seŋk-man] ofa βerənt-l-ətlat.
 shaman.drum-LOC beat-CVB learn-PRS-PL<3PL
 ‘They get to know these through their shamanhood, **either** by hanging an axe **or else** by beating a shaman drum.’ (OUDB 878)

We also attest structures with non-adjacent disjuncts, which are likely to result from conjunction reduction. Thus, the most likely derivation for (35) is stripping, as indicated (unless the constituent preceded by *muj* has been extraposed or is an afterthought).

- (35) *pro joxt-əm jə:x-lat-ən [jiŋk-ən tɛm-l-aj-ət muj pa:*
 come-PTCP.PST people-PL-LOC water-LOC pour-PRS-PASS-3PL or else
[tə:s'-ən tɛm-l-aj-ət].
 snow-LOC pour-PRS-PASS-3PL
 ‘They are splattered by the coming people with water **or else** with snow.’ (OUDB 1022)

3.3 Eastern Khanty texts from the 1990s: generalization of conjunctions, spreading of phrasal coordination

Our Eastern Khanty corpus from the 1990s, collected by Márta Csepregi (Csepregi 1998; 2002, annotated as OUIDB 728, 730, 732-735, 737, 1076, 1081, 1083-1087), consists of 11 fairy tales and three brief stories from the everyday life of the Khanty. It demonstrates that the spread of conjunctions attested in the Rédei-corpus from 1964 continued into the 1990s. Whereas the 3700-word Rédei corpus from 1964 contains 56 occurrences of *pa:*, the 4200-word Csepregi corpus contains 126 instances of *pə:n(ə)*. Furthermore, it also contains 58 instances of the additive particle *o:s*, which also developed into a conjunction.

As before, *pə:nə* and *o:s* do not always appear between two conjuncts; their first conjunct can also be distant, or implicit. In (40a), *pə:nə* conjoins three clauses; in (40b), it adds a sentence to the preceding independent sentence. (41) contains both types of *o:s*: a conjunction linking clauses, and an additive particle with a distant or implicit antecedent:

- (40) a. [əv'ə su:βəm-ət məj-i] *pə:nə* [pərt pu:l-ət
again reel-of-thread-INS give-PST.PASS.3SG and wood piece-INS
məj-i], *pə:nə* [mən].
give-PST.PASS.3SG and go.PST.3SG
'Again he was provided with a reel of thread **and** he was provided with a piece of wood, **and** he set off.' (OUIDB 736)
- b. *vu: ke:n'ər qo:-pərili n'əβmił: "mə: nny-ət ələ*
that poor man-like_thing say.PST.3SG I you-ACC just
n'u:łtiptə-l-əm." *pə:n* pro pro *ələ vi n'u:łtiptə-təy.*
make.swear-PRS-1SG and he him just so make.swear-PST.SG<3SG
'That poor devil says: "I make you swear." **And** he made him swear.' (OUIDB 737)
- (41) *vu: βe:ti-t o:s, βe:ti-t u:lə ki:t-lə-t, o:s jaqə*
those reindeer-PL also reindeer-PL down catch-PRS-PASS.3PL and home
βəjə-t, qv:t ləypinə u:l lətəlt-ət.
take-PST.PASS.3PL house inside down melt-PST.PASS.3PL
'Those reindeer, **too**, the reindeer are caught, **and** were taken home, they were melted off in the house.' (OUIDB 730)

As is also clear from (41), where the last clause is attached to the previous one without any connective element, the use of conjunctions still isn't general; asyndetic coordination of clauses is still common. To illustrate, (42) describes three subsequent events: the first one is expressed by a nonfinite verbal projection, but the second and third clauses are simply juxtaposed:

- (42) *vi su:łtə-m-əm-ə ulək n'u:r mə:-n kət'əy-nət əβətəm-i*
so slip-PTCP.PST-1SG-LAT harness tether I-LOC knife-COM cut-PST.PASS.3SG
*mə:n-ə βe:ləy qoβit u:tnəm qu:yl-əm.*¹²
I-LOC driving-pole along up.to.bank climb-PST.1SG
'Upon my having slipped, the harness tether was cut with a knife by me, I climbed along the driving pole up to the bank.' (OUIDB 730)

The use of *pə:nə* is still much more general between clauses than between phrases – apart from VPs that can also be interpreted as conjoined clauses with a pro subject (43).

¹² The first locative pronoun encodes the agent of the passive verb; the second one is a locative subject, which often occurs with active verbs to mark subjects functioning as shifted topics (Sosa 2017).

- (43) *jv: tu: aβət-v tət-am, tjaqə pɐ:n jaqə sv.γəltə-m.*
 well that sledge-LAT sit_down-PST.1SG well and home gallop-PST.1SG
 ‘Well, I sat down on the sledge and galloped home.’ (OUDB 730)

Sporadically, NPs conjoined by *pɐ:nə* are also attested. The use of *pɐ:nə* is not restricted to binary combinations of nominals, unlike co-compounding:

- (44) *əj mətə tɐ:t-nə [tu:fəŋ v:γən] pɐ:nə [βɐ:tʲ kʰr βejγən] pɐ:n*
 one some time-LOC Bearded_Chin and Two_Thin_Legs and
 [*o:γ tɔləŋkən*] *βət-t-ət.*
 Two_Temples live-PRS-3PL
 ‘There once lived Bearded Chin and Two Thin Legs and Two Temples.’ (OUDB 1346)

Remarkably, *pɐ:nə* is by now an alternative to the dual suffix. One of the best-known Khanty tales has the title *pi:tʲiŋγəli-γən o:pisv:-γən* ‘little.bird-DU older.sister-DU’ in the 1901 version and in one of the versions from the 1990s; however, in a version recorded in 1993, it is entitled *pi:təŋkəli pɐ:nə o:pi*, i.e., the dual suffixes of the noun phrases have been replaced by the conjunction *pɐ:nə*. The disappearance of the dual indicates that the co-compound projection, with a single NumP subsuming the juxtaposed NPs, has been replaced by an Indo-European-type coordination construction. However, this process has barely begun; the general way of coordinating NPs is still co-compounding, e.g.:

- (45) a. *mʉβ li:tot-ət qu:t-ət əntə tɐ:pət-t-o?*
 what food-INS fish-INS not feed-PRS-PASS.2SG
 ‘Aren't you fed with food [and] fish?’ (OUDB 737)
- b. *tʲi i:ki tɐ:s-ət βay-ət jaqə i:t-ət.*
 this old_man wealth-PL money-PL home take-PST.3PL
 ‘They took home this old man's riches [and] money.’ (OUDB 734)

Disjunctive coordination is fairly frequent at the phrase level. The most common disjunctive coordinator is *mʉβ* (in addition to disjunctive meaning, it still preserves its adverbial meaning that indicates uncertainty or approximate quantity).

- (46) *mɐ: paqqə βət-m-əm-nə [je:ŋ ʉrəkkə qo:-ləm v:t-nə] mʉβ*
 I little.boy-TRNS live-PTCP.PST-1SG-LOC thirteen year-LOC or
 [*je:ŋ ʉrəkkən'ətə v:t-nə*] *βət-m-əm-v*
 fourteen year-LOC live-PTCP.PST-1SG-LAT
 ‘Me being a little boy, thirteen years old **or** fourteen years old, ...’ (OUDB 730)

Parallel coordinate constructions with identical sub-constituents still display no conjunction reduction in many cases. In (40a) above, the first and second clauses have parallel structures with identical verbs ([*he was provided with a reel of thread*] and [*he was provided with a piece of wood*]), but both instances of the verb are spelled out. In (47), the identical subjects of the three clauses are spelled out in each clause:

- (47) *sv:pət tɔβ u:tə kəɾəγ-m-ət tɐ:t-nə sv:pət tɔβ-ət tət rək-kən*
 neck bone off fall-PTCP.PST-3SG time-LOC neck bone-3SG here fly-PST.3DU
sv:pət tɔβ o:s noq tɑqqən-təγ.
 neck bone also up sit_back-PST.3DU
 ‘When the neck bone [cut into two] fell off, his neck bone flew up, and the neck bone sat back to its place.’ (OUDB 737)

Gapping, nevertheless, is attested:

- (48) *pə: v:ntəp jəmsi qv:t pələk-ə qatəltə-l-ləl, pə: v:ntəp pəyi*
 some cradle right house side-LAT carry-PRES-PL<3PL other cradle left
qv:t pələk-ə.
 house side-LAT
 ‘They carried some of the cradles to the right side of the house, the other cradles to the left side of the house.’ (OUDB 735)

In (46), the disjunctive particle conjoins pairs of locative-marked and comitative-marked noun phrases (PPs) that have identical sub-constituents but display no ellipsis ([*in thirteen years*] or [*in fourteen years*]). Apparently, the spreading of conjunctive/disjunctive particles and the use of ellipsis in parallel constructions proceed top down in the language; they can target clauses, clausal adjuncts, and the oblique complements of the verb, but they still cannot target the sub-constituents of noun phrases.

3.4 Prosodic facts

With the appearance of high-quality recordings of Khanty speech towards the end of the 20th century, some generalizations may be made about the prosodic properties of coordinated constituents and co-compounds.

Overtly coordinated constituents, but not co-compounds, (i) may include pauses between the constituents, (ii) may have a pitch reset between the constituents, and (iii) may not be prosodically parallel. The sub-constituents in co-compounds, in contrast, have a tighter prosodic connection: no pauses or pitch reset. Each of the sub-constituents carries the same pitch contour, often downstepped on the second constituent (i.e., with lower absolute F0 values).

The typical prosodic realization of two coordinated nominals is illustrated in (49) and Figure 1: there is a pause and a pitch reset between the constituents, and they are not prosodically parallel to each other.

- (49) (*əj mətə lə:t-nə*) *tu:ʃəŋ v:γən pə:nə βv:t kər βejyən pə:n o:y toləŋkən*
 one some time-LOC Bearded_Chin and Two_Thin_Legs and Two_Temples
βal-l-ət.
 live-PRS-3PL
 ‘There once lived Bearded Chin and Two Thin Legs and Two Temples.’
 (OUDB, 1346)

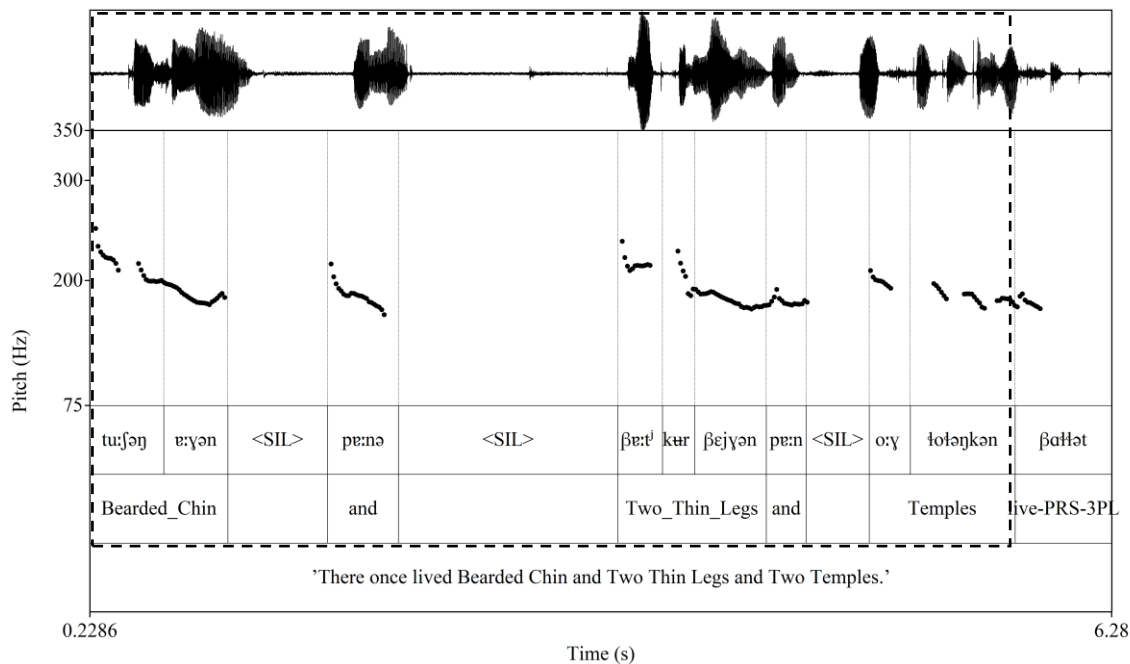


Figure 1. Realization of overtly coordinated nominals in (52) (boxed)

In turn, co-compounds include no pitch reset or pauses, and carry the same pitch contour on both constituents, downstepped on the second one:

- (50) [DP *βoqu-ən tʃe:βər-γən*] *ləβ jot-əl jə-s-γən*.
 fox-DU hare-DU 3SG with-3SG come-PST-3DU
 ‘The fox and the hare came with him.’ (OUIDB, 1315: 129)

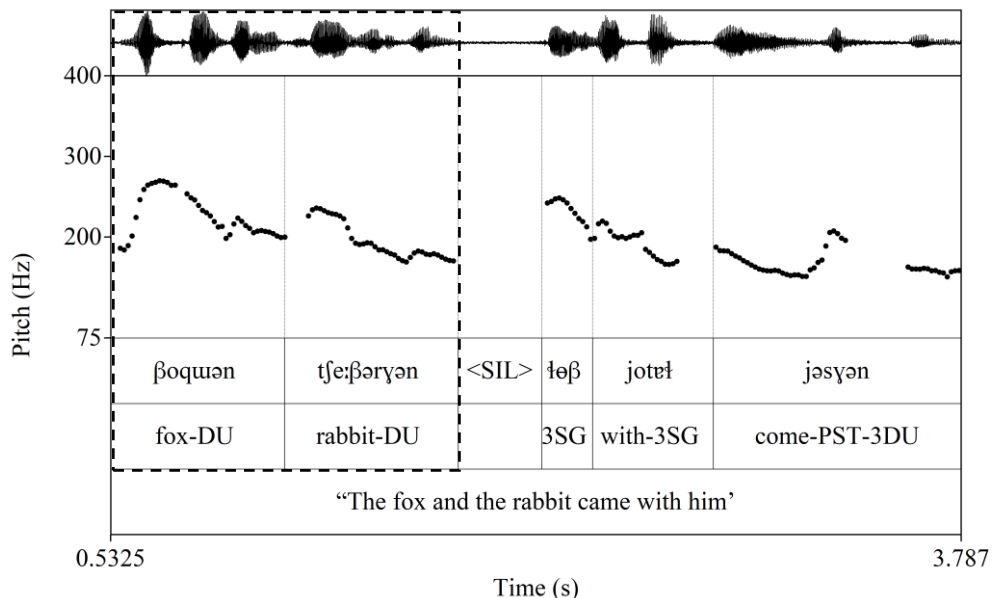


Figure 2. Realization of a nominal co-compound in (53) (boxed)

In sum, the prosodic facts lend additional support to the intuition that the sub-constituents of a co-compound or juxtaposed parallel clauses/verbal projections are more tightly connected to each other than overtly coordinated constituents. In the absence of a coordinator, prosody alone

is relied upon to signal that two adjacent constituents form a higher-order one. In contrast, the presence of a coordinating head allows for a less rigid prosodic realization.

3.5 Interim summary

In the course of the 20th century, the Khanty language has been subject to the increasing dominance of Russian. Parallel with the advancement of Khanty-Russian bilingualism, the Khanty texts recorded in 1936, in 1964, and in the 1990's show the gradual emergence of clausal coordination by means of conjunctive and disjunctive morphemes. Whereas in the document from 1936, we only find sporadic occurrences of the conjunctions *i* and *a* borrowed from Russian, the texts recorded in 1964 already show the recurring use of conjunctive and disjunctive morphemes grammaticalized from native words. In the texts from the 1990s, their use is already systematic, albeit the mere juxtaposition of coordinate clauses is also common.

In the document from 1936, conjunction reduction and phrasal coordination are practically absent except that morphologically parallel NPs representing closely associated concepts can be combined in co-compounds. In the later texts, we attest a growing number of coordinated VPs and coordinated clausal adjuncts, which tend to involve an overt conjunction. In the most recent texts, the first signs of the attrition of co-compounding (the replacement of double dual marking by a conjunction) have also appeared. Non-adjacent coordinates also occur. In such cases, the second coordinate, preceded by a conjunction, is likely to be the result of conjunction reduction (gapping or stripping). Phrasal coordination and conjunction reduction below the NP level are not attested.

4. Coordinated constructions in 21st century Khanty

In today's Eastern (Surgut) Khanty, overt conjunctions are ubiquitous. In a fashion familiar from many better-studied languages (e.g., English), overt conjunctions are not only possible but also preferred with both clausal and phrasal conjuncts, as shown in 4.1 and 4.2.

Evidence provided in Section 4.3 suggests that phrasal coordination in contemporary Khanty, in different contexts, may result either from clausal coordination accompanied by conjunction reduction, or by genuine phrasal coordination that does not involve ellipsis. The latter is attested with so-called symmetric predicates (Lakoff & Peters 1966), such as *intertwine* or *be alike*; multiple subjects that such predicates appear with cannot result from conjunction reduction. Similarly, non-singular agreement on verbs in clauses with multiple singular subjects suggests that they cannot result from conjunction reduction. Additional evidence against the wide-spread use of conjunction reduction in Khanty comes from the rather limited use of structures that are commonly analyzed as relying on ellipsis, such as gapping and stripping.

At the same time, co-compounding is also productive in today's Khanty. In many contexts it may be used interchangeably with phrasal coordination, though it can only be used with constituents that are semantically related and morphologically parallel; more on this in 4.6.

4.1 Syndetic clausal coordination

In today's Eastern Khanty, there is a strong preference for full clauses to be coordinated by an overt conjunction. Examples without overt conjunctions, according to the speakers, sound incomplete (though not strictly speaking ungrammatical). The choice of conjunction corresponds to the relative order of events described by the conjoined clauses: *p̄v:nə* is used for consecutive events, and *o:s*¹³ for contemporaneous ones. This fits well with the fact that these

¹³ There is considerable variability with respect to vowel pronunciation both between subtypes of the same dialect spoken in different villages and even between individual speakers of Khanty. To facilitate recognition of lexical items, the data in this section is transliterated based on standard Surgut Khanty orthography, with differences between individual speakers overlooked. In order to ensure consistency with the previous sections, words of

conjunctions can also be used as adverbials, *o:s* meaning ‘also’, and *pə:nə* meaning ‘otherwise’, ‘again’, ‘also’, ‘additionally’.

- (51) *I:t i:ttən. Mə:fə nə:j ʉl-ət, ⁽ⁱ⁾o:s/ ⁽ⁱⁱ⁾pə:nə Mi:fə jəŋk*
 now evening Masha fire light-PRS.3SG and Misha water
tu:-l.
 bring-PRS.3SG

(i) ‘It is evening now. Masha is making a fire, and Misha is bringing water.’

(ii) ‘It is evening now. Masha makes a fire, and (then) Misha brings water.’

Similarly, disjunctions between clauses are obligatorily overt – otherwise, the disjunctive reading is not rendered:

- (52) *Tem qatəl-nə v:ŋk-em sv:rt tə:rt-l antə v:ʃ-em qu:l qatʃəm*
 this day-LOC mother-1SG pike fry-PRS.3SG or father-1SG fish hot
ji:nk βar-l.
 water cook-PRS.3SG
 ‘Today, either mother fries pike or father cooks fish soup.’

An initial *pə:nə* in today’s Khanty may also be used in its adverbial meaning, ‘also/otherwise/then’:

- (53) *I:t i:ttən-γə jə-y. Pə:nə Mə:fə nə:j ʉl o:s*
 now evening-TRANS become-PST.3SG and Masha fire light-PST.3SG and
Mi:fə jəŋk tu:β.
 Misha water bring-PST.3SG
 ‘Evening came now. Then Masha made a fire and Misha brought water.’

As Khanty is a pro-drop language, conjoined clauses that share a non-overt subject may result either from clausal coordination with *pro* subjects or VP coordination (with overt objects). Similarly, coordinated verbs that share both a subject and an object may be analyzed either as coordination of verbs or that of larger projections (VP/TP/CP) with non-overt arguments.

There is a strong preference for an overt conjunction in these contexts, as shown in (54). In (54a), non-topical objects, not cross-referenced on the verbs, are repeated in each conjunct. This is consistent with Nikolaeva (1999: 68), who showed that only an object that elicits agreement on the verbs can be shared between them. In (54b), the object agreement suffixes on the verbs license object drop.¹⁴

- (54) a. *Mi:fə qu:l qv:təp-i noq βəj, qu:l-ət v:r-γə pan*
 Misha fish net-ABL up take.PST.3SG fish-PL many-TRA put.PST.3SG
pə:nə qu:l i:lə nʹaqəs.
 and fish away scale.PST.3SG
 ‘Misha took the fish out from the net, sorted it and scaled it.’

particular interest (conjunctions *o:s* and *pə:nə*) are rendered in the form that is standard in the OUIDB, in line with the preceding sections of the paper, as opposed to transliteration based on standard orthography.

¹⁴ In a different set of examples, the generalization about object agreement as necessary for object drop did not hold, speakers judged coordinated verbs without object agreement as felicitous. This option seems to be more marked in conjunction constructions though.

(i) *Mi:fə ne:βi βe:li kv:tət pə:nə βe:l.*
 Misha white reindeer catch.PST.3SG and kill.PST.3SG
 ‘Misha caught and killed a white reindeer.’

- b. *Mi:fə sv:rt qv:təmt-əγ, jaqə tu:β-təγ pɐ:nə ni:k*
 Misha pike catch-PST.3SG home bring-PST.3SG<SG and to_water
mol-təγ.
 boil-PST.3SG<SG
 ‘Misha caught a pike, brought it home and cooked it in a cauldron.’

Similarly to conjunctions, disjunctions in clauses with a shared subject need to be overt. In the context of a disjunction, though, the second iteration of the shared object is optional, even in the absence of agreement. This is the case with different markers of disjunction:¹⁵

- (55) a. *Qv:ntəγ jv:γ qu:t βer-lət antɐ (qu:t) tar-lət.*
 Khanty people fish boil-PRS.3PL or fish fry-PRS.3PL
 ‘The Khanty people boil or fry fish.’
- b. *v:ŋki qatəl jatfə (maβət) ti:tə pətə:n muβə sv:rt mɐ:l-əl muβə*
 mother day middle for food due.to or pike boil-PRS.3SG or
(sv:rt) tar-l.
 pike fry-PRS.3SG
 ‘Mum is either boiling or frying pike for lunch.’

To recap, both overt conjunctions and disjunctions are strongly preferred with coordinate clauses.

4.2 Phrasal coordination

Phrasal coordination in contemporary Khanty is as ubiquitous as the coordination of clauses/VPs. First, consider coordinated nominals. Subjects and direct and indirect objects are used with overt conjunctions and disjunctions (in many contexts, co-compounding is still used as an alternative to coordination; more on this in Section 4.6).

Example (56) shows subjects (agents and patients) coordinated with the help of an overt conjunction; dropping the conjunction is degraded. When two (as opposed to more) nominals are conjoined, the verb carries dual agreement; unlike in co-compounding, the dual morpheme does not appear on either of the conjuncts.

- (56) a. *Mv:fə(*-γən) pɐ:nə Mi:fə(*-γən) i:rək-kən.*
 Masha-DU and Misha-DU sing-PST.3DU
 ‘Masha and Misha sang/used to sing (together or not).’
- b. *Mv:-nə sv:rt pɐ:nə jεβ əj qu:tʰəŋ-v pəsən aβti-jε*
 I-LOC pike and perch one space.near-LAT table top-LAT
pan-iyən.
 put-PASS.PST.3DU
 ‘(A) pike and (a) perch were put beside each other on the table by me.’

In contrast, in the context of disjunction, the verb bears singular agreement; it may also accompany both disjuncts.

- (57) *Mv:fə (v:rəγ-ət) muβə (pə)/βəs Mi:fə v:rəγ-ət.*
 Masha sing-PRS.3SG or Misha sing-PRS.3SG
 ‘Masha or Misha is singing/sings.’

¹⁵ There are multiple disjunctions in today’s Khanty, all of which derive from adverbials: *antɐ* ‘or’, *antəqv (pə)* ‘or, perhaps, alternatively’, *antəqv ... antəqv (pə)* ‘either ... or’, *muβ(ə)* ‘or’, *muβə (pə)* ‘or’, *muβə ... muβə (pə)* ‘either ... or’, *mu:j (pə)* ‘or’, *βəs* ‘or’. Usage is determined by the semantics of the utterance, dialectal and idiolectal preferences. The examples preserve the respective speaker’s choice of a disjunction and also list several options where available.

Direct and indirect objects, too, are readily coordinated with overt conjunctions and disjunctions:

(58) *Mv: sv:rt pɐ:nə/muβə jεβ qv:ləmt-əm.*
 I pike and/or perch catch/get-PST.1SG
 ‘I caught (a) pike and/or (a) perch.’¹⁶

(59) *v:ti Mi:fɐ-yɐ pɐ:nə/muβə Pe:tʃv-yɐ βεt tʃarɐs məj.*
 father Misha-LAT and/or Petja-LAT five thousand give.PST.3SG
 ‘Father gave Misha and/or Petja 5000 rubles.’

With a prominent patient, one accompanied by a demonstrative, agents may be coordinated either in active or passive voice (preferred):

(60) *Te:m sv:rt Sɐ:fɐ-nə muβə Mi:fɐ-nə kɐ:tʃ-i.*
 that pike Sasha-LOC or Misha-LOC catch-PASS.PST.3SG
 ‘That pike was caught by Sasha or Misha.’

Like nominals, adjectives that apply to the same referent are coordinated with an overt conjunction. This is true of attributive adjectives, if they describe different dimensions of the same referent, as in (61a), as well as predicative adjectives, as in (61b). Disjunction works in a parallel way to conjunction, as shown in (62).

(61) a. *Qoβ pɐ:nə norəq ju:γ noβ ojevʃlə-tɐγə ru:pɐk.*
 long and straight wood branch find-INF difficult
 ‘It is difficult to find a long and straight stick.’

b. *I:ttən ti:tot ke:βrəm pɐ:nə ɛpləŋ βv:t.*
 evening meal hot and tasty be.PST.3SG
 ‘The dinner was hot and tasty.’

(62) a. *Mv:nt-əm pə:ytə antɐ βastə qantʃʃə ot mɔ:səl.*
 I-ACC-1SG black or blue draw thing necessary
 ‘I need a blue or a black pencil.’

b. *Qantʃʃə ot pə:ytə βəs βastə βot (mɐ: əntə nom-ləm).*
 draw thing black or blue be.PST.3SG I NEG remember
 ‘The pencil was black or blue (I don’t remember).’

Adverbs that describe different dimension of an action are overtly coordinated, too; the same is true for numerals.

(63) a. *Ke:ʃkɐ ju:γ-v ɐstɐγə pɐ:nə sɐj-ləγ qu:ŋəl.*
 cat tree-LAT quickly and sound-ABESS climb.PST.3SG
 ‘A cat quickly and quietly climbed up a tree.’

b. *Mi:fɐ kɐ:t-yən βəs/antəqɐ (pə) qo:ləm sv:rt tu:β.*
 Misha two-DU or/possibly three pike bring.PST.3SG
 ‘Misha brought two or three pikes.’

Is it still possible to coordinate individual phrases via coordinating full clauses, as was the case in earlier varieties of Khanty? Speakers report that that such constructions sound cumbersome and redundant, and come across as old-fashioned. Moreover, full coordinated clauses/TPs in contemporary language do not have the same meaning as single clauses that contain phrasal coordination. According to speakers, (64) cannot be used interchangeably with

¹⁶ There is no mass vs. count distinction with nouns like those denoting fish, so this utterance may be interpreted as catching one fish of each species or catching an unspecified amount of fish of each species.

a clause *I caught (a) pike and (a) perch*. Instead, it can be uttered to emphasize that catching (a) pike was a more significant event, that there were more pikes than perches caught, or that the speaker’s primary intention was to catch (a) pike. That is, the second clause is interpreted as communicating something that is an addition to the main message of the utterance.

- (64) *Mə: sv:rt qv:ləmt-əm pɛ:nə jεβ qv:ləmt-əm.*
 I pike catch-PST.1SG and perch catch-PST.1SG
 ‘I caught (a) pike and caught (a) perch.’

Similarly, (65) cannot be interpreted in such a way that the adjectives apply to a single entity – that is, this sentence cannot mean that a stick that is both long and straight is hard to find. Instead, it can only mean that long sticks are hard to find, and so are straight sticks; cf. (61a) above.

- (65) *Qoβ ju:y noβ ru:pək ojeɣlə-tɛɣə (pɛ:nə) norəq ju:y noβ ru:pək*
 long wood branch difficult find-INF and straight wood branch difficult
ojeɣlətɛɣə.
 find-INF
 ‘It is difficult to find a long stick, it is difficult to find a straight stick.’

4.3 Conjunction reduction or phrasal coordination?

As the data in the previous section showed, phrasal coordination is possible in today’s Khanty, and overt coordinators in phrasal coordination are ubiquitous. The question that this fact gives rise to is what syntax phrasal coordination has. There are several options, which can be grouped according to whether they take phrasal coordination to result from coordination of two individual phrasal constituents, or a process that takes two full clauses and renders unpronounced certain parts of them, such that the result is a single clause with two phrasal constituents coordinated. The ‘phrasal’ approaches include those that take individual phrases to be coordinated, be it with the help of a dedicated projection, such as &P (Munn 1987; Kayne 1994; Johannessen 1996), or phrasal adjunction (Munn 1992; 1993). In turn, according to clausal-coordination approaches, coordination of two phrases may result from coordination of two full clauses followed by ellipsis of all material other than the conjunct in one of them (both the &P- and adjunction-style options are logically possible here as well) (Gleitman 1965; Wilder 1994; Schwarz 1999). Alternatively, in so-called parallel-structure approaches, the two clauses that the two conjuncts are part of undergo the process of Union, whereby the identical constituents in them (e.g., all other than the conjuncts) are fused and only spelled-out once (Goodall 1987).

The data from today’s Khanty may not necessarily provide definitive support for any one of these approaches. As the data in this section shows, though, phrasal coordination is more readily attested in Khanty, while ellipsis in the context of clausal coordination is considerably more restricted.

A line of argumentation traditionally taken to tease apart phrasal coordination from clausal coordination followed by conjunction reduction is based on agreement facts. If the verb in a clause with conjoined singular subjects takes singular agreement, this is compatible with a conjunction reduction approach, as in (66a). On the contrary, if the verb carries non-singular agreement, these facts support a phrasal coordination approach, as in (66b)

- (66) a. [S \forall :SG] and [S V.SG]
 b. [S and S V.PL/DU]

Both agreement patterns are attested, which suggests that both structural configurations are possible:

- (67) a. *Pu:pi, ofər kərəp ot, βoqu, tʃe:βər pɐ:nə kəmləɣ mɐ: jv:tam*
 bear wolf fox hare and wolverine I with-1SG
janɣ-ɐl.
 go-PRS.3SG.
 ‘The bear, wolf, fox, hare and wolverine go with me.’
- b. *Sv:rt, jeβ pa:nə v:ɣərnə pi:ryi ji:ŋk-a ne:βrəm-ət.*
 pike perch and ide back water-LAT jump-PST.3PL
 ‘A pike, a perch, and an ide jumped back into the water.’

In contrast with (67), other coordination constructions can only be analyzed as resulting from phrasal coordination. First, phrasal coordination in Khanty is possible with so-called collective/symmetrical predicates, such as *be alike*, *get divorced*, or *mix (together)*. Such predicates cannot be used with a semantically single argument, and, accordingly, underlying clausal coordination in their context is ill-formed (Curme 1931; Peters 1966; Lakoff & Peters 1966; Wilder 2019):

- (68) a. *John and Mary are alike.*
 b. **John is alike and Mary is/are alike.*

The availability of such constructions in today’s Khanty, as illustrated in (69), attests to the possibility of phrasal coordination that does not result from ellipsis. Note that dual marking on the verbs in (69a-b) cannot be omitted.

- (69) a. [*Mv:fɐ pɐ:nə Pe:tʃɐ*] *ki:t-ɣə mən*(-ɣən).*
 Masha and Petja two-TRANSLAT go-PST.3DU
 ‘Masha and Petja got divorced.’
- b. [*Mv:fɐ pɐ:nə Pe:tʃɐ*] *əj qorɐsəp*(-ɣən).*
 Masha and Petja one alike-DU
 ‘Masha and Petja are alike.’
- c. *Mv: (əj) v:nəɣ-v [qu:t mɐ:rən pɐ:nə qu:t βoj] nʲu:lɐ rʉβt-əm.*
 I one bowl-LAT fish caviar and fish oil together mix-PST.1SG
 ‘I mixed caviar and fish oil together in a bowl.’

The same argument can be made for some postposition-like lexical items, such as *between*: they also cannot take a singular argument and, accordingly, must rely on coordination of phrases that identify the two arguments of *between*. Such postpositions exist in Khanty and apply to coordinated constituents, as in (70). Note that these postposition-like constructions result from the grammaticalization of a noun with spatial semantics (*kɜ:təp* ‘middle’), and, accordingly, we take them to have the corresponding syntactic structure; more on this in the next section.

- (70) *Pu:ɣət qarə [pɐqɐ:t pɐ:nə jaβən kɜ:təp-nə] βət-ət.*
 village space house and river middle-LOC be-PRS.3SG
 ‘The yard is between the house and the river.’

In a similar vein, a single focus particle, such as *only*, can apply to two coordinated nominals in Khanty. An underlying structure with ellipsis, either with a single or iterated *only*, would also be infelicitous here:

- (71) a. *Only Masha and Katja saw a fox.*
 b. **Only Masha saw a fox and only Katja saw a fox.*
 c. **Only Masha saw a fox and Katja saw a fox.*

- (72) *Top Mɛ:fɛ pɛ:nə Kɛ:tʃe βoqi βu:j-γən.*
 Only Masha and Katja fox see-PST.3DU
 ‘Only Masha and Katja saw a fox.’

To recap, contemporary Khanty allows for coordination of phrasal constituents as well as clausal coordination followed by conjunction reduction. The next sections provide further evidence that the latter is more restricted, as is ellipsis in general. Iteration of overt larger structures, such as clauses or TPs, is no longer a felicitous way to coordinate smaller constituents in Khanty.

4.4 Coordination within nominal phrases

Conjunction reduction within nominal projections is not allowed. Three contexts are considered here: (i) phrases with grammaticalized nouns acting as postpositions, (ii) possessive constructions, and (iii) nouns modified by adjectives. In all of these, ellipsis of the head noun is banned.

First, consider postposition-like phrases with spatial semantics that evolved from grammaticalized nouns such as *space*, *inner space* or *other side*.¹⁷ Given that spatial nouns in these constructions preserve their semantics, carry a locative or lative case-marker, and, with a pronominal possessor, carry a possessive suffix, we take them to be nominal phrases, as opposed to P⁰-heads. The nominal referent of the spatial specification is a modifier of the spatial noun and acts as a possessor.

- (73) a. [DP *sɥymət* [NP *tɔmpi-nə*]]
 birch_tree other_side-LOC
 ‘behind the birch tree’ (Lit.: ‘on the birch tree’s other side’)
- b. [DP *tuβ* [NP *qu:tʃi-t-nə*]]
 he nearby_space-3SG-LOC
 ‘next to him/her’ (Lit.: ‘in her/his nearby space’)

In a coordination, both spatial nouns must be spelled out:

- (74) *Ju:γ-ət* [[DP *qɔ:t* [NP *(*qu:tʃəŋ-nə*)]] *pɛ:nə* [DP *ki:βri* [NP *(*qu:tʃəŋ-nə*)]]]
 tree-PL house nearby_space- LOC and well nearby_space-LOC
oβr-et.
 tall-PL
 ‘The trees in the vicinity of the house and in the vicinity of the well are tall. (=two distinct groups of trees)’

If one of the spatial nouns is omitted, the resulting phrase is interpreted as both nominal referents/modifiers applying to the same spatial noun. Accordingly, these contexts rely on coordination of the modifiers, as opposed to ellipsis of the spatial noun; cf. also (70) above.¹⁸

- (75) *Ju:γ-ət* [DP *qɔ:t pɛ:nə ki:βri* [NP *qu:tʃəŋ-nə*]] *oβr-et.*
 tree-PL house and well nearby_space-LOC tall-PL
 ‘The trees in the vicinity of the house and the well are tall. (=a single group of trees)’
 NOT: ‘The trees in the vicinity of the house and in the vicinity of the well are tall. (=two distinct groups of trees)’

¹⁷ The other type of postpositions, non-spatial, evolved from grammaticalized adverbs (Nikolaeva 1999: 36).

¹⁸ These constructions are currently undergoing further change: younger speakers allow (75) to be interpreted as resulting from ellipsis of one of the spatial nouns.

In a similar way, ellipsis of the head noun is banned in possessive constructions. Possessors in Khanty have no overt marking. If the possessor is a noun, the possessum is not overtly marked either, with the two nouns simply juxtaposed (Csepregi 2017).¹⁹

- (76) [DP *I:βen* [NP *ruut*]]
 Ivan boat
 ‘Ivan’s boat’

Possessive constructions may be coordinated, as in (77a). If one of the possessors is omitted, as in (77b), the only available interpretation is that of a single possessum being shared by the two possessors. Structurally, this reading relies on coordination of the possessors. The other reading, according to which each possessor would have their own possessum, with the first of the possessa elided, is not available. Accordingly, ellipsis of the head nominal is also impossible in possessive constructions.

- (77) a. [DP *I:βen* [NP *ruut*]] *pɛ:nə* [DP *Mɛ:ʃɛ* [NP *ruut*]]
 Ivan boat and Masha boat
 ‘Ivan’s boat and Masha’s boat (=two boats)’
 b. [DP *I:βen* *pɛ:nə* *Mɛ:ʃɛ* [NP *ruut*]]
 Ivan and Masha boat
 ‘Ivan and Masha’s boat (= a single boat that belongs to both)’
 NOT: ‘Ivan’s boat and Masha’s boat (=two boats)’

Finally, consider nominals that have other kinds of modifiers, such as demonstratives or adjectives, and are coordinated, as in (78). Here, too, neither head noun may be omitted, which attests to the impossibility of head noun ellipsis.

- (78) *Mi:ʃɛ* [DP *tɛ:m* *nɛ:βi* [NP **(βe:ti)*]] *mɯβə* [DP *tom* *pəytə* [NP **(βe:ti)*]]
 Misha this white deer or that black deer
βɛł-təy.
 kill-PST.3SG<SG
 ‘Misha killed this black deer or that white one.’

There are, nevertheless, two contexts that allow for coordinated modifiers to apply distributively to two head nouns if only one head noun is present: (i) with a non-singular head noun and (ii) in the context of disjunction. A single head noun with dual marking may be used to indicate that each of the modifiers applies to one of the two head nouns:

- (79) a. [DP *nɯy* *pɛ:nə* *nɯy* [NP *qu:tʲi-n-nə*]].
 you and you nearby_space-2DU-LOC
 ‘next to you_i and you_j’ (felicitous if accompanied by gesturing; lit.: ‘in your_i and your_j nearby space’)
 b. *Mɛ:* [DP *I:βen* *pɛ:nə* *Mɛ:ʃɛ* [NP *ruut*(-yən)*]] *kəntʃ-t-əm.*
 1SG Ivan and Masha boat-DU seek-PRS-1SG
 ‘I am looking for Ivan’s and Masha’s boats (=each has a boat).’
 ‘I am looking for Ivan and Masha’s boats (=two shared boats).’
 c. *Mɛ:* [DP *pəytə* *pɛ:nə* *nɛ:βi* [NP *βe:ti*(-yən)*]] *βətɛ* *ki:tʃəm* *βv:t.*
 1SG black and white deer-DU take-INF desire be.PST.3SG
 ‘I wanted to buy a black and a white deer. (=two distinct deer)’

¹⁹ If the possessor is pronominal, the possessum is inflected for the person of the possessor (Csepregi 2017).

In the context of disjunction, the nominal head carries singular but may be modified by disjoint modifiers (arguably, because there is a single referent: only one of the modifiers may apply at a time):

- (80) a. *Mɛ:ʃɛ βv:njtj-ɛnt-əl* [DP *ʃari* *ante sɛ:βən* [NP *ʔəypi-jɛ*]].
 Masha berry.pick-RES-PRS.3SG bucket or basket inner_space-LAT
 ‘Masha is picking berries into a bucket or into a basket.’
- b. *Mɛ:* [DP *I:βɛn* *ante Mɛ:ʃɛ* [NP *ruut*]] *kəntʃ-t-əm*.
 I Ivan or Masha boat seek-PRS-1SG
 ‘I am looking for Ivan’s or Masha’s boat. (=each has one boat)’
- c. *Mɛ:* [DP *pəytə* *ante ne:βi* [NP *βe:ti*]] *βətɛ ki:tʃəm βv:t*.
 I black or white deer-DU take.INF desire be.PST.3SG
 ‘I wanted to buy a black or a white deer.’

To recap, ellipsis of head nouns within nominal phrases is banned, while coordination of modifier phrases (possessors and adjectives) is allowed. At the same time, it can be shown that the nominal constituents that undergo coordination are full DPs: this is demonstrated by the fact that coordinated nominals cannot share a modifier – instead, the modifier only applies to the first of the coordinated nominals. This is true of adjectives and nominal referents with spatial postpositions:

- (81) *Mɛ:* [DP *ənəl/ təm* [NP *sv:rt*]] *pɛ:nə* [DP *v:γərnə*] *qɛ:t-əm*.
 I big/ this pike and ide catch-PST.1SG
 ‘I caught a big pike and an ide.’/ ‘I caught this pike and (an) ide.’
 (Not: ‘I caught a big pike and a big ide.’/ ‘I caught this pike and this ide.’, ‘I caught these pike and ide.’)
- (82) *Mɛ:* *i:mp-əm* *mɛ:-nt* [DP *qv:t* [NP *i:ʔpi-nə*]] *βəs/antəqɛ (pə)*
 I dog-POSS.1SG I-ACC house front_space-LOC or/conversely
 [DP *(*qv:t*) [NP *tɔmpi-nə*]] *tɛ:γət-əl*.
 house back_space-LOC wait- PRS.3SG
 ‘My dog is waiting for me in front of the house or behind the house.’

In sum, Khanty has developed true phrasal coordination by now but has adopted only very limited use of ellipsis.

4.5 Other kinds of coordination

Certain types of coordinate structures – notably, coordination of unlike constituents and coordination that involves ellipsis, such as stripping and forward and backward gapping are easier to account for under conjunction reduction approaches. However, even these contexts in Khanty do not provide strong evidence for ellipsis, which is needed for conjunction reduction.

First, let us consider evidence that comes from coordination of unlike constituents. Generally, coordination requires for both conjuncts to be semantically and/or syntactically parallel, which has come to be known as the *Law of Coordination of Likes* (LCL) (Williams 1981). Violating this rule commonly results in ungrammaticality:

- (83) a. **The scene [of the movie and that I wrote] was in Chicago* (Chomsky 1957: 36)
 b. **John drank [whisky and on Sunday]*. (Wilder 2019)

At the same time, certain ‘unlikes’ can be felicitously coordinated, as in (88). Accounting for the apparent violation of LCL becomes possible if coordination is analyzed as resulting from coordination of phrases of the same type, followed by conjunction reduction (Beavers & Sag 2004; Chaves 2006; for criticism, cf. Levine 2011):

(84) *John* [[*is a Republican*] and [~~*is proud of it*~~]].

The equivalent of (84) is allowed in Khanty but is not informative, because copulas in present tense are null. Stronger evidence comes from similar constructions with overt semi-copulas, as in (85): both must be overt, suggesting that conjunction reduction is not an option, unlike in English.

(85) *Pe:tʃ* – [[*pioneer-γə* *(*jəy*)] *pɛ:nə* [*tʃu:t-ət-nɛt* *jarəŋ-kə*
 Petja pioneer-TRNS become.PST.3SG and DEM-3SG-with proud-TRNS
 *(*jəy*)]].
 become.PST.3SG
 ‘Petja became a pioneer and became proud of that.’

Next, consider gapping, a kind of ellipsis that happens in coordinated clauses and targets the iterated verb (Ross 1968). The remaining lexical material in the clause that contains the ellipsis site is contrasted with its correlates in the preceding clause; one of the remaining constituents is typically the subject, while the other one may be an object or an adjunct (Johnson 1996; Winkler 2005). (86) illustrates forward gapping, where the ‘gapped’ verb is found in the second conjunct. Many verb-final languages also allow for backward gapping, where the ellipsis site is in the first clause.

(86) *John likes ice-cream, and Mary ~~likes~~ chocolate cake.*

In Khanty, the felicity of forward gapping varies by speaker: some speakers accept it, while others insist that the lack of an overt verb in the second conjunct is most readily interpreted as a null copula, as shown in (87a). Younger speakers accept forward gapping more readily than older ones. An overt coordinator is preferred in these contexts, but it may also be dropped. The choice of *o:s* as a coordinator in these contexts indicates juxtaposition between the two clauses and is similar to the Russian particle/complementizer *a* ‘whereas, while, as for’. Only a non-reduced sentence, as in (87b), is judged as felicitous by all speakers.

(87) a. %*Mi:χɐ sɔ:rt qɐ:təl*, (*o:s*) *ʃɔ:ntu* – *jɛβ*.
 Mikha pike catch.PST.3SG and Shonty perch
 ‘Mikha caught a pike, and Shonty [caught] a perch.’
 ‘Mikha caught a pike, and Shonty [is] a perch.’
 b. *Mi:χɐ sɔ:rt qɐ:təl*, *pɛ:nə ʃɔ:ntu jɛβ qɐ:təl*.
 Mikha pike catch.PST.3SG and Shonty perch catch.PST.3SG
 ‘Mikha caught a pike, and Shonty caught a perch.’

Similar variation is found in forward gapping with intransitive predicates, where the remnants in the second conjunct consist of a subject and an adverb, as opposed to a subject and a direct object, as in (87). Here, the correlation between age and ellipsis ease did not hold: an older speaker offered an utterance with forward gapping, as in (88a), while a younger one preferred a paraphrase without ellipsis, as in (88b).

(88) a. *Sɛ:ʃɐ kənəkkə sɛ:p ɯ:lti ne:βrəm-əy*, *o:s pɛ:tʃ* – *ru:pəkkə*.
 Sasha easily creak across jump-PST.3SG and Petja with_effort
 ‘Sasha easily jumped over the creak and Petja did so with effort.’
 b. *ʃɔ:ntu pɛ:stɛ sɛ:p-əli ɯ:lti ne:βrəm-əy*, *o:s pɛ:tʃ əntə ti:kim*
 Shonty quickly creak-DIM across jump-PST.3SG and Petja NEG that
lɛ:tək βɔ:l.
 dexterous be.PST.3SG
 ‘Sasha quickly jumped over the creak but Petja wasn’t so dexterous.’

Backward gapping/RNR, in contrast, is overall judged as more felicitous, though an overt verb in the first conjunct is still preferred. There is also a preference for an overt conjunction, which in this case may be either *o:s* or *pɛ:nə*.

- (89) a. ?*Mi:fɛ sv:rt, Sɛ:fɛ jɛβ qv:təl.*
 Misha pike, Sasha perch catch.PST.3SG
 ‘Misha caught a pike, and Sasha caught a perch.’
- b. *Mi:fɛ sv:rt qv:təl, (pɛ:nə/ o:s) Sɛ:fɛ jɛβ qv:təl.*
 Misha pike catch.PST.3SG, and Sasha perch catch.PST.3SG
 ‘Misha caught a pike, and Sasha caught a perch.’

Backward gapping is judged as more felicitous with intransitive predicates and adjuncts:

- (90) *ʃɔ:ntu latəqqə, o:s pɛ:tʃv tap sv:p-əli ɯ:lti ne:βrəm-əɣ.*
 Shonty handily and Petja barely creak-DIM across jump-PST.3SG
 ‘Shonty easily jumped over the creak and Petja barely did so.’

The contrast between gapping with objects and adjuncts is likely rooted in their structure: examples like (89a) rely on the ellipsis of the V⁰-head, while examples like (90) involve VP-ellipsis. The former process is marginal in Khanty, while the latter is less restricted. One of the reasons for this distinction may be the fact that ellipsis of the verb results in a string of nominals, as in (89a). Lack of overt case marking may lead to ambiguity and higher processing cost, which is why such structures are avoided.

Finally, consider stripping, or bare argument ellipsis: here, all constituents in the second clause are deleted under identity with the first one, except for a single constituent, which may be accompanied by an adverb (*perhaps, as well, too*) or negation (Ross 1969; Hankamer & Sag 1976):

- (91) a. *John left yesterday, and Mary too.*
 b. *John drank whisky last night, or maybe tequila.*
 c. *Mary likes oysters, but not John.*

Stripping is not considered fully felicitous in Khanty. Stripping with the subject as the remnant (VP ellipsis) sounds somewhat colloquial to some speakers, but hard to interpret to others. The preferred version spells out the second conjunct in full; an overt conjunction is preferred too.

- (92) a. ??*Mi:fɛ sv:rt qv:təl, Sɛ:fɛ ətʃə.*
 Misha pike catch.PST.3SG, Sasha too
 ‘Misha caught a pike, Sasha too.’
- b. *Mi:fɛ sv:rt qv:təl, (pɛ:nə) Sɛ:fɛ ətʃə sv:rt qv:təl.*
 Misha pike catch.PST.3SG, and Sasha too pike catch.PST.3SG
 ‘Misha caught a pike (and) Sasha caught a pike too.’

Stripping with the subject as the remnant in the second clause is equally bad with disjunction:

- (93) **Sɛ:fɛ təm sv:rt qv:təl mʉβə Mi:fɛ.*
 Sasha this pike catch.PST.3SG or Misha
 (‘Sasha or Misha caught this pike.’)

Stripping with an object remnant is even more restricted: here, only two fully spelled out clauses, coordinated with an overt *pɛ:nə* are judged as fully felicitous, while various more reduced options are less so:

- (94) a. **Mi:fɛ sv:rt qv:təl pɛ:nə jɛβ ətʰə.*
 Misha pike catch.PST.3SG and perch too
 ('Misha caught a pike and a perch, too.')
- b. ?*Mi:fɛ sv:rt qv:təl pɛ:nə jɛβ ətʰə qv:təl.*
 Misha pike catch.PST.3SG and perch too catch.PST.3SG
 'Misha caught a pike and a perch, too.'
- c. *Mi:fɛ sv:rt qv:təl pɛ:nə lʉβ jɛβ ətʰə qv:təl.*
 Misha pike catch.PST.3SG and 3SG perch too catch.PST.3SG
 'Misha caught a pike and he caught a perch, too.'

In a parallel fashion, stripping with an object remnant is infelicitous in the context of disjunction.

- (95) **Mi:fɛ sv:rt qv:təl muβə jɛβ.*
 Misha pike catch.PST.3SG or perch
 ('Misha caught a pike or a perch.')

As was the case with gapping, stripping with adjuncts is, in contrast, considered felicitous:

- (96) *Mi:fɛ te:m qatəl qu:t kəntʃtʃv-yə mənə-t muβə qoltvɔit.*
 Misha this day fish hunt-TRANS go-PRS.3SG or tomorrow
 'Misha is going fishing today or tomorrow.'

As was the case with gapping, stripping with adjuncts may be less restricted since it relies on VP-deletion as opposed to ellipsis of the verb.

4.6 Co-compounding

Widespread use of overt conjunctions in phrasal coordination in Khanty did not obliterate the traditional use of co-compounds instead of overt coordination. Nouns (as well as adjectives, numerals, verbs) and larger phrases, such as nouns modified by adjectives, can form co-compounds in Khanty. In terms of meaning, co-compounds in most contexts can be used interchangeably with overtly coordinated constituents:

- (97) a. *Mv:fɛ-yən (*pɛ:nə) Mi:fɛ-yən v:rəy-ləyən.*
 Masha-DU and Misha-DU sing-PRS.3DU
 'Masha and Misha sing (in general).' / 'Masha and Misha are singing now.'
- b. *Mv:fɛ(*-yən) pɛ:nə Mi:fɛ(*-yən) v:rəy-ləyən.*
 Masha-DU and Misha-DU sing-PRS.3DU
 'Masha and Misha sing (in general).' / 'Masha and Misha are singing now.'

Like phrasal coordination, co-compounding can be used with collective predicates:

- (98) a. *Mv:fɛ-yən Pe:tʃv-yən ki:t-yə mən-yən.*
 Masha-DU Petja-DU two-TRANSLAT go-PST.3DU
 'Masha and Petja got divorced.'
- b. *Mv:fɛ-yən Pe:tʃv-yən əj qorvəsəp-yən.*
 Masha-DU Petja-DU one alike-DU
 'Masha and Petja are alike.'

Co-compounding and overt coordination also behave similarly with respect to distributivity: both strategies can have an interpretation whereby the predicate applies to each of the conjuncts either separately or together:

- (99) a. *Mv:fə ke:fkə pɐ:nə v:mp-əli mɛs ji:ŋk-ət ji:n'əlt-əɣ.*
 Masha cat and dog-DIM cow water-INSTR water-PST.3SG
 ‘Masha gave the cat and the little dog milk (together/separately).’
- b. *Mv:fə ke:fkə-ɣən v:mp-əli-ɣən mɛs ji:ŋk-ət ji:n'əlt-əɣ.*
 Masha cat-DU dog-DIM-DU cow water-INSTR water-PST.3SG
 ‘Masha gave the cat and the little dog milk (together/separately).’

The preferred interpretation is determined pragmatically and is consistent between the two structures:

- (100) a. *ʃv:ntu pɐ:nə Pe:tʰa rɔ:pitlətə maβəl βet tʰarɛs-ət məj-at.*
 Shonty and Petja work for five thousand-INSTR give-ASS.PST.3PL
 ‘Shonty and Petja were given 5000 [rubles] for their work (altogether).’
- b. *ʃv:ntu-ɣən Pe:tʰa-ɣən rɔ:pitlətə maβəl βet tʰarɛs-ət məj-at.*
 Shonty-DU Petja-DU work for five thousand-INSTR give-PASS.PST.3PL
 ‘Shonty and Petja were given 5000 [rubles] for their work (altogether).’

There are some differences between co-compounding and overt coordination, though. First, phrasal coordination and co-compounding behave differently with respect to the two types of disjunction, which yield alternative questions or yes/no-questions, respectively (Romero & Han 2003; Han & Romero 2004; Pruitt & Roelofsen 2013). With low disjunction (alternative questions), phrasal coordination is used, while co-compounding is used with high disjunction (yes/no-questions):

- (101) a. *Mi:fə sv:rt muβə jɛβ qv:ləmt-əɣ?* (alternative questions)
 Misha pike or perch catch-PST.3SG
 ‘Did Misha catch a pike or a perch?’
- b. *Mv:fə v:ləŋ kv:pə ante fɛ:j ji:n'tʰ?*
 Masha morning coffee or tea drink.PST.3SG
 ‘Did Masha drink coffee or tea in the morning?’
- (102) a. *Mi:fə sv:rt-moqsəŋ qv:ləmt-əɣ?* (yes/no-questions)
 Misha pike-muksun catch-PST.3SG
 ‘Did Misha catch a pike or a muksun (i.e., did he catch some fish)?’
- b. *Mv:fə v:ləŋ kv:pə-fɛ:j ji:n'tʰ?*
 Masha morning coffee-tea drink.PST.3SG
 ‘Did Masha drink coffee or tea in the morning (i.e., did she drink something)?’

These facts are consistent with the intuition that, more generally, co-compounds cannot be used to express disjunction of the *either... or* type, where only one of the disjuncts can be true at a given time. This is compatible with the structural analysis of co-compounds proposed in Section 2.3: co-compounds share functional projection(s) that combines two element(s), but there is no room for a dedicated head that can express a disjunctive reading. In contrast, a conjunctive reading can be expressed with iconic means, such as the iterated dual suffix and prosodic parallelism, even in the absence of a dedicated head.

4.7 Summary

The data from contemporary Khanty presents a natural continuation of the process that started and gained speed in Khanty during the second half of the 20th century. First, overt coordination became commonplace. In the traditional varieties of Khanty, there were no overt coordinators, and conjunction reduction was not allowed. Gradually, adverbials meaning ‘then’ and ‘also’ came to be used as coordinators, while still retaining their adverbial meaning in other contexts.

The change from no overt coordinators to a strong preference for overt coordinators in most contexts is by now complete. Second, coordination by now also applies to constituents smaller than clauses. We have demonstrated that many of these contexts rely on the coordination of phrasal constituents, as opposed to clausal coordination followed by conjunction reduction, based on some of the agreement facts, the availability of collective predicates, and patterns of coordination within nominal constituents. Third, we have shown that the use of ellipsis in Khanty is quite restricted (marginal within the VP and prohibited within the DP), based on data from gapping, stripping, and VP-ellipsis. Finally, co-compounds are still in use in today's Khanty, and, for the most part, are used interchangeably with overtly coordinated constituents.

5. Accounting for the correlation of overt conjunctions and phrasal coordination

The data surveyed above outline an evolutionary path from a stage of Khanty where it had no conjunctive/disjunctive particles and no phrasal coordination (only co-compounding) to a stage with overt conjunctions and disjunctions, and phrasal coordination widely available. This diachronic process suggests that there is an intrinsic correlation between overt conjunctions and phrasal coordination. The question arises about what motivates this correlation.

The language of the Khanty texts recorded in 1901, displaying no conjunction reduction in partially identical parallel clauses and phrases, involves plenty of repeated material. These texts are folklore texts, which might suggest that repetition is a genre-specific, rhetorical means; however, it is also present in Marenjanin's text from 1936, which represent an informal but not folklore-like register. The prevalence of this much repetitiveness may seem striking and uneconomical. However, general principles of economy in grammar (e.g., Haiman 1983, 1985) are so fundamental that repetitiveness of this kind is expected to be licensed only if it pays off elsewhere. We argue that this is indeed the case: in the language type represented by traditional Khanty, the lack of the ellipsis of repeated material in parallel structures, or, more generally, the lack of phrasal coordination facilitates processing.

In Khanty, an SOV language with un-casemarked subjects and objects, with unmarked possessors and possessa, and with both subject and object pro-drop, an [NP₁ NP₂ V] string can, in principle, be analyzed in multiple ways:

- (103) [NP₁ NP₂ V]
- (i) NP₁ = subject, NP₂ = object;
 - (ii) NP₁ = possessor, NP₂ = subject;
 - (iii) NP₁ = possessor, NP₂ = object (subject = *pro*).

The grammatical functions of the NPs can be disambiguated based on the selectional properties of the verb, the agreement morphemes on it, and the context – that is to say, their disambiguation is not complete until the verb has been processed. As traditional Khanty only had asyndetic coordination, the replacement of NP₁ or NP₂ or both in the [NP₁ NP₂ V] sentence by an (asyndetically) coordinated expression would have extended the string of juxtaposed un-casemarked NPs, thereby multiplying the interpretive options:

- (103') [NP₁ NP₂ V] (with asyndetic phrasal coordination allowed)
- (iv) NP₁, NP₂ = subjects;
 - (v) NP₁, NP₂ = objects (subject = *pro*).

In the case of [NP₁ NP₂ NP₃V], the possibilities would multiply, resulting in garden-path situations, i.e., initial misinterpretations necessitating the reanalysis of the string.

Therefore, we propose that, asyndetic phrasal coordination in earliest attested Khanty must have been blocked, in order to ensure processing efficiency, as required by Hawkins' (2004) principle of Maximize On-line Processing, discussed in Section 1.3. According to this principle,

the language processor aims to maximize the set of properties that are assignable to item X as soon as X is processed; the optimality of a given construction is determined as “a function of properties that are unassigned or misassigned to X in a structure/sequence S, compared to the number in an alternative” (Hawkins 2004: 51).

In traditional Khanty, the structural relations/ grammatical functions of the NPs in an [NP₁ NP₂ V] sequence cannot be unambiguously assigned until the listener has parsed the clause-final verb and its agreement morphology. Let us illustrate that with a pseudo-Khanty SOV sentence *Masha children watch+AGR*. If the agreement suffix cross-references a singular subject, *Masha* acts as the subject, and *children* as the object. Alternatively, if the context suggests that the subject/topic is identical with that of the preceding sentence and is represented by a *pro*, then *Masha* is to be interpreted as the possessor of the object, *children*. If the agreement cross-references a plural subject, *children* is the subject and *Masha* is the possessor of the subject.

Now, in the hypothetical version of Khanty that has asyndetic phrasal coordination, the number of NPs with temporarily unassigned or misassigned properties would increase. The two NPs of the string *Masha children watch+AGR* could also represent either the subject, or the object (assuming a *pro* subject). If the string included three or four caseless NPs, the possibilities and the potential misalignments would multiply. The delayed assignments and the rounds of property misassignment + reassignment, in Hawkins’ terms, would impose an excessive load on working memory.

The lack of phrasal coordination, therefore, reduces the chance of garden-path situations. The price to pay for it is some repetitiveness – which actually may not be high. Structural parallelism has been shown to facilitate both comprehension and production; see Frazier et al. (2000) and the experimental studies cited therein. The advantage of parallel forms is assumed to be due to the reuse of templates, and/or shortcuts in the mapping of form and meaning. Lexical parallelisms are likely to further increase these effects.

Coordination can be unavoidable – e.g., in the case of two agents jointly performing a collective action such as fighting or shaking hands. In such cases, traditional Khanty used co-compounding – i.e., it unified noun phrases of the same grammatical function in a single extended nominal projection, marking their unity by parallel morphology.

In the 20th century, the emergence of conjunctions has licensed the use of coordinated maximal projections, including coordinated noun phrases. Because a conjunction linking two constituents indicates that the two NPs share the same grammatical function and the same mother nodes, it thereby eliminates several unassignment and misassignment possibilities. In another pseudo-Khanty sentence *Masha and children watch+AGR*, *Masha* cannot be a possessor, due to the conjunction immediately following it. The possibility of *Masha* being the subject and *children* being the object is also excluded for the same reason. The coordinated nouns are therefore parsed as the subject of the clause – unless the context provides a *pro* subject, in which case the coordinated nouns are interpreted as the object.

The diachronic process observed in Khanty has another theoretical implication. The use of conjunctions is spreading top-down: they first appeared between clauses, then between VPs, between subjects and between adjuncts. The syndetic coordination of VP-internal objects and that of NP-internal constituents are even more recent developments. NP-internal ellipsis is still rejected. A similar top-down spread of change has also been observed in languages undergoing a change of the head–complement order. Hungarian, a descendant of SOV Proto-Ugric, for example, has developed a head-initial VP and head-initial functional projections (CP, TopP, FocP, NegP, and DP), but the NP and PP projections still preserve the head-final structure.

The top-down direction of word order change has been related to the Final-Over-Final Condition (FOFC), a syntactic constraint disallowing structures where a head-initial phrase is contained in a head-final phrase in the same extended projection (Biberauer 2018). It is,

however, unclear how FOFC could be applied to the evolution of syndetic coordination. In the widely accepted theory of coordination, &P is head-initial, but it is an adjunct (Munn 1992; 1993). Therefore, &P is not part of the extended projections that constitute the spine of clause structure. Furthermore, the Khanty clause is still mostly head-final, hence, even if &P was the extension of a lexical projection, FOFC would only allow it above the clause level. It seems, therefore, that the direction of diachronic changes that affect several levels of syntactic structure is controlled by a more general principle. A natural source for a more general principle would be a processing requirement, as in the account developed here.

6. Conclusion

In this paper, we summarized the stages of the development of overt conjunctions and phrasal coordination in Khanty and provided a processing account of the correlation between the availability of conjunctions and coordination of phrasal constituents. We showed, based on corpus data, that the earliest attested Khanty used clausal juxtaposition for both clausal and phrasal coordination, and did not have overt conjunctions, phrasal coordination, or conjunction reduction. As Russian-Khanty bilingualism became more widespread over the 20th century, we find successively more overt conjunctions/disjunctions, and more evidence for phrasal coordination, in our corpora of Khanty texts from 1936, 1964, and the 1990s. Drawing upon elicitation data, we demonstrated that in contemporary Khanty overt conjunctions/disjunctions are ubiquitous, and so is phrasal coordination. Based on a variety of syntactic tests, we showed that what looks like coordination of phrasal constituents in contemporary Khanty indeed is derived via coordination of phrase-size constituents, as opposed to clausal coordination, followed by conjunction reduction. The conclusion that this allowed us to make is that there is a correlation between the emergence of conjunctions/disjunctions and the availability of phrasal coordination in Khanty. We showed that the restriction on phrasal coordination in the absence of dedicated coordinating particles follows from the Maximize On-line Processing principle, which restricts the number of morphosyntactically unmarked constituents that can be strung together. Once conjunctions/disjunctions grammaticalized (likely, due to contact with Russian), a way of disambiguating a string of constituents emerged which opened the door for the emergence of phrasal coordination.

References

- Beavers, John & Ivan A. Sag. 2004. Coordinate ellipsis and apparent non-constituent coordination. In *The Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar*, 48–69.
- Biberauer, Theresa 2018. Probing the nature of the final-over-final condition: The perspective from adpositions. In Laura R. Bailey & Michelle Sheehan (eds.), *Order and structure in syntax I: Word order and syntactic structure*, 177–216. Berlin: Language Science Press.
- Bogoras, Waldemar 1922. Chukchee. In *Handbook of American Indian Languages*, Part 2. *Bureau of American Ethnology Bulletin* 40: 631–903.
- Chafe, Wallace. 1985. Linguistic differences produced by differences between speaking and writing. In David R. Olson, Nancy Torrance, and Angela Hildyard (eds.) *Literacy, Language, and Learning*, 105–123. Cambridge: Cambridge University Press.
- Chafe, Wallace. 1987. Cognitive constraints on information flow. In Russell Tomlin (ed.) *Coherence and Grounding in Discourse*, 21–51. Amsterdam: John Benjamins.
- Chaves, Rui P. 2006. Coordination of unlikes without unlike categories. In *Proceedings of the 13th International Conference on Head-driven Phrase Structure Grammar*, 102–122. CSLI Publications Stanford, CA.
- Chomsky, Noam. 2013. Problems of projection. *Lingua* 130: 33–49.
- Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.
- Cole, Peter 1982. *Imbabura Quechua*. *Lingua Descriptive Studies*. Amsterdam: North Holland.
- Craig, Colette Grinevald 1977. *The Structure of Jacalteco*. Austin: University of Texas Press.
- Csepregi, Márta. 1998. Szurguti osztják chrestomatia. Szeged: JATE.
- Csepregi, Márta. 2002. Texte in chantischer Sprache vom Fluss Agan. In: E. Helimski – A. Widmer (eds.): *Sei gegrüsst! Beiträge zur Finnougristik zu Ehren von Gert Sauer dargebracht zu seinem siebzigsten Geburtstag. Veröffentlichungen der Societas Uralo-Altaica* 57, 85–93. Wiesbaden: Harrassowitz Verlag.
- Csepregi, Márta. 2017. *Surgutskij dialekt khantyjskogo jazyka [The Surgut dialect of Khanty]*. Khanty-Mansijsk: Department of Education and Youth Policy of KhMAO; The Ob-Ugric Institute of Applied Research and Development.
- Curme, G. 1931. *Syntax*. Boston, Mass: Heath.
- DEWOS = Steinitz, Wolfgang (ed.) 1966–1993.
- 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.
- Drellishak, Scott 2005. Coordination and processing. *University of Washington Working Papers in Linguistics* 24.
- É. Kiss Katalin, Orsolya Tánzos 2018. From possessor agreement to object marking in the evolution of the Udmurt -jez suffix: A grammaticalization approach to morpheme syncretism. *Language* 94 :733-757.
- Frazier, Lynn, Alan Munn, Charles Clifton, Jr. 2000. Processing coordinate structures. *Journal of Psycholinguistic Research* 29(4): 343–370.
- Gleitman, Lila R. 1965. Coordinating conjunctions in English. *Language* 41(2). 260–293.
- Goodall, Grant. 1987. *Parallel Structures in Syntax*. Cambridge: Cambridge University Press.
- Haiman, John. 1983. Iconic and economic motivation. *Language* 59. 781–819.
- Haiman, John. 1985. *Natural syntax. Iconicity and erosion* (Cambridge Studies in Linguistics). Vol. 44.
- 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.

- Han, Chung-hye & Maribel Romero. 2004. Disjunction, focus, and scope. *Linguistic Inquiry*. MIT Press 35(2). 179–217.
- Hankamer, Jorge & Ivan Sag. 1976. Deep and surface anaphora. *Linguistic inquiry* 7(3). 391–428.
- Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
- Heycock, Caroline & Roberto Zamparelli. 2005. Friends and colleagues: Plurality, coordination, and the structure of DP. *Natural language semantics*. Springer 13(3). 201–270.
- Johannessen, Janne Bondi. 1996. Partial agreement and coordination. *Linguistic Inquiry* 27. 661–676.
- Johnson, Kyle. 1996. In search of the middle field. University of Massachusetts, Amherst, ms. http://people.umass.edu/kbj/homepage/index_johnson.htm.
- Kálmán, Béla. 1976. *Chrestomathia Vogulica*. Budapest: Tankönyvkiadó.
- Kayne, Richard S. 1994. *The antisymmetry of syntax*. Vol. 25. Cambridge MA: MIT Press.
- Lakoff, George & Stanley Peters. 1966. Phrasal conjunction and symmetric predicates.
- Levine, Robert. 2011. Linearization and its discontents. In *The Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar*, 126–146. Citeseer.
- Lewy, Ernst 1911. *Zur finno-ugrischen Wort- und Satzverbindung*. Vandenhoeck, Göttingen.
- Mikola, Tibor. 1973–74. Ugor ikerszavak. *Néprajz és Nyelvtudomány* 17-18: 54–62.
- Mikola, Tibor. 1985. Ikerszók és parallelizmusok az obi-ugor nyelvekben. *Néprajz és Nyelvtudomány* 29-30: 143–149.
- Mithun, Marianne. 1988. The grammaticalization of coordination. In: John Haiman and Sandra A. Thompson (eds.) *Clause Combining in Grammar and Discourse*, 331–359. Amsterdam: John Benjamins.
- Munn, Alan Boag. 1987. Coordinate structure, and X-bar theory. *McGill Working Papers in Linguistics* 4(1). 121–140.
- Munn, Alan. 1992. A null operator analysis of ATB gaps. *The Linguistic Review*. Walter de Gruyter, Berlin/New York 9(1). 1–26.
- Munn, Alan Boag. 1993. *Topics in the syntax and semantics of coordinate structures*. University of Maryland at College Park PhD Thesis.
- Nikolaeva, Irina. 1999. *Ostyak*. München: Lincom Europa.
- Nikolaeva, Irina. 2003. Possessive affixes in the pragmatic structuring of the utterance: evidence from Uralic. In: Pirkko Suihkonen, and Bernard Comrie (eds.) *International symposium on deictic systems and quantification in languages spoken in Europe and North and Central Asia. Collection of papers*, 130–145. Izhevsk and Leipzig: Udmurt State University and Max Planck Institute for Evolutionary Anthropology.
- Ohuri, Toshio. 2004. Coordination in Mentalese. In: Martin Haspelmath (ed.) *Coordinating Constructions*, 41–66. Amsterdam/Philadelphia: John Benjamins.
- OUIDB EM Ob-Ugric Database Eastern Mansi Corpus http://www.babel.gwi.uni-muenchen.de/index.php?abfrage=EM_corpus&subnavi=corpus_pub
- Peters, Stanley. 1966. *Coordinate Conjunction in English*. Massachusetts Institute of Technology Doctoral dissertation.
- Pruitt, Kathryn & Floris Roelofsen. 2013. The interpretation of prosody in disjunctive questions. *Linguistic inquiry*. MIT Press 44(4). 632–650.
- Ravila, Paavo. 1986. Über die Verwendung der Numeruszeichen in den uralischen Sprachen. *Finnisch-Ugrische Forschungen* 27: 1-136.
- Rédei, Károly. 1968. *Nord-ostjakische Texte (Kazym-Dialekt) mit Skizze der Grammatik*. Göttingen: Vandenhoeck & Ruprecht.

- Romero, Maribel & Chung-hye Han. 2003. Focus, ellipsis and the semantics of alternative questions. *Empirical issues in formal syntax and semantics* 4. 291–307.
- Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT.
- Ross, John Robert. 1968. *Constraints on Variables in Syntax*. Massachusetts Institute of Technology Doctoral dissertation.
- Ross, John Robert. 1969. Guess who? In *Proceedings of the fifth regional meeting of the Chicago Linguistic Society*. Chicago: University of Chicago Press.
- Ross, John Robert. 1970. Gapping and the order of constituents. In M. Bierwisch, and K.E. Heidolph (eds.), *Progress in Linguistics*. The Hague: Mouton.
- Schwarz, B. 1999. On the syntax of either... or. *Natural Language and Linguistic Theory* 17(2). 339–370.
- Sipos, Mária. 2015. Időhatározói összetett mondatok serkáli hanti szövegekben *Nyelvtudományi Közlemények* 111: 131–149.
- Sosa, Sachiko. 2017. Functions of morphosyntactic alternations, and information flow in Surgut Khanty discourse. PhD dissertation, University of Helsinki.
- Stassen, Leon. 2003 Noun phrase conjunction: The coordinative and the comitative strategy, In: Frans Plank (ed.) *Noun Phrase Structure in the Languages of Europe*, 761–819. Berlin: Mouton de Gruyter.
- Steinitz, Wolfgang (ed.). 1966–1993. *Dialektologisches und etymologisches Wörterbuch der ostjakischen Sprache*. Berlin: Akademie-Verlag.
- Steinitz, Wolfgang. 1941. *Ostjakische Volksdichtung und Erzählungen aus zwei Dialekten II*. Stockholm.
- Suárez, Jorge A. 1983. *The Mesoamerican Indian Languages*. Cambridge Language Surveys. Cambridge: Cambridge University Press.
- Szabolcsi, Anna. 1990. Osztyák parallelizmusok és mellérendelő összetételek. *Nyelvtudományi Közlemények* 91: 221–225.
- Terjoshkin, N. I. 1959. *Bukvar*. Leningrad: Gosudarstvennoje Uchebno-Pedagogicheskoye Izdatelstvo.
- Terrill, Angela. 2004. Coordination in Lavukaleve. In: Martin Haspelmath (ed.) *Coordinating Constructions*, 427–444. Amsterdam/Philadelphia: John Benjamins.
- Vértes, Edit. 2001. H. Paasonens surgutostjakische Textsammlungen am Jugan. Neu transkribiert, bearbeitet, übersetzt und mit Kommentaren versehen von Edith Vértes. *Mémoires de la Société Finno-Ougrienne* 240.
- Wälchli, Bernhard. 2005. *Co-Compounds and Natural Coordination*. Oxford: Oxford University Press.
- Welmers, William E. 1973. *African Language Structures*. Berkeley: University of California Press.
- Wilder, Christopher. 1994. Coordination, ATB and ellipsis. In C. Jan-Wouter Zwart (ed.), *Minimalism and Kayne's Antisymmetry Hypothesis* (Groningen Arbeiten Zur Germanistischen Linguistik 37), 291–331.
- Wilder, Christopher. 2018. Conjunction Reduction and Right-Node Raising. In: Jeroen van Craenenbroeck and Tanja Temmerman (eds.) *The Oxford Handbook of Ellipsis*. DOI: 10.1093/oxfordhb/9780198712398.013.2
- Williams, Edwin S. 1981. Transformationless grammar. *Linguistic Inquiry*. JSTOR 645–653.
- Winkler, Susanne. 2005. *Ellipsis and Focus in Generative Grammar*. Berlin; New York: Mouton de Gruyter.
- Worth, Dean Stoddard. 1961. *Kamchadal Texts Collected by W. Jochelson*. The Hague: Mouton.
- Zwart, Jan Wouter. 2005. Some notes on coordination in head-final languages. *The Linguistics in The Netherlands* 232–241.