The emergence of conjunctions and phrasal coordination in Khanty

Lena Borise
Hungarian Research Centre for Linguistics
Budapest, Benczúr utca 33, 1068 Hungary
ORCID: 0000-0002-6356-7472
lena.borise@nytud.hu

Katalin É. Kiss, corresponding author
Hungarian Research Centre for Linguistics
Budapest, Benczúr utca 33, 1068 Hungary
ORCID: 0000-0003-2054-4593
e.kiss.katalin@nytud.hu

Abstract
Prior to widespread contact with Russian, Khanty (Uralic; Finno-Ugric) did not have overt conjunctions or phrasal coordination. Instead, Khanty texts from the 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 appeared in the context of clausal coordination, followed by coordination of smaller phrases. Based on novel elicitation data, we demonstrate that, in contemporary Khanty, (i) overt conjunctions are commonplace, and, additionally, (ii) coordinated clausal constituents may be derived via phrasal coordination or clausal coordination with conjunction reduction/ellipsis, but (iii) ellipsis of syntactic heads is banned (nouns & postpositions) or dispreferred (verbs). Based on this diachronic picture, we conclude that the 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), the West-Siberian Chukchee (Bogoras 1922), and Kamchadal (Worth 1961) – that syndetic coordination has arisen in 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. 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 is 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 the coordination of

clausal constituents has become possible, we show that, in today's Khanty, it may arise from the coordination of phrasal constituents or coordination of clauses followed by ellipsis, but ellipsis of syntactic heads is banned (nouns & postpositions) or dispreferred (verbs).

We interpret the facts 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 the interaction of particular features of Khanty morphosyntax and a processing maxim, the principle of Maximize On-line Processing (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 Sources and methodology

The data from older stages of Khanty come from four corpora. We chose corpora of comparable sizes that were recorded at intervals of roughly 30 years in the 20th century (around 1900, in the thirties, sixties, and nineties). Three of the four periods could be represented by easily accessible prose texts of the Ob-Ugric Database (OUDB), converted to standardized IPA transcription and annotated morphologically.

The oldest variety of Khanty considered here is described based on 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). 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. Maremjanin, a speaker of the northern Sherkaly dialect, written in 1936 and published by Steinitz in 1989, were consulted. The next stage of Khanty is represented by a 3740-word corpus of tales and ethnographical narratives collected by Rédei in the northern Kazym area in 1964, annotated in the OUDB. Finally, we consulted a 4200-word corpus of Eastern Khanty tales and narratives collected in the 1990s by Márta Csepregi (Csepregi 1998, 2002), also annotated in the OUDB.

Contemporary Khanty data come from elicitation sessions with three female speakers of Eastern (Surgut) Khanty (age range: 21-69 y.o.). During the elicitation sessions, the speakers were asked to evaluate pre-constructed sentences in Surgut Khanty that contained coordinate structures. If the pre-constructed examples turned out to be ungrammatical or infelicitous, the speakers were asked to provide their own paraphrases. Additionally, a small spoken corpus, collected from two female speakers (age: 29 and 68) in the Surgut area in 2017 by Márta Csepregi and Katalin Gugán and containing short narratives and conversations on everyday topics, was consulted for further examples of coordination (http://www.nytud.hu/oszt/elmnyelv/urali/adatbazisok_szurgutihanti.html).

Our data sources come from a variety of genres. As such, we take them to be representative of Khanty in general. It should be noted that, in traditional texts and folk tales, lack of ellipsis and the resulting repetitiveness may be used as a rhetorical device, which gives the text a certain rhythm, and may help its memorization and performance. However, the data in Sections 3.1 and 3.2 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. Maremjanin's (1936) autobiographical notes are of special significance because, owing to their content, they are closer to everyday speech, and exempt from conventionalized constructions to a greater degree than the folklore texts.

Though Khanty dialects show significant phonological and morphological variation, their descriptive grammars – Nikolaeva's (1999) grammar of Northern Khanty, Gulya's (1966) and Csepregi's (1998, 2017) grammars of Eastern Khanty, and Filchenko's (2010) grammar of Vasyugan Khanty – do not show any differences with reference to the syntax of coordination. Neither did we find any differences with respect to the processes of language change described

here.[1] Therefore, 'Khanty' is used here to refer collectively to the Khanty varieties that data are available for. Relevant grammatical properties of individual Khanty varieties are described in the context of particular examples.

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 spread 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, 1964 and the 1990s, respectively. Section 4 discusses coordination in contemporary Khanty. Section 5 proposes an analysis of the observed phenomena. Section 6 concludes.

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 find 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.

(1) a. [*βɒ:jəx qu:ɬ kəntʃ-tɐɣə mənn-əs i:ki*], [*i:mi jaqqən quɯ:tʲ*].[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. [*ruɯ:t βɛr-ɬ-ətəɣ, βɛr-ɬ-ətəɣ*],[3] [*i:ttən torrəm pəttəɣłə-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:ɬ-s-i*], [*qu:ɬ łɐ:tʲ-s-i*],
go-PST.3SG fish_basket up pull-PST-PASS.3SG fish catch-PST-PASS.3SG
[*łɐ:rəɣ łɐ:tʲ-s-i*], [*sɒ:rt łɐ:tʲ-s-i*], [*jɐ:β łɐ:tʲ-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 cannot be distinguished from juxtaposed predicate phrases. Such examples do not have overt conjunctions either:

(2) [*qo:ɬəm sɒ:t tʃɒ:ət mə-ss-əm*], [*kɐ:rkɐm 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)

[1] A reviewer has called our attention to the fact that the Yugan dialect, represented by our oldest corpus, came into contact with Russian at a later time than the other dialects analyzed here. This can mean that the process of the introduction of conjunctions and phrasal coordination that has taken place in the Khanty language may have started later in the Yugan dialect.

[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. Glosses follow the Leipzig Glossing Rules, with the following additions: ABESS – abessive, APPR – approximative, LAT – lative, TRNS – translative.

[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 *βɛrłətəɣ* in (1b) could also be analyzed as a full clause with a *pro*-subject and a *pro*-object.

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

(3) a. [*qoːəpti βaɬ-ɬ-ət*] [*βɐːnəpti βaɬ-ɬ-ət*].[4]
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 tɐːə-nə qɒːɬ-s*], [*kɐːt tɐːə-nə qɒːɬ-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)

The use of adverbs or particles as connectives is extremely rare in the Paasonen corpus. We find a few occurrences of *pɐːnə/paːne*, the locative form of *pa* 'other', to be interpreted additively, meaning 'otherwise' or 'on the other hand'. By today, *paːne* has grammaticalized into a conjunction meaning 'and', but these early texts still show no sign of its grammaticalization into a regular linking element:[5]

(4) *βɛɬi βɛɬ,* ***pɐːnə*** *nʲoβ βɛɬ.*
reindeer kill-PST.3SG 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'. Languages without a lexical equivalent of *or* as a disjunctive coordinator often employ lexical elements expressing uncertainty, or irrealis markers, to encode disjunction (Mauri 2008, Jing-Schmidt & Peng 2015).[6] (5b), with the second *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.

(5) a. *tot* ***məβ*** *əj qatɬ βaɬ-s-əɣən* ***məβə*** *kɐːt qatɬɣən βaɬ-s-əɣən.*
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.'
(OUDB 1314)

b. *əj łɐːpət łiː-s-tən,* ***məβə*** *kɐːt łɐːpət-ɣən łiː-s-tən.*
one week eat-PST-SG<3DU about two week-DU eat-PST-SG<3DU
'They ate it for one week, perhaps they ate it for two weeks.' (OUDB 1316)

Additionally, instead of clausal juxtaposition, propositions describing simultaneous or consecutive actions are often connected by the subordinating one of the propositions in the form of a non-finite verbal projection: a converb (6a), or a present or past participle (6b) (see also Sipos 2015). This type of structure is often analysed as a type of cosubordination, distinct from subordination (Van Valin 2005).

[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.

[5] In languages that develop overt coordinators, conjunctions most often result from grammaticalization of comitatives and adverbials with the meaning 'also' (Mithun 1988). While *pɐːnə* does not fit into either of these categories, the other Khanty conjunction, *oːs* 'and', discussed in Section 4, derives from the adverbial 'also'.

[6] A reviewer points out that in the eastern Vakh and Vasjugan dialects of Khanty, a different indefinite pronoun, *qam*, is also used when expressing approximate numbers (cf. also Steinitz 1966-1993).

(6) a. *jəŋk qɒːnəŋ-nə nʲeːβər łiːβ-min suːtʃəɣtə-ł-ɣən.*
water bank-LOC foam eat-CVB walk.around-PRS-3DU
'They walk around, eating foam on the river bank.' (OUDB 1314)

b. *iːmi βɐːn-ɣə joβt-əm łɐːt-nə juːβ toj-ɐ quːŋt-əs.*
woman near-TRNS come-PTCP.PST time-LOC tree top-LAT climb-PST.3SG
'The woman having come close, he climbed to the treetop.'[7] (OUDB 1315)

In the construction type illustrated in (6b), the initial participial clause often repeats the content of the previous finite clause, thereby performing a bridging function – as discussed by Filchenko (2010: 529).[8]

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', borrowed from Russian, that may occasionally occur in sentence coordination. In contrast, the Paasonen-corpus we examined contains no such examples. This is expected, given that they come from Eastern (Yugan) dialects, which did not yet experience 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 a traditional Khanty clause, each grammatical function is represented by an expression denoting a single entity. If we described the event in (1c) in English, the catching of different kinds of fish in the same fish-basket at the same time would be described by a single clause involving coordinate noun phrases; in Khanty, however, the same predicate is predicated of each kind of fish in a separate clause. Situations where multiple participants have the same property or perform the same action (7a), or multiple objects are affected in a similar way by the same agent (7b) are expressed by multiple juxtaposed clauses. (The anaphoric temporal adverbial 'after that' linking the last clause of (7b) to the previous one is not a grammaticalized connective.)

(7) a. *torrəm jiːr βɛr-tɐɣə mɒːst-ł, məɣ jiːr βɛr-tɐɣə*
sky animal _sacrifice do-INF need-PRS.3SG earth animal_sacrifice do-INF
mɒːst-ł.
need-PRS.3SG
'A sky animal sacrifice needs to be made, an earth animal sacrifice needs to be made.' (OUDB 1313)

b. pro *puːpi toβə quɯːj-s-ətəɣ, por-βɒːjəɣ toβə quɯːj-s-ətəɣ, βoqɯ*
bear there leave-PST-SG<3SG wolf there leave-PST-SG<3SG fox
toβə quɯːj-s-ətəɣ, tʃeːβər toβə quɯːj-s-ətəɣ, tʲuːt puɯːrnə
there leave-PST-SG<3SG hare there leave-PST-SG<3SG that after
kəmłəɣ quɯːj-s-ətəɣ.
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, then he left behind the wolverine.' (OUDB 1315)

When a subject performs multiple actions, clausal juxtaposition is usually indistinguishable from vP-juxtaposition, as a repeated subject tends to be represented by a silent *pro*.

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

[8] This strategy of clause-chaining has also been described for Oceanic languages – see Ohori (2004) and Terrill (2004).

Nevertheless, there are also examples with a repeated overt subject, like (8). (The additive particle *o:s*, 'also', linking the third clause to the second one, has by today grammaticalized into a conjunction.)

(8) [*ɐ:łəŋ łɐ:tnə o:pi-ł łi:-s-i.*] [*tʲu:t puɯ:rnə*
first time older_sister-3SG eat-PST-PASS.3SG that after
pi:tʃənɣəli *kɐ:tł-s-i,*] [*o:s* ***pi:tʃiŋɣəli*** *noq łi:-s-i jəppəɣ-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 (9). (9) has a complex structure: the juxtaposed clauses form correlative units pairwise, with the indefinite subject of the first member of the pair co-referring with the *pro* subject of the second member:

(9) *qo: əj βəłi toj, ɯ:łə βɛł-təɣ, qo: əj łɐ:β taj-ɐs,*
someone reindeer have.PST.3SG down kill-PST.SG<3SG someone horse have-PST.3SG
ɯ:łə βɛł-təɣ, qo: əj mɛs taj-ɐs, ɯ:łə βɛł-təɣ,
down kill-PST-SG<3SG someone cow have-PST.3SG down kill-PST-SG<3SG
qo: əj ɒ:tʃ taj-ɐs, ɯ:łə βɛł-təɣ pɒ:ri βɛr-tɐɣə.
someone sheep have-PST.3SG down 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 to circumvent the coordination of subject or object NPs participating in the same event is to supply all but one of them with a comitative suffix.[9]

(10) *tʲu: qo: i:mi-ł-**nɐt** nʲe:βrem-əł-**nɐt** tot ɒ:məs-ł-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 apply to the same constituent, the modified constituent is repeated. Example (11a) involves repeated VPs with adverbial modifiers of the same type. As illustrated by (11b), 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.

(11) a. *i:ttən łi:tot əntem, ɐ:łəŋ łi: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. *nɐ:j-nɐm ɒ:məst-ł səjłək ɒ:məst-ł.*
fire-APPR sit-PRS.3SG silently sit-PRS.3SG
'He sits next to the fire, he sits silently.' (OUDB 1313)

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

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 *βɒːjəx quːɬ*, 'game fish', i.e., 'animals hunted on land and in water', in (1). 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 (for prosodic evidence, see Section 3.4). 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* (12a), *rain-wind* (12b), *woman-man* (13a), *house-barn* (13c), or temporary pairs determined by a particular situation such as a fox and a hare acting jointly in a fairy tale (13b). The co-compound *game-fish* in (12c) has a collective meaning.

(12) a. ***kət-iɬ*** ***kər-iɬ*** *iːɬə* *moːritə-min* *təβt-ɐ* *pan-tɐɣə* *mɒːst-ɬ*
hand-3SG foot-3SG off break-CVB fire-LAT put-INF need-PRS.3SG
nʲeːβrem-əɬ-nɐt.
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-ɣə*** ***βɐːt-ɣə*** *jə-s.*
weather rain-TRNS wind-TRNS become-PST.3SG
'The weather turned into rain [and] wind.' (OUDB 1313)

c. *tʲ* *βɛɬ-m-ɐɬ* ***βɒːjəx*** ***quːɬ*** *məŋɐti* *piːtɬ*
so catch-PTCP.PST-3SG game fish we.DAT fall-PRS.3SG
'So the game [and] fish caught by him will fall on us.' (OUDB 1313)

d. ***juːβ*** ***poːm*** *əj tʃɐːmə* *ɬiː-ɬ-təɣ*
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 (13a–c).[10] Note that in (13b), the event involves three referents with the same function. Only two of them can form a co-compound, which is signaled by dual marking; the third one figures in a separate clause.

(13) a. ***iːmi-ɣən*** ***iːki-ɣən*** *paɣ* *taj-s-əɣən.*
woman-DU man-DU son have-PST-DU
'The woman [and] the man had a son.' (OUDB 1315)

b. ***βoqɯ-ɣən*** ***tʃeːβər-ɣən*** *ɬəβ* *jot-ɐɬ* *jə-s-ɣən,* *kəmɬəɣ* *ɬəβ*
fox-DU hare-DU he with-3SG come-PST-3DU wolverine he

[10] 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 find that dual marking is also present on the members of inanimate co-compounds if they have unique singular referents.

jot-ɐ̈ł *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* *ōməs-ta* ***xōt-eŋen,*** ***tabas-eŋen-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 (12c) and (14), and the object in (12d) – elicit singular verbal agreement.

(14) ***ruːtʲ*** ***qantəɣ*** *qoɬnə* *sɐːrnɐm* *βaɬ-ɬ?*
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 dual marking in a co-compound 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 (10) above.

A co-compound may be also dominated by a PossP, with possessive agreement spelled out on both nominals – see (12a) and (15b). The members of a co-compound share the same case, and the case suffix is usually present on both – see (12b) and (15c). In (13c) and (15a,b), 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 (15a,b). When there is no other suffix to signal their morphological parallelism, the case morpheme must be spelled out on both nouns, as in (15c).[11]

(15) 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 (12c) and (13c) each contain a modifier. Although 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 (13c).[12] A bare co-compound, e.g. the one in (12c), is presumably dominated by an NP node.

[11] Examples (13c) and (15a,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.

[12] 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.

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:

(16) a. *Mɐ:* ***ənəɬ βojən*** *sɒ:rt* *qɒ:tɬ-əm.*
1SG big fatty pike catch-PST.1SG
'I caught a big-fatty pike.'

b. *Mi:ʃɐ* ***ki:t qaɬəm*** *sɒ:rt* *tu:β.*
Misha two three pike bring.PST.3SG
'Misha brought two-three pikes.'

c. *je:ji-ɣən* *manʲi-ɣən* ***ɬi:k-kən*** ***jinʲtʲi-ɣən***
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. This suggests 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 heads post-syntactically, before lexical insertion. The inflectional morphemes are realized on both juxtaposed elements.

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 Maremjanin's autobiographical notes from 1936: the first conjunctions

In the autobiographical notes of K. M. Maremjanin, written in 1936, the typical way of coordinating clauses is juxtaposition, too – see (17), where the two juxtaposed clauses describe consecutive events:

(17) *Wɔš-na* *kăt* *mŭj* *χutəm* *χătḷ* *ut-əs,* *kara* *peləк* *peta*
city-LOC two or three day be-PST.3SG Kara side in.the.direction.of
kĭt-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 (18).

(18) *Tŭjt-em jĭŋk-a nŏp-əs. Tŭjt-em jĭŋk-a nŏp-əm*
sledge-1SG water-LAT drift-PST.3SG sledge-1SG water-LAT drift- PTCP.PST
tăχə-na nŏwə 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:

(19) a. *Jaj-em tɔw-ŋ-ət kĭr-əs* ***ĭ*** *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-ŏt-na tusa tapət-s-əte,* ***a*** *tumət-sŏχ-na ănt*
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, i.e., no omission of the repeated material in all but one of the conjuncts. The juxtaposed clauses of (20), for example, contain the same verb iterated four times; the coordination of the objects is avoided:

(20) *Jŏnttə tɔw-ət wer-s-əm, jŏnttə uχt-ət wer-s-əm, jŏnttə*
playing horse-PL make-PST-1SG, playing sledge-PL make-PST-1SG, playing
sese-t wer-s-əm, jŏnttə 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 (21), linked by the Russian conjunction *i*, have different subjects and a shared VP. The VPs are spelled out without ellipsis; they are iterated three times:

(21) *Tăm zawod-ət fabrikaj-ət ŭw-t-ət* ***ĭ*** *tŭtəŋ-tŭjt-ət ŭw-t-ət*
this plant-PL factory-PL roar-PST-3PL and fiery-sledge-3PL roar-PST-3PL
ĭ *awtomobil-ət ŭw-t-ət.*
and car-PL roar-PST-3PL
'These plants-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:

(22) *Men jăχ-s-amn̥ sŏta-jŏχan-a* ***ĭ*** *muχtəŋ-jŏχan-a χŭt kăš-ta.*
1DU go-PST-1DU Sŏta-river-LAT and Muχtəŋ-river-LAT fish look.for-INF
'We went to Sŏta-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 mǐγ-na **ǐ** χăr mǐγ-na* ‘forest ground-LOC and tundra ground-LOC’ (Steinitz 1989: 185), *aśe-m **ǐ** śatśaśe-m* ‘father-1SG and grandfather-1SG’ (Steinitz 1989: 175). Nevertheless, genuine co-compounds are still common and productive. Though some of them may be conventionalized or lexicalized (*tɔwn̩ sŭsn̩* ‘in spring - in autumn', or *mŏn tɔwtəw mistəw* ‘our horses - cows’ in (23a) below), there are also novel expressions such as *zawodət-fabrikajət* ‘plants-factories’ in (21).

(23) a. ***mŏŋ tɔwtəw mĭstəw*** *śŏras χu-na arə pŭš*
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. ***nŏwə jɔχ wŭrtə jɔχ*** *t'at'əs-ta pĭt-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 *mŏn tɔwtəw mistəw* 'our horses - cows' in (23a) share an overt possessor, which elicits agreement on both of them, and the members of the co-compound *tumət-sŏχ-na* 'clothes-overcoat-LOC' in (19b) 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 Yugan dialect illustrated in (12), the case suffix in Maremjanin’s text can be spelled out only once, whether or not the members of the co-compound bear identical agreement suffixes.

Above, (17) contains coordinated numerals connected by *mŭj*, the Northern Khanty equivalent of the indefinite/approximative pronoun *məβə* (*kăt mŭj χutəm χătl̩* ‘two or three days’). In the example (5a) 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 *kăt mŭj χutəm χătl̩* ‘two or three days’ is an instance of phrasal disjunction and conjunction reduction, with *mŭj* functioning as a disjunctive particle, or the pattern 'numeral *mŭj* 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 Maremjanin 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 see the occasional use of the conjunction *i/ij* between clauses, and sporadically, between noun phrases as well. In this respect, Maremjanin's language use is likely to be ahead of the general course of evolution in Khanty because he became a balanced Khanty-Russian bilingual before the majority of Khanty did. As we learn from his autobiography, he worked for Russian merchants from 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 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:

(24) *βuɬi* *sox* *jɛməŋ* *taxi-ja* *ixət-ɬ-a,* *sʲata* *xaj-ɬ-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 see spreading of a native conjunction, *paː*, a cognate of the additive adverb *pɐːnə* '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.[13] While the texts collected by Paasonen in 1901 contain 4 occurrences of *pɐːnə,* and a single occurrence of *məβə* functioning as a linking adverb, the Rédei texts contain 56 instances of *paː*, 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, it displays a marked increase in the use of overt conjunctions. 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.

The Rédei corpus contains no instance of *i*, the Russian conjunction used by Maremjanin. In later Kazym texts of the Ob-Ugric Database *i* does occur sporadically; we find 13 occurrences of *i* as opposed to 142 occurrences of *paː*. The eastern Surgut texts contain only *pɐːnə,* whereas in the recent Yugan texts of the database, not analyzed in this paper, the numbers of *i* and *paː* is practically even. Here we only give account of the particles attested in the corpora that we have analyzed. The choice between *i* and *paː(nə)* across the Khanty language area appears to have dialectal, regional, sociolectal, and idiolectal components, the examination of which is beyond the scope of this paper.

In the Rédei corpus, *paː* occurs both between clauses (25a), and in the initial or post-topic position of independent sentences (25b) – i.e., it functions as an additive particle with an adjacent, distant, or implicit first associate. Because of *pro*-drop, clausal coordination is often indistinguishable from VP-coordination (25c):

(25) a. *nʲɔː-ɬ* *juβtəs-əɬ* ***paː*** *mɔːjpər* *xɔːj-ɬ-a.*
arrow shoot-PRS.3SG and bear hit-PRS-PASS.3SG
'The arrow shoots **and** the bear is hit.' (OUDB 1022)

b. *iki-le-ɬ* *iɬ* *oɬ-əs.* *imi-le-ɬ* *łɛtot-łaɬ* *pasan-a*
man-DIM-3SG down lie-PST.3SG woman-DIM-3SG food-PL<3SG table-LAT
ʃaβi-s-əɬɬe. *łuβ* ***paː*** *iɬi* *oɬ-əs.*[14]
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. [*ʃɛpan* [*sʲart-əɬ*]] ***paː*** [pro [*lop-əɬ,* *muj* *kanʃə-ɬ* *jiŋ*
shaman shamanize-PRS.3SG and tell-PRS.3SG what search-PRS.3SG water
βɵːrt]]
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:

[13] Please note that we do not claim that the rate of grammaticalization of *pɐː*(*nə*) was necessarily the same across the Khanty varieties, which we compare at different time points – only that the starting point, *pɐː*(*nə*) used adverbially, and the end point, *pɐː*(*nə*) as a conjunction, are the same.

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

(26) *jeːtn-a* *ji-ł,* *pasan-ən* *isʲiti* *łɛtoti*
evening-LAT become-PRS.3SG table-LOC same.way full_of_food
xaj-ł-ɛm, ***aː*** *min* *ant* *oł-ł-əmən* *łaːβəł-ti* *pit-ł-əmən.*
leave-PRS-SG<1SG 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 find a disjunctive connective, grammaticalized from *muj* 'what' (the equivalent of the Eastern Khanty *məβ*), originally an indefinite used as an approximator, as discussed in connection with (5). A similar pair of sentences in the Paasonen tales, cited under (3a), are still juxtaposed without a connective.

(27) pro *xuβ* *man-əs* ***muj*** *βaːn* *man-əs*
long go-PST.3SG or short go- PST.3SG
'He went for a long time, **or** he went for a short time.' (OUDB 1117)

While phrasal coordination was barely present in the Maremjanin notes, the 1964 texts already contain different types of constituents conjoined by *paː,* among them modified NPs (28a) and predicative adjectives (28b).

(28) a. *łuβ sorm-a* *ji-te-ł* *jupijən* *sʲar-łał* *meːt* *aːj*
he death-LAT become-PTCP.PRS-3SG after shaman-PL.3SG most small
pox-łał-a ***paː*** *mɛt* *aːj* *eːβi-łał-a* *pit-ł-ət.*
son-PL.3SG-LAT and most small daugher-PL.3SG-LAT pass-PRS-3PL
'After he dies, his shamanic skills go to his youngest sons **and** youngest daughters.' (OUDB 878)

b. *jəːxəł-łał* *məːjpər* *βoxəłtə-ti* *pata* *βɛrən-s-aj-ət* [*βəːna-ʃək*]
bow-PL<3SG bear overcome-PTCP.PRS for make-PST-PASS-3PL big-COMP
paː [*taːka-ʃək*].
and fast-COMP
'His bows for shooting a bear were made bigger **and** faster.' (OUDB 1022)

Paː 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 (29):

(29) *sʲi* *βeːsʲəŋ* *nɛŋ-əł,* *xorasəŋ* *nɛŋ-əł* *piła* *aŋk-eł* *aːsʲe-ł*
this pretty woman-3SG beautiful woman-3SG with mother-3SG father-3SG
xosʲa *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 with *muj*. In most of these cases, *muj* has already lost its approximator function. *Muj* is sometimes strengthened by *paː,* in which case *paː* has its adverbial meaning, 'else'. If *muj* precedes both disjuncts, as in (30b), only the one between the two disjuncts is combined with *paː*.

(30) a. *ałmənti* *ki* *sʲi* *βułi* *is-əł* [*sɛmsajot-ət-a*] ***muj*** ***paː*** [*təːrəm-a*]
as if that reindeer soul-3SG spirit-PL-LAT or else god-LAT
man-əł.
go-PRS.3SG
'supposedly that reindeer's soul goes to the spirits **or else** to god.' (OUDB 878)

b. *łiβ* pro *ʃɛpan-łał-ən* ***muj*** [*ła:jəm ixət-man*] ***muj*** ***pa:***
they them shamanhood-3PL-LOC either axe hang-CVB or else
[*pe:nʲsʲar-ən sɛŋk-man*] *oʃ-a βɛrənt-ł-əłłał*.
shaman.drum-LOC beat-CVB effort-LAT 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)

Sporadically, we also find structures with non-adjacent disjuncts, which are likely to result from conjunction reduction. Thus, a possible derivation for (31) is stripping, as indicated; or else the constituent preceded by *muj* has been extraposed or is an afterthought.

(31) pro *joxt-əm jɔ:x-łał-ən* [*jiŋk-ən tɛm-ł-aj-ət* ***muj*** ***pa:***
come-PTCP.PST people-PL-LOC water-LOC pour-PRS-PASS-3PL or else
[*łɔ:sʲ-ən* ~~*tɛm-ł-aj-ət*~~].
snow-LOC pour-PRS-PASS-3PL
'They are splattered by the coming people with water **or else** with snow.' (OUDB 1022)

Phrasal coordination and conjunction reduction also occur sporadically in constructions with idiosyncratic coordinating elements, which suggests that the role of conjunctions in licensing these phenomena is pragmatic rather than formal. In (32), the context shows that the clause-initial *isʲiti* 'in similar ways' is not anaphoric but cataphoric; it expresses that the NPs following it are subjects of the same burning event:

(32) *isʲiti βɔ:j sox-ət, sʲaʃkan sa-xət oxʃam-ət, sʲaʃkan-ət tut-a*
same.way animal fur-PL calico coat-PL headscarf-PL calico-PL fire-LAT

a:ptijəł-s-aj-ət.
feed-PST-PASS-3PL
'Calico coats, headscarves, calico cloths were fed into the fire in similar ways.'
(OUDB 878)

In (33), three parallel clauses are juxtaposed. They are linked by the numeral modifiers *one*, *second*, *third* at the beginning of the clauses. The verb is only spelled out in the first clause, i.e., the second and third clauses involve gapping. Gapping is defined for SVO languages as an elided verb(al complex) flanked by the subject and a remnant (the object, an oblique argument, or an adjunct) in the non-initial conjunct of a coordinate construction. In SOV languages like Khanty, the elided verb is final (and ellipsis can also take place in the first conjunct – see Ross (1970)); see also Section 4.5.

(33) [pro *ij sə:n xint-əŋ imi jə:ʃ kutəp-a ɔ:məs-s-əłłe*],
one vessel knapsack-ADJ woman path middle-LAT place-PST-SG<3SG
[*kimət sə:n kə:rt xɔ:nəŋa*], [*xə:łmit sə:n ɔ:β kimpija*].
second vessel village beside third vessel door outside
'He placed one of the vessels in the middle of the path of the woman with the knapsack, the second vessel at the border of the village, the third vessel in front of the door.'
(OUDB 883)

The spreading of conjunctions and the emergence of phrasal coordination have not rendered co-compounding obsolete; on the contrary, we find more co-compounds (and less repetition of clauses with identical constituents) than in the earlier texts.

(34) a. ***ox-əł*** ***sɛm-əł*** *montəłmə-s*
head-3SG eye-3SG wrap-PST.3SG
'He covered his head [and] eyes.' (OUDB 883)

b. ***xɔ:t-ət*** ***łopas-ət*** *βɛrən-s-aj-ət*
house-PL storehouse-PL make-PST-PASS-3PL
'Houses [and] storehouses were made.' (OUDB 878)

In (35), the juxtaposed nominals have a single locative suffix:

(35) ***łə:ŋ tur*** ***ka:ł tur-ən*** *xojat*
male_protective_spirit's_voice femal_protective_spirit's_voice-LOC somebody
uβ-te-ł *sa:tʲ-əł*
shout-PTCP.PRS-3SG be_heard-PRS.3SG
'Somebody was heard shouting in the male protective spirits's voice [and] the female protective spirit's voice.' (OUDB 1117)

In sum, the 1964 Kazym Khanty corpus shows the emergence of the conjunctions *pa:* 'and' and *muj* 'or', grammaticalized from native words. Parallel with the spreading of conjunctive particles, phrasal coordination also appeared.

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 OUDB 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. While the 3700-word Rédei corpus from 1964 contains 56 occurrences of *pa:*, the 4200-word Csepregi corpus contains 126 instances of *pɐ:n(ə)*. Furthermore, it also contains 58 instances of the additive particle *o:s*, which also developed into a conjunction.

As before, *pɐ:nə* and *o:s* do not always appear between two conjuncts; their first conjunct can also be distant, or implicit – recall example (8), from the Paasonen corpus, and example (25b), from the Rédei corpus. In (36a), *pɐ:nə* conjoins three clauses; in (36b), it adds a sentence to the preceding independent sentence. (37) contains both types of *o:s*: a conjunction linking clauses, and an additive particle with a distant or implicit antecedent.

(36) a. [*ətʲə* *su:βəm-ɐt* *məj-i*] ***pɐ:nə*** [*pɛrt* *pu:l-ɐt*
again reel_of_thread-INS give-PST.PASS.3SG and wood piece-INS
məj-i], ***pɐ: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.' (OUDB 736)

b. *tʲu:* *ke:nʲɐr* *qo:-pərili* *nʲaβmił*: "*mɐ:* *nɯŋ-ɐt* *ɛłə*
that poor man-like_thing say.PST.3SG 1SG 2SG-ACC just
nʲu:łtiptə-ł-əm." ***pɐ:n*** pro pro *ɛłə* *tʲi* *nʲu:łtiptə-təɣ*.
make.swear-PRS-1SG and 3SG 3SG just so make.swear-PST.SG<3SG
'That poor devil says: "I make you swear."' **And** he made him swear.' (OUDB 737)

(37) *tʲu:* *βe:łi-t* ***o:s***, *βe:łi-t* *u:łə* *ki:t-łɐ-t*, ***o:s*** *jaqə*
those reindeer-PL also reindeer-PL down catch-PRS-PASS.3PL and home
βəjɐ-t, *qɒ:t* *łəɣpinə* *u:ł* *łołɐłt-ɐt*.
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.' (OUDB 730)

As is also clear from (37), 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, (38) describes three consecutive events: the first one is expressed by a nonfinite verbal projection, but the second and third clauses are simply juxtaposed:

(38) *tʲi suːłtə-m-ɐm-ɐ ʉlɐk nʲuːr mɐː-n kətʃəɣ-nɐt ɛβətəm-i*
so slip-PTCP.PST-1SG-LAT harness tether 1SG-LOC knife-COM cut-PST.PASS.3SG
mɐː-nə βɐːləɣ qoβit uːtnɐm quːɣł-əm.[15]
1SG-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.' (OUDB 730)

The use of *pɐːnə* 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 (39).

(39) *jɐː tʲuː aβəł-ɐ łɛł-əm, tʲaqɐ* ***pɐːn*** *jaqə sɒːɣəltə-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 *pɐːnə* are also attested. The use of *pɐːnə* is not restricted to binary combinations of nominals, unlike co-compounding:

(40) *əj mətɐ łɐːt-nə* [*tuːʃəŋ ɐːɣən*] ***pɐːnə*** [*βɐːtʲ kʉr βɛjɣən*] ***pɐːn***
one some time-LOC Bearded_Chin and Two_Thin_Legs and
[*oːɣ łołəŋkən*] *βał-ł-ət.*
Two_Temples live-PRS-3PL
'There once lived Bearded Chin and Two Thin Legs and Two Temples.' (OUDB 1346)

Remarkably, *pɐːnə* is by now an alternative to the dual suffix. One of the best-known Khanty tales has the title *piːtʃiŋɣəli-ɣən oːpisɐː-ɣən* '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 *piːtʲəŋkəli pɐːnə oːpi*, i.e., the dual suffixes of the noun phrases have been replaced by the conjunction *pɐːnə.* 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 is only beginning; co-compounding is still a common way of coordinating NPs, e.g.:

(41) a. *mʉβ* ***liːtot-ɐt*** ***quːł-ɐt*** *əntə łɐːpət-ł-o?*
what food-INS fish-INS not feed-PRS-PASS.2SG
'Aren't you fed with food [and] fish?' (OUDB 737)

b. *tʲi iːki* ***tɐːs-ət*** ***βaɣ-ət*** *jaqə iːłt-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 at the phrase level is fairly frequent. The most common disjunctive coordinator is *mʉβ* (in addition to disjunctive meaning, it still preserves its adverbial meaning that indicates uncertainty or approximate quantity).

(42) *mɐː paq-qə βał-m-ɐm-nə* [*jeːŋ ʉrəkkə qoː-łəm ɒːł-nə*] ***mʉβ***
I little.boy-TRNS live-PTCP.PST-1SG-LOC thirteen year-LOC or
[*jeːŋ ʉrəkkənʲəłə ɒːł-nə*] *βał-m-ɐm-ɐ*
fourteen year-LOC live-PTCP.PST-1SG-LAT
'me being a little boy, thirteen years old **or** fourteen years old, ...' (OUDB 730)

[15] 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).

Parallel coordinate constructions with identical sub-constituents still display no conjunction reduction in many cases. In (36a) 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 (43), the identical subjects of the three clauses are spelled out in each clause:

(43) ***sɐ:pəł*** ***łoβ*** *ɯ:łə* *kərəɣ-m-ɐł* *łɐ:t-nə* ***sɐ:pəł łoβ-əł*** *təт* *rək-kən*
neck bone down fall-PTCP.PST-3SG time-LOC neck bone-3SG here fly-PST.3DU
sɐ:pəł ***łoβ*** *o:s* *noq* *łaqqən-təɣ.*
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:

(44) *pɐ:* *ɒ:ntəp* *jəmsi* *qɒ:t* *pɛlək-ɐ* *qatəłtə-ł-łɐł,* *pɐ:* *ɒ:ntəp* *pəɣi*
some cradle right house side-LAT carry-PRES-PL<3PL other cradle left
qɒ:t *pɛlək-ɐ.*
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 (42), the disjunctive particle conjoins pairs of locative-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. At this point, 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 can be made about the prosodic properties of coordinated constituents and co-compounds.

Overtly coordinated constituents, but not co-compounds, may contain (i) pauses and/or (ii) a pitch reset between the constituents, and (iii) may be not intonationally parallel. The sub-constituents in co-compounds, in contrast, have a tighter prosodic connection: they show 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 by (45) and Figure 1: there is a pause and a pitch reset between the first and the second constituent, and the constituents are not intonationally parallel.

(45) (*əj* *mətɐ* *łɐ:t-nə*) *tu:ʃəŋ* *ɐ:ɣən* *pɐ:nə βɐ:tʲ kʉr βɛjɣən* *pɐ:n* *o:ɣ* *łołəŋkən*
one some time-LOC Bearded_Chin and Two_Thin_Legs and Two_Temples
βał-ł-ət.
live-PRS-3PL
'There once lived Bearded Chin and Two Thin Legs and Two Temples.'
(OUDB, 1346)

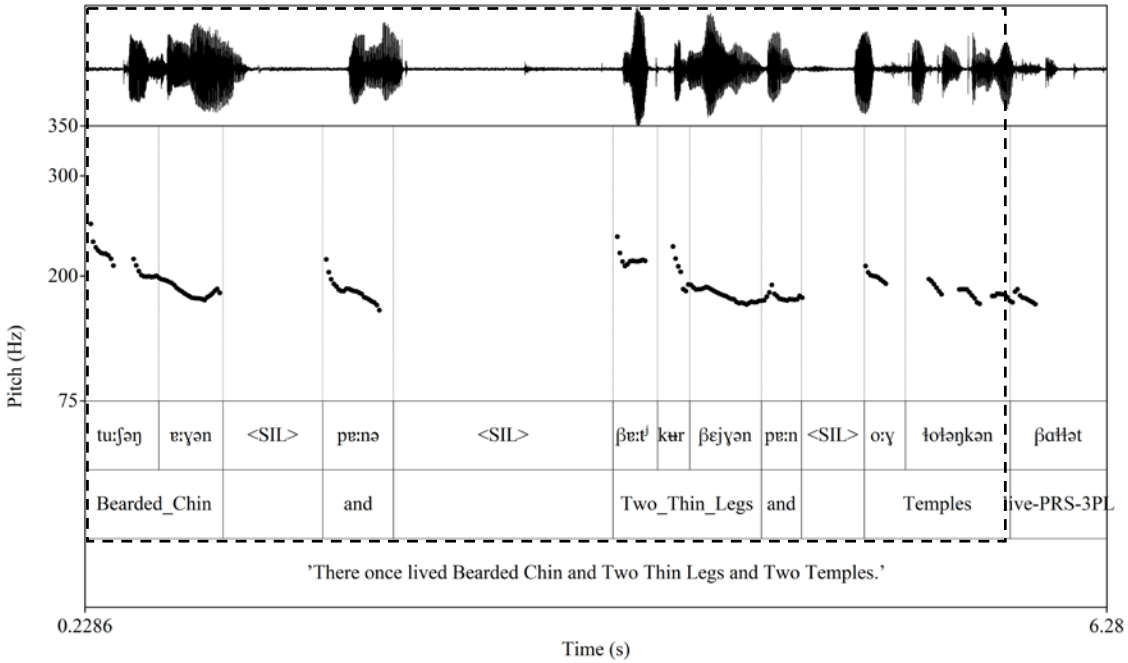

Figure 1. Realization of overtly coordinated nominals in (52) (boxed)

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

(46) [DP *βoqɯ-ən* *tʃeːβər-ɣən*] *ɬəβ* *jot-ɐɬ* *jə-s-ɣən.*
fox-DU hare-DU 3SG with-3SG come-PST-3DU
'The fox and the hare came with him.' (OUDB, 1315: 129)

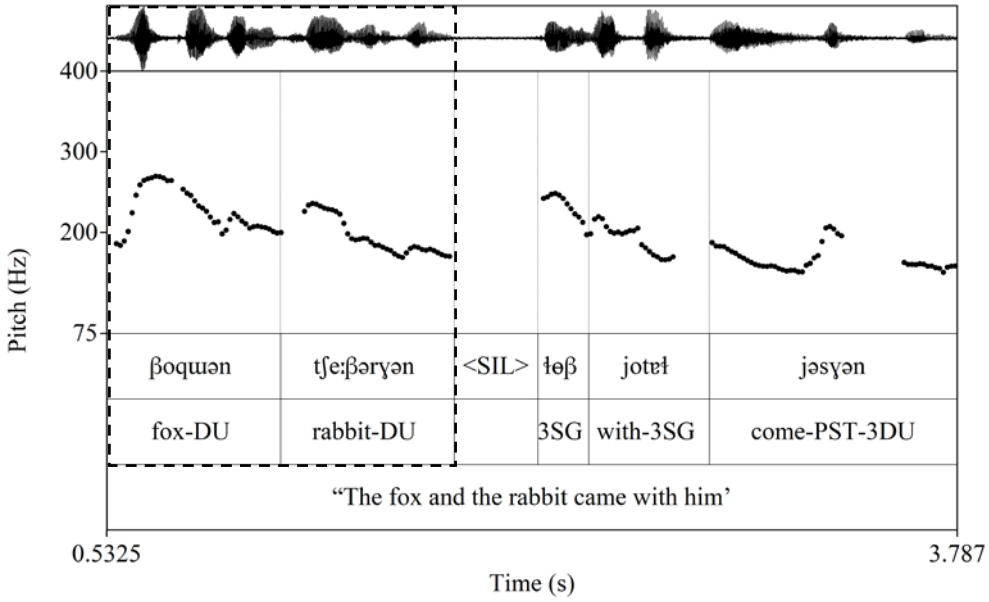

Figure 2. Realization of a nominal co-compound in (53) (boxed)

The prosodic facts lend support to the intuition that the sub-constituents of a co-compound are more tightly connected than overtly coordinated constituents. In the absence of a coordinator, prosody alone signals that the two adjacent constituents form a higher-order one. In contrast, the presence of a coordinator 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, 1964, and the 1990s show the gradual emergence of clausal coordination by means of conjunctive and disjunctive morphemes. In the document from 1936, we only find sporadic occurrences of the conjunctions *i* and *a* borrowed from Russian, while the texts from 1964 already show the recurring use of native, grammaticalized conjunctive and disjunctive morphemes. In the texts from the 1990s, their use is systematic, though 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 into 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 (replacement of double dual marking by a conjunction) also appear. Non-adjacent coordinates (resulting from gapping or stripping) are exceptional. 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 4.3 suggests that phrasal coordination in contemporary Khanty may result either from clausal coordination with conjunction reduction, or coordination of phrasal constituents. Sections 4.4 and 4.5, show that, in the context of coordination, ellipsis of nominal and postpositional heads is banned, and ellipsis of verbs is dispreferred.

Co-compounding is also productive, as shown in 4.6. It can often be used interchangeably with phrasal coordination, though only with constituents that are semantically related and morphologically parallel.

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 ungrammatical). The choice of conjunction is determined by the relative order of events described by the conjoined clauses: *pɐːnə* is used for consecutive events, and *oːs* [16] for contemporaneous ones.

(47) *Iːt* *iːttən.* *Mɐːʃɐ* *nɐːj* *ɯ̆ł-əł,* [(i)]***oːs/*** [(ii)]***pɐːnə*** *Miːʃɐ* *jəŋk*
now evening Masha fire light-PRS.3SG and Misha water
tuː-ł.
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.'

[16] There is considerable variability in vowel pronunciation both between subtypes of the same dialect spoken in different villages and even between individual speakers. 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. To ensure consistency, words of interest (e.g., conjunctions *oːs* and *pɐːnə*) are rendered in the form that is standard in the OUDB, in line with the preceding sections of the paper, as opposed to transliteration based on standard orthography.

Similarly, disjunctions between clauses are obligatorily overt:

(48) *Tem qatəɬ-nə ɐ:ŋk-em sɒ:rt tɐ:rt-ɬ* ***antɐ*** *ɐ:ɬʲ-em qu:ɬ qatʃəm*
this day-LOC mother-1SG pike fry-PRS.3SG or father-1SG fish hot
ji:nk βar-ɬ.
water cook-PRS.3SG
'Today, my mother is frying pike or my father is cooking fish soup.'

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

(49) *I:t i:ttən-ɣə jə-ɣ.* ***Pɐ:nə*** *Mɐ:ʃɐ nɐ:j ɯɬ* ***o:s***
now evening-TRNS become-PST.3SG and Masha fire light.PST.3SG and
Mi:ʃɐ 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. Coordinated verbs that share both a subject and an object may also be analyzed either as coordination of verbs or that of larger projections (VP/TP/CP) with non-overt arguments. The latter analysis seems more appropriate in conjoined clauses, where non-topical objects, not cross-referenced on the verbs, tend to be spelled out in each conjunct (50a). In (50b), the conjoined VPs/TPs/CPs involve pro-drop licensed by object–verb agreement.[17]

(50) a. *Mi:ʃɐ qu:ɬ qɒ:təp-i noq βəj, qu:ɬ-ət ɐ:r-ɣə pan*
Misha fish net-ABL up take.PST.3SG fish-PL many-TRNS put.PST.3SG
pɐ:nə *qu:ɬ i:ɬə nʲaqəs.*
and fish away scale.PST.3SG
'Misha took the fish out from the net, sorted it and scaled it.'

b. *Mi:ʃɐ sɒ:rt qɒ:ləmt-əɣ, jaqə tu:β-təɣ* ***pɐ:nə*** *ni:k*
Misha pike catch-PST.3SG home bring-PST.3SG<SG and to_water
moɬ-təɣ.
boil-PST.3SG<SG
'Misha caught a pike, brought it home and cooked it in a cauldron.'

4.2 Phrasal coordination

Phrasal coordination in contemporary Khanty is as ubiquitous as the coordination of clauses/VPs. First, consider coordinated nominals. Subjects as well as direct and indirect objects are used with overt conjunctions and disjunctions (though in many contexts, co-compounding is still used, as shown in 4.6).

Example (51) shows subjects (agents and patients) coordinated with the help of an overt conjunction; dropping the conjunction leads to a decrease in acceptability. When two (as opposed to more) nominals are conjoined with *pɐ:nə*, the verb carries dual agreement; however, the conjuncts are not marked by dual, unlike in co-compounding.

[17] The generalization about object agreement as necessary for object drop is consistent with Nikolaeva's observations (1999: 68). At the same time, in a different set of examples, our consultants judged coordinated verbs without object agreement as felicitous. This option seems to be more marked.

(i) *Mi:ʃɐ ne:βi βe:ɬi kɐ:təɬ pɐ:nə βɛɬ.*
Misha white reindeer catch.PST.3SG and kill.PST.3SG
'Misha caught and killed a white reindeer.'

(51) a. *Mɐːʃɐ*(*-*ɣən*) ***pɐːnə*** *Miːʃɐ*(*-*ɣən*) *iːrək-kən.*
Masha-DU and Misha-DU sing-PST.3DU
'Masha and Misha sang/used to sing (together or not).'

b. *Mɐː-nə* *sɒːrt* ***pɐːnə*** *jɛβ* *əj* *quːtʲəŋ-ɐ* *pəsɐn* *aβti-jɐ*
1SG-LOC pike and perch one space.near-LAT table top-LAT
pan-iɣən.
put-PASS.PST.3DU
'(A) pike and (a) perch were put beside each other on the table by me.'

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

(52) *Mɐːʃɐ* (*ɐːrəɣ-əɬ*) ***mɯβə (pə)/βəs*** *Miːʃɐ* *ɐːrəɣ-əɬ.*
Masha sing-PRS.3SG or Misha sing-PRS.3SG
'Masha or Misha is singing/sings.'

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

(53) *Mɐː* *sɒːrt* ***pɐːnə/mɯβə*** *jɛβ* *qɒːɬəmt-əm.*
1SG pike and/or perch catch/get-PST.1SG
'I caught (a) pike and/or (a) perch.'[18]

(54) *ɐːtʲi* *Miːʃɐ-ɣɐ* ***pɐːnə/mɯβə*** *Peːtʲɐ-ɣɐ* *βɛt* *tʲarɐs* *məj.*
father Misha-LAT and/or Petja-LAT five thousand give.PST.3SG
'Father gave Misha and/or Petja 5000 [rubles].'

Like nominals, adjectives that modify the same referent are coordinated with overt conjunctions. This is true of attributive adjectives, if they describe different dimensions of the same referent, as in (55a), and predicative adjectives, as in (55b). Disjunction works in a parallel way, as in (56).

(55) a. *Qoβ* ***pɐːnə*** *norəq* *juːɣ* *noβ* *ojɐɣɬə-tɐɣə* *ruːpɐk.*
long and straight wood branch find-INF difficult
'It is difficult to find a long and straight stick.'

b. *Iːttən* *ɬiːtot* *keːβrəm* ***pɐːnə*** *ɛpɬəŋ* *βɒːɬ.*
evening meal hot and tasty be.PST.3SG
'The dinner was hot and tasty.'

(56) a. *Mɐːntəm* *pəːɣtə* ***antɐ*** *βastə* *qantʃtʃə* *ot* *mɒːsəɬ.*
1SG.DAT black or blue draw thing necessary
'I need a blue or a black pencil.'

b. *Qantʃtʃə ot* *pəːɣtə* ***βəs*** *βastə* *βɒːɬ* *(mɐː* *əntə* *nom-ɬəm).*
draw thing black or blue be.PST.3SG 1SG NEG remember-PRS.1SG
'The pencil was black or blue (I don't remember).'

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

(57) a. *Keːʃkɐ* *juːɣ-ɐ* *pɛstɐɣə* ***pɐːnə*** *sɯj-ɬəɣ* *quːŋəɬ.*

[18] 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.

cat tree-LAT quickly and sound-ABESS climb.PST.3SG
'A cat quickly and quietly climbed up a tree.'

b. *Miːʃɐ kɐːt-ɣən **βəs/antəqɐ (pə)** qoːɬəm spːrt tuːβ.*
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, like in earlier varieties of Khanty? Speakers report 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 our consultants, (58) cannot be used interchangeably with a neutral, unmarked utterance *I caught a pike and a perch.* Instead, (58) emphasizes that catching a pike was more significant, that there were more pikes than perches caught, or that the speaker's primary intention was to catch a pike.

(58) *Mɐː spːrt qpːɬəmt-əm **pɐːnə** jɛβ qpːɬəmt-əm.*
I pike catch-PST.1SG and perch catch-PST.1SG
'I caught (a) pike and caught (a) perch.'

Similarly, the adjectives in (59) cannot be interpreted as applying to a single entity – that is, (59) cannot mean that a stick that is both long and straight is hard to find. Instead, it only means that long sticks are hard to find, and so are straight sticks; cf. (55a) above.

(59) a. *Qoβ juːɣ noβ ruːpɐk ojɐɣɬə-tɐɣə **(pɐːnə)** norəq juːɣ noβ ruːpɐk ojɐɣɬə-tɐɣə.*
long wood branch difficult find-INF and straight wood branch difficult 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 previous section showed, phrasal coordination is possible in today's Khanty, and overt coordinators are ubiquitous. The question that this gives rise to is what syntax phrasal coordination has. There are several options, which can be grouped based on whether they take phrasal coordination to result from coordination of phrasal constituents, or a process that takes full clauses and renders unpronounced certain parts of them. The 'phrasal' approaches include those that take individual phrases to be coordinated, with the help of a dedicated projection like &P (Munn 1987, Kayne 1994, Johannessen 1996) or phrasal adjunction (Munn 1992, 1993). According to clausal-coordination approaches, coordination of phrases may result from the coordination of full clauses followed by ellipsis of the identical material (Gleitman 1965, Wilder 1994, Schwarz 1999). The data from today's Khanty does not provide definitive support for either of these approaches.

Agreement facts are often taken to distinguish phrasal coordination from clausal coordination followed by conjunction reduction. If the verb in a clause with conjoined singular subjects takes singular agreement, this is compatible with a conjunction reduction approach. Non-singular agreement supports a phrasal coordination approach. Both agreement patterns are attested in Khanty:

(60) a. *Puːpi, oβər-kʉrəp-ot, βoqɯ, tʃeːβər **pɐːnə** kəmɬəɣ mɐː jɒːtam **jaŋq-ɐɬ**.*
bear wolf fox hare and wolverine I with.1SG go-PRS.3SG.
'The bear, wolf, fox, hare and wolverine go with me.'

b. *Spːrt, jɛβ paːnə ɐːɣərnə piːryi jiːŋk-a **neːβrəm-ət**.*

pike perch and ide back water-LAT jump-PST.3PL
'A pike, a perch, and an ide jumped back into the water.'

In the context of a disjunction, the second iteration of the shared object is possible though optional, even in the absence of agreement – cf. (50) – which speaks for conjunction reduction (unless we assume verb-coordination).

(61) *Qɐːntəɣ jɒːɣ quːɬ βɛr-łət **antɐ** (quːɬ) tar-łət.*
Khanty people fish boil-PRS.3PL or fish fry-PRS.3PL
'The Khanty people boil or fry fish.'

In contrast, there are other coordination constructions that can only be analyzed as resulting from phrasal coordination. First, phrasal coordination is possible with so-called collective/symmetrical predicates, such as *be alike* or *get divorced.* These predicates require a non-singular argument, and, accordingly, underlying clausal coordination is ill-formed (Curme 1931, Peters 1966, Lakoff & Peters 1966, Wilder 2019).The availability of such constructions in today's Khanty, as in (62), attests to the possibility of phrasal coordination that does not result from ellipsis. The dual marking in (62a,b) cannot be omitted.

(62) a. [*Mɐːʃɐ **pɐːnə** Peːtʲɐ*] *kiːt-ɣə mən*(-ɣən).*
Masha and Petja two-TRNS go-PST.3DU
'Masha and Petja got divorced.'

b. [*Mɐːʃɐ **pɐːnə** Peːtʲɐ*] *əj qorɐsəp*(-ɣən).*
Masha and Petja one alike-DU
'Masha and Petja are alike.'

c. *Mɐː (əj) ɐːnəɣ-ɐ* [*quːɬ mɐːrən **pɐːnə** quːɬ βoj*] *nʲuːłɐ 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.'

Second, a single focus particle, like *only*, can apply to two coordinated nominals in Khanty. This speaks against ellipsis, because an underlying structure either with a single or iterated *only* would be infelicitous:

(63) *Top Mɐːʃɐ **pɐːnə** Kɐːtʲɐ βoqi βuːj-ɣən.*
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 (especially in the context of disjunction).

4.4 Coordination within nominal phrases and PPs

In coordination, ellipsis of one of the head nouns or postpositions is banned. For postpositions, we illustrate this with Ps grammaticalized from spatial nouns. For nouns, we use possessive constructions and nouns modified by adjectives.

First, consider postpositional phrases.[19]

(64) a. [PP [DP *sʉɣymət*] [P *tompinə*]]
birch_tree behind
'behind the birch tree' (Lit.: 'on the birch tree's other side')

[19] In these PPs, the P has been grammaticalized from a spatial noun (*tompi* means 'other side', *quːtʲəŋ* means 'nearby place'), which has been incorporated into the locative case marker *-ne*. The DP complement of P was originally the possessor of the spatial noun (É. Kiss (to appear)).

b. [PP [DP *ɬɯβ*] [P *quːtʲŋiɬnə*]]
he nearby
'next to him/her' (Lit.: 'in her/his nearby space')

In a coordination, both postpositions must be spelled out:

(65) *Juːɣ-ət* [PP [PP [DP ***qnːt***] [P *(***quːtʲəŋnə***)]] *pɐːnə* [PP [DP ***kiːβri***] [P *(***quːtʲəŋnə***)]]]
tree-PL house nearby and well nearby
oβr-et.
tall-PL
'The trees in the vicinity of the house and in the vicinity of the well are tall. (=two groups of trees)'

If one of the postpositions is omitted, the interpretation is that both DPs apply to the same postposition. Accordingly, these contexts rely on the coordination of the DP complements and not the ellipsis of the postposition.[20]

(66) *Juːɣ-ət* [PP [DP ***qnːt*** ***pɐːnə kiːβri***] [P ***quːtʲəŋnə***]] *oβr-et.*
tree-PL house and well nearby tall-PL
'The trees in the vicinity of the house and the well are tall. (=one group of trees)'
NOT: 'The trees in the vicinity of the house and in the vicinity of the well are tall. (=two groups of trees)'

Turning to nouns, 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).

(67) [DP *Iːβɐn* [NP *rɯt*]]
Ivan boat
'Ivan's boat'

Possessive constructions may be coordinated, as in (68a). If one possessor is omitted, as in (68b), the only available interpretation is that of a single shared possessum, which results from the coordination of the possessors. The other reading, with two possessors and possessa, and the first possessum elided, is not available. Accordingly, ellipsis of the head nominal is impossible.

(68) a. [DP *Iːβɐn* [NP *rɯt*]] *pɐːnə* [DP *Mɐːʃɐ* [NP *rɯt*]]
Ivan boat and Masha boat
'Ivan's boat and Masha's boat (=two boats)'

b. [DP *Iːβɐn pɐːnə Mɐːʃɐ* [NP *rɯt*]]
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, when coordinated nominals are accompanied by modifiers – e.g., demonstratives or adjectives, as in (69) – neither head noun may be omitted, which also attests to the impossibility of head noun ellipsis.

(69) *Miːʃɐ* [DP ***teːm neːβi*** [NP *(***βeːɬi***)]] *mɯβə* [DP ***tom pəɣtə*** [NP *(***βeːɬi***)]]
Misha this white deer or that black deer
βɛɬ-təɣ.
kill-PST.3SG<SG

[20] These constructions are currently undergoing further change: younger speakers allow (66) to be interpreted as resulting from ellipsis of one of the postpositions.

‘Misha killed this black deer or that white one.’

In sum, Khanty has developed true phrasal coordination but does not allow for ellipsis of nominal or postpositional heads.

4.5 Other kinds of coordination

Certain types of coordinate structures – notably, coordination of ‘unlikes’ and coordination that involves ellipsis, like 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. Variable acceptability of gapping and stripping suggests that ellipsis of the verbal head is dispreferred.

First, let us consider coordination of ‘unlikes’. Coordination requires for the conjuncts to be semantically and/or syntactically parallel (Williams, 1981), but certain ‘unlikes’ can be felicitously coordinated, as in (70). Accounting for this is possible if coordination is taken to result from coordination of phrases of the same type, followed by conjunction reduction (Beavers and Sag, 2004; Chaves, 2006):

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

In the Khanty equivalent, both verbs must be overt, suggesting that conjunction reduction is not an option (semi-copulas are used because present-tense copulas are null):

(71) *Peːtʲɐ* – [[*pioner-ɣə* *(*jəɣ*)] ***pɐːnə*** [*tʲuːt-əɬ-nɐt* *jarəŋ-kə* *(*jəɣ*)]].
Petja pioneer-TRNS become.PST.3SG and DEM-3SG-with proud-TRNS become.PST.3SG
‘Petja became a pioneer and became proud of that.’

Next, consider gapping, a kind of ellipsis that targets the iterated verb in coordinated clauses (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, and the other one may be an object or adjunct (Johnson 1996, Winkler 2005). (72) illustrates forward gapping, with the ‘gapped’ verb in the second conjunct. Many verb-final languages also allow for backward gapping, with the ellipsis site in the first clause.

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

In Khanty, the felicity of forward gapping varies: some speakers accept it, while others interpret the lack of an overt verb in the second conjunct as a null copula, as in (73a), which shows that ellipsis of the verb is unavailable. Younger speakers accept forward gapping more readily than older ones. An overt coordinator is preferred; the choice of *oːs* as a coordinator indicates juxtaposition between the two clauses and is similar to the Russian particle/complementizer *a* ‘whereas’. Only a non-reduced utterance, (73b), is accepted by all speakers.

(73) a. *%Miːχɐ* *sɒːrt* *qɐːtəɬ,* (***oːs***) *ʃɒːntɯ* – *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. *Miːχɐ* *sɒːrt* *qɐːtəɬ,* ***pɐːnə*** *ʃɒːntɯ* *jɛβ* *qɐːtəɬ.*
Mikha pike catch.PST.3SG and Shonty perch catch.PST.3SG
‘Mikha caught a pike, and Shonty caught a perch.’

Backward gapping/RNR, in contrast, is more felicitous, though an overt verb in the first conjunct is still preferred. There is also a preference (but not a requirement; 74a) for an overt conjunction (74b).

(74) a. ?*Mi:ʃɐ sɒ:rt, Sɐ:ʃɐ jɛβ qɐ:təɬ.*
Misha pike, Sasha perch catch.PST.3SG
'Misha caught a pike, and Sasha caught a perch.'

b. *Mi:ʃɐ sɒ:rt qɒ:təɬ,* ***(pɐ:nə/ o:s)*** *Sɐ:ʃɐ jɛβ qɐ:təɬ.*
Misha pike catch.PST.3SG, and Sasha perch catch.PST.3SG
'Misha caught a pike, and Sasha caught a perch.'

Finally, in stripping, all constituents in the second clause are deleted under identity with the first one, except for one, which may be accompanied by an adverb (*perhaps, as well, too*) or negation (Ross 1969, Hankamer & Sag 1976). Stripping is not fully felicitous in Khanty. Stripping with the subject as the remnant (VP ellipsis) sounds somewhat colloquial to some speakers, but is hard to interpret to others. The optimal version spells out the second conjunct in full; an overt conjunction is preferred (though not required).

(75) a. ??*Mi:ʃɐ sɒ:rt qɐ:təɬ, Sɐ:ʃɐ ətʲə.*
Misha pike catch.PST.3SG, Sasha too
'Misha caught a pike, Sasha too.'

b. *Mi:ʃɐ sɒ:rt qɐ:təɬ,* ***(pɐ:nə)*** *Sɐ:ʃɐ ətʲə sɒ:rt qɐ:təɬ.*
Misha pike catch.PST.3SG, and Sasha too pike catch.PST.3SG
'Misha caught a pike (and) Sasha caught a pike too.'

Stripping with an object remnant is even more restricted: here, only two fully spelled out clauses, coordinated with an overt *pɐ:nə*, are judged as fully felicitous. Note that an overt verb in the second conjunct improves acceptability.

(76) a. **Mi:ʃɐ sɒ:rt qɐ:təɬ* ***pɐ:nə*** *jɛβ ətʲə.*
Misha pike catch.PST.3SG and perch too
('Misha caught a pike and a perch, too.')

b. ?*Mi:ʃɐ sɒ:rt qɐ:təɬ* ***pɐ:nə*** *jɛβ ətʲə qɐ:təɬ.*
Misha pike catch.PST.3SG and perch too catch.PST.3SG
'Misha caught a pike and caught a perch, too.'

c. *Mi:ʃɐ sɒ:rt qɐ:təɬ* ***pɐ:nə*** *ɬɯβ jɛβ ətʲə qɐ:təɬ.*
Misha pike catch.PST.3SG and 3SG perch too catch.PST.3SG
'Misha caught a pike and he caught a perch, too.'

To recap, unavailability of coordination of 'unlikes' provides evidence against conjunction reduction. Restricted use of gapping and stripping speaks against widespread ellipsis of the verb.

4.6 Co-compounding

Widespread use of overt conjunctions in phrasal coordination in Khanty did not obliterate the use of co-compounds. Nouns as well as adjectives, numerals, and verbs, can still form co-compounds. In terms of meaning, co-compounds in most contexts can be used interchangeably with overtly coordinated constituents:

(77) a. *Mɐ:ʃɐ-****ɣən*** *(*pɐ:nə) Mi:ʃɐ-****ɣən*** *ɐ:rəɣ-łəɣən.*
Masha-DU and Misha-DU sing-PRS.3DU
'Masha and Misha sing (in general)/are singing now.'

b. *Mɐːʃɐ*(***-ɣən**) *pɐːnə* *Miːʃɐ*(***-ɣən**) *ɐːrəɣ-łəɣən.*
Masha-DU and Misha-DU sing-PRS.3DU
'Masha and Misha sing (in general)/are singing now.'

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

(78) *Mɐːʃɐ-**ɣən*** *Peːtʲɐ-**ɣən*** *kiːt-ɣə* *mən-ɣən.*
Masha-DU Petja-DU two-TRNS go-PST.3DU
'Masha and Petja got divorced.'

There are some differences between co-compounding and overt coordination, though: e.g., they 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):

(79) a. *Miːʃɐ* *spːrt* ***mɯβə*** *jɛβ* *qpːłəmt-əɣ?* (alternative question)
Misha pike or perch catch-PST.3SG
'Did Misha catch a pike or a perch?'

b. *Miːʃɐ* *spːrt-moqsəŋ* *qpːłəmt-əɣ?* (yes/no-question)
Misha pike-muksun catch-PST.3SG
'Did Misha catch a pike or a muksun (i.e., did he catch some fish)?'

These facts are consistent with our consultants' intuition that co-compounds cannot be used to expresses disjunction of the *either… or* type, where only one of the disjuncts can be true at a time. This is compatible with the structural analysis of co-compounds sketched out in Section 2.3: co-compounds share the functional projections that combine the two elements, 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 with an additive meaning came to be used as coordinators, while still retaining their adverbial function in other contexts. 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. Third, we have shown that the use of ellipsis in Khanty is quite restricted (marginal with respect to coordinated verbs and prohibited with respect to coordinated nominal and postpositional heads). 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 between 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 that this gives rise to is 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 Maremjanin's text from 1936, which represents an informal but not folklore-like register. The prevalence of this much repetitiveness may seem striking and uneconomical. 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 [NP_1 NP_2 V] string can be analyzed in multiple ways:

(80) [NP_1 NP_2 V]
- (i) NP_1 = subject, NP_2 = object;
- (ii) NP_1 = possessor, NP_2 = subject;
- (iii) NP_1 = possessor, NP_2 = object (subject = *pro*).

The grammatical functions of the NPs are 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 NP_1 or NP_2 (or both) in the [NP_1 NP_2 V] sentence by an (asyndetically) coordinated expression would have extended the string of juxtaposed un-casemarked NPs, thereby multiplying the interpretive options:

(99′) [NP_1 NP_2 V] (with asyndetic phrasal coordination allowed)
- (iv) NP_1, NP_2 = subjects;
- (v) NP_1, NP_2 = objects (subject = *pro*).

In the case of [NP_1 NP_2 NP_3 V], the possibilities further 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 to ensure processing efficiency, as defined by Hawkins (2004).

Hawkins (2004) argued that processing efficiency is governed by two rules: ‘Express the most with the least’ and ‘Express it earliest’. Based on the latter, he proposed the principle of Maximize On-line Processing, provided in (81). ‘On-line Property to Ultimate Property ratios’ 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). The intuition behind (81) 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).

(81) **Maximize On-line Processing** (Hawkins 2004:510)
> 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:

(82) *Unassignment factors* (Hawkins 2004:52)
   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).

(83) *Misassignment factors (selected)* (Hawkins 2004:53)
   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);
   c. 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 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 e.g., coordination. We propose that phrasal coordination in earliest attested Khanty was unavailable due to (i) the absence of overt conjunctions, (ii) the lack of overt case on subjects, objects, and possessors, 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.

Specifically, in traditional Khanty, the structural relations/ grammatical functions of the NPs in an [NP_1 NP_2 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 (Sheehan et al. 2017). 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 the 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, in most cases, via the coordination of phrase-size constituents, but may also result from conjunction reduction (granted that the restrictions on ellipsis of nouns, postpositions, and verbs

are obeyed). 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 stringed 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.

Funding information
This research was supported by grants KKP 129921 and K 135958 of NKFIH, the Hungarian National Research, Development and Innovation Office, which we gratefully acknowledge.

Acknowledgements
We owe thanks to two anonymous reviewers and the editor of the *Journal of Historical Linguistics*, as well as to our colleagues in the Uralic syntax research group of the Hungarian Research Centre for Linguistics, especially Márta Csepregi, Éva Dékány, and Katalin Gugán for their useful comments. Special thanks are due to our Khanty consultants.

References

Beavers, John & Ivan A. Sag. 2004. Coordinate ellipsis and apparent non-constituent coordination. *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar* ed. by Stefan Müller, 48-69. Stanford: CSLI Publications.

Bogoras, Waldemar. 1922. Chukchee. *Handbook of American Indian Languages, Part 2*. (=*Bureau of American Ethnology Bulletin* 40) ed. by Franz Boas, 631-903. Washington: Government Printing Office.

Chafe, Wallace. 1985. Linguistic differences produced by differences between speaking and writing. *Literacy, Language, and Learning* ed. by David R. Olson, Nancy Torrance & Angela Hildyard, 105-123. Cambridge: Cambridge University Press.

Chafe, Wallace. 1987. Cognitive constraints on information flow. *Coherence and Grounding in Discourse* ed. by Russell Tomlin, 21-51. Amsterdam: John Benjamins.

Chaves, Rui P. 2006. Coordination of unlikes without unlike categories. *Proceedings of the 13th International Conference on Head-driven Phrase Structure Grammar* ed. by Stefan Müller, 102-122. Stanford: CSLI Publications.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130:1.33-49.

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 Jacaltec*. Austin: University of Texas Press.

Csepregi, Márta. 1998. Szurguti osztják chrestomatia. Szeged: JATE.

Csepregi, Márta. 2002. Texte in chantischer Sprache vom Fluss Agan. *Sei gegrüsst! Beiträge zur Finnougristik zu Ehren von Gert Sauer dargebracht zu seinem siebzigsten Geburtstag* (=*Veröffentlichungen der Societas Uralo-Altaica 57*) ed. by Eugene Helimski & Anna Widmer, 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, George. 1931. *Syntax*. Boston: Heath.

DEWOS = Steinitz, Wolfgang, ed. 1966-1993.

Di Sciullo, Anna Maria. 2002. *Asymmetry in Grammar*. Amsterdam & Philadelphia: John Benjamins.
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. 2021. Definiteness effect in the PP. *Linguistic Inquiry* OnlineEarly.
É. Kiss, Katalin & Orsolya Tánczos. 2018. From possessor agreement to object marking in the evolution of the Udmurt *-jez* suffix: A grammaticalization approach to morpheme syncretism. *Language* 94:4.733-757.
Filchenko, Andrey Yury. 2010. *Aspects of the Grammar of Eastern Khanty*. Tomsk: TSPU-Press.
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.
Gulya, János. 1966. *Eastern Ostyak Chrestomaty*. Bloomington: Indiana University Publication.
Haiman, John. 1983. Iconic and economic motivation. *Language* 59:4.781-819.
Haiman, John. 1985. *Natural Syntax. Iconicity and Erosion* (=*Cambridge Studies in Linguistics Vol. 44*). Cambridge: Cambridge University Press.
Halle, Morris & Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. *The View from Building 20: Essays in linguistics in honor of Sylvain Bromberger* ed. by Kenneth Hale & S. Jay Keyser, 111-176. Cambridge MA: MIT Press.
Han, Chung-hye & Maribel Romero. 2004. Disjunction, focus, and scope. *Linguistic Inquiry* 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* 13:3.201-270.
Jing-Schmidt, Zhuo & Xinjia Peng. 2015. The emergence of disjunction: A history of constructionalization in Chinese. *Cognitive Linguistics* 27:1.101-136. https://doi.org/10.1515/cog-2015-0073
Johannessen, Janne Bondi. 1996. Partial agreement and coordination. *Linguistic Inquiry* 27:4.661-676.
Johnson, Kyle. 1996. In search of the middle field. Ms. Amherst: University of Massachusetts. http://people.umass.edu/kbj/homepage/index_johnson.htm.
Kálmán, Béla. 1976. *Chrestomathia Vogulica*. Budapest: Tankönyvkiadó.
Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
Lakoff, George & Stanley Peters. 1966. Phrasal conjuntion and symmetric predicates. *Modern Studies in English* ed. by David A. Reibel & Sanford A. Schane, 113-142. Englewood Cliffs: Prentice-Hall.
Lewy, Ernst. 1911. *Zur finno-ugrischen Wort- und Satzverbindung*. Göttingen: Vandenhoeck.
Mauri, Caterina. 2008 . The irreality of alternatives: Toward a typology of disjunction. *Studies in Language* 32:1.22-55.
Mikola, Tibor. 1973-74. Ugor ikerszavak [Ugric twin words]. *Néprajz és Nyelvtud*omány 17-18.54-62.
Mikola, Tibor. 1985. Ikerszók és parallelizmusok az obi-ugor nyelvekben [Twin words and parallelisms in the Ob-Ugric languages]. *Néprajz és Nyelvtud*omány 29-30.143-149.

Mithun, Marianne. 1988. The grammaticalization of coordination. *Clause Combining in Grammar and Discourse* ed. by John Haiman & Sandra A. Thompson, 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* 9:1.1-26.
Munn, Alan Boag. 1993. Topics in the syntax and semantics of coordinate structures. Doctoral dissertation, University of Maryland at College Park.
Nikolaeva, Irina. 1999. *Ostyak*. München: Lincom Europa.
Nikolaeva, Irina. 2003. Possessive affixes in the pragmatic structuring of the utterance: evidence from Uralic. *International Symposium on Deictic Systems and Quantification in Languages Spoken in Europe and North and Central Asia. Collection of Papers* ed. by Pirkko Suihkonen & Bernard Comrie, 130-145. Izhevsk & Leipzig: Udmurt State University & Max Planck Institute for Evolutionary Anthropology.
Ohori, Toshio. 2004. Coordination in Mentalese. *Coordinating Constructions* ed. by Martin Haspelmath, 41-66. Amsterdam/Philadelphia: John Benjamins.
OUDB 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. Doctoral dissertation, MIT.
Pruitt, Kathryn & Floris Roelofsen. 2013. The interpretation of prosody in disjunctive questions. *Linguistic Inquiry* 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: Vandenhoech & 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. 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. *Progress in Linguistics* ed. by Manfred Bierwisch & K. E. Heidolph, 249-259. The Hague: Mouton.
Schwarz, B. 1999. On the syntax of *either... or*. *Natural Language and Linguistic Theory* 17:2.339-370.
Sheehan, Michelle, Theresa Biberauer, Ian Roberts & Anders Holmberg. 2017. *The Final-Over-Final Condition: A Syntactic Universal*. Cambridge, MA: MIT Press.
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. Doctoral dissertation, University of Helsinki.
Stassen, Leon. 2003 Noun phrase conjunction: The coordinative and the comitative strategy, *Noun Phrase Structure in the Languages of Europe* ed. by Frans Plank, 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. Osztják parallelizmusok és mellérendelő összetételek [Ostyak parallelisms and coordinate compounds]. *Nyelvtudományi Közlemények* 91.221-225.
Terjoshkin, N. I. 1959. *Bukvar*. Leningrad: Gosudarstvennoje Uchebno-Pedagogicheskoje Izdatelstvo.
Terrill, Angela. 2004. Coordination in Lavukaleve. *Coordinating Constructions* ed. by Martin Haspelmath, 427-444. Amsterdam & Philadelphia: John Benjamins.
Van Valin, Robert Jr. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: Cambridge University Press.
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*). Helsinki: Suomalais-Ugrilainen Seura.
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. *Minimalism and Kayne's Antisymmetry Hypothesis* (=*Groningen Arbeiten Zur Germanistischen Linguistik 37*), ed. by Jan-Wouter Zwart, 291–331. Groningen: University of Groningen.
Wilder, Christopher. 2018. Conjunction Reduction and Right-Node Raising. *The Oxford Handbook of Ellipsis* ed. by Jeroen van Craenenbroeck & Tanja Temmerman. DOI: 10.1093/oxfordhb/9780198712398.013.2
Williams, Edwin S. 1981. Transformationless grammar. *Linguistic Inquiry* 12:4.645–653.
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. *Linguistics in the Netherlands 2005* ed. by Jenny Doetjes & Jeroen van de Weijer, 232–241. Amsterdam: John Benjamins.