

# Word order constraints in the Tundra Nenets wh-question

Nikolett Mus

Hungarian Research Institute for Linguistics

Siberian Life 2021, Hamburg, 4–5 February 2021

## 1 Background

### 1.1 Initial considerations

- In Siberia, there are over 30 indigenous languages (belonging to 8 language families; Pakendorf, 2010).
- Nowadays, the majority of these languages are threatened or even moribund.
- Nevertheless, a number of speakers still remain.  
⇒ This situation urges the need for documenting these languages.
- Documenting:
  - collecting and archiving, i.e. processing, (written and spoken) data as far as possible
  - making the collected data available for the speaker and the researcher community
  - describing these languages in such a way that makes it possible to reconstruct their structure
- The present paper addresses different data gathering techniques, e.g. using corpus data, or syntactic elicitation test, and different approaches, e.g. a corpus-based approach, a typological approach, a formal syntactic approach.
- It will be shown, that the systematic combination of different methods and approaches leads us to a more complex picture.
- Tundra Nenets (Samoyedic, Uralic) wh-questions will be target constructions in the study.

## 1.2 The Tundra Nenets language

- Tundra Nenets belongs to the Samoyedic branch of the Uralic language family.
- It is spoken by c. 20,000 people.
- Its status is 6b, i.e. *threatened*, on the EGIDS scale.



## 2 Wh-questions in the literature

- In Tundra Nenets genuine single wh-questions, the wh-phrase usually remains *in situ* (1)–(2).

(1) *xibá Irina-m? meńe?*  
 who Irina-ACC love.3SG  
 ‘Who loves Irina?’ [Subject, *in situ*]

(2) *Sergei xibá-m? meńe?*  
 Sergei who-ACC love.3SG  
 ‘Whom does Sergei love?’ [Object, *in situ*]

- Other configurations are also observed (3)–(4).

(3) *Irina-m? xibá meńe?*  
 Irina-ACC who love.3SG  
 ‘Who goes to school today?’ [Subject, *ex situ*]

(4) *xibá-m? Sergei meńe?*  
 who-ACC Sergei love.3SG  
 ‘Whom does Sergei love?’ [Object, *ex situ*]

- The literature suggests, that the different orders may represent different discourse-pragmatic interpretation, but it is not necessarily the case. Thus, the *in situ* and *ex situ* (or initial, middle, preverbal) positions of a wh-phrase reflect free variations (Nikolaeva, 2014).
- It is also implied that the wh-phrase does not have a dedicated syntactic position (Nikolaeva, 2014).
- These implications, however, does not explain the ungrammaticality of the configuration in (5), in which the ‘only’-focus cannot precede the *wh-in situ*.

(5) \**Masha-ri* *ɲamge-mʔ* *ɲawor-ɲa?*  
 Masha-LIM what-ACC eat-CO.3SG  
 ‘What does only Masha eat?’ [Object, \**in situ*]

(6) *ɲamge-mʔ* *Masha-ri* *ɲawor-ɲa?*  
 what-ACC Masha-LIM eat-CO.3SG  
 ‘What does only Masha eat?’ [Object, *ex situ*]

[Q1:] Where can (not) appear the wh-phrase in single wh-questions?

[Q2:] What syntactic constraints can we formulate in single wh-questions in Tundra Nenets?

- Only those configurations will be discussed here that contains the following constituents: S, O, V; and either of these constitutes is a wh-phrase.
- Focus on constructions in which the position of the wh-phrase is invariant.

### 3 Approaches, methods and results

#### 3.1 A corpus-based approach (Mus, 2015)

- Tundra Nenets Monolingual Corpus (TNMC) consisting of published and/or electronically accessible texts.<sup>1</sup>
- The texts in the corpus dominantly represent the written version of the Tundra Nenets language.
- The current size of TNMC (as of 2021) is 452,930 tokens.
- The number of attested wh-questions is 143.

Word order	Position of wh-phrase	Nr
SOV	<i>in situ</i>	25
OSV	<i>ex situ</i>	14
SOV	<i>in situ</i>	91
OSV	<i>ex situ</i>	13

<sup>1</sup>The corpus is available on the following website <https://tundranenetsdata.nytud.hu/bonito>

- In OSV, the agreeing 3rd person object (possibly topical) occupies the sentence initial position preceding the Subject wh-phrase in 13 cases out of 14 clauses.
- In OSV, the object is expressed by a complex interrogative phrase in 9 examples out of the 13 clauses.
- In OSV, the subject preceded by the (complex) interrogative phrase is a pronominal (and possibly focussed) one in 9 clauses out of the 13 occurrences.

**Generalization[1]:** SOV is the most frequent interrogative order attested in the corpus. *Ex situ* position is possibly due to discourse-pragmatic factors.

### 3.2 (Functional) Typological approach to word order

- Consultations have been undertaken with a native speaker of Tundra Nenets (Khadry Okotetto, born in 1991) in Moscow in 2017, 2019, and online in 2020.
- Visual stimuli, i.e. short cartoons (drawn by Barbara Egedi) were provided to the speaker. The answers were recorded both in audio and in writing.
- What to test: all the typologically possible permutations of S, O, and V.

#### 1 Questioning the subject



- The expected answer: ‘Who gave/brought the cake?’  
(Note that the object, i.e. ‘the cake’ is expected to be marked information structurally: it can be interpreted as a topic, or a contrastive focus.)

- Results of the test:

Word order	Position of wh-phrase	Syntactic explanation
OSV (7)	<i>ex situ</i>	presumably the object is topical
SOV	<i>in situ</i>	presumably the object is a contrastive focus
SVO	n.i.	postposed topical object
*OVS	postverbal	wh-restriction
*VSO	postverbal	wh-restriction
*VOS	postverbal	wh-restriction

(7) *torta-mʔ xiba ta-sa(-da)?*  
 cake-ACC who give-INTERR(-SG.3SG)  
 ‘Who gave a/the cake?’ [Subject, *ex situ*]

(8) *\*torta-mʔ ta-sa(-da), xiba?*  
 cake-ACC give-INTERR(-SG.3SG) who  
 intended ‘Who gave a/the cake?’ [Subject, *postverbal*]

## 2 Questioning the object



- The expected answer: ‘What did the rabbit give/bring?’

(Note that the subject, i.e. ‘the rabbit’ is expected to be marked information structurally: it is either interpreted as a topic, or a contrastive focus.

- Results of the test:

Word order	Position of wh-phrase	Syntactic explanation
SOV (10)	<i>in situ</i>	presumably the subject is topical
OSV	<i>ex situ</i>	presumably the wh-phrase is emphasized
	<i>ex situ</i>	presumably the subject is a contrastive focus
OVS	n.i.	postponed topical subject
*SVO	postverbal	wh-restriction
*VSO	postverbal	wh-restriction
*VOS	postverbal	wh-restriction

- (9) *ńawa-ko-r*                      ***ŋamge-m*** *ta-sa?*  
 rabbit-DIM-2SG.POSS what-ACC give-INTERR.3SG  
 ‘What did the rabbit (lit. your rabbit) give?’ [Object, *in situ*]
- (10) \**ńawa-ko-r*                      *ta-sa,*                      ***ŋamge-m?***  
 rabbit-DIM-2SG.POSS give-INTERR.3SG what-ACC  
 intended: ‘What did the rabbit (lit. your rabbit) give?’ [Object, \**postverbal*]

(Note that VX is otherwise possible in Tundra Nenets.)

⇒ Even the neutral interpretation of the clause is ruled out, the test was not appropriate to differentiate between the *in situ* and *ex situ* positions.

**Generalization[1]** still holds for these data.

**Generalization[2]:** the wh-phrase cannot appear after the finite verb (\*[Vwh]).

- (11) Q: What does the man read?

A: \**Xasawa tolabi,* *kniga-m?*  
 man read.3sg book-acc  
 intended meaning: ‘The man reads THE BOOK.’

**Generalization[2’]:** the information focus cannot appear after the finite verb (\*[Vfoc]).

### 3.3 A formal syntactic approach

- Questionnaires containing (i) target language manipulation elicitation, and (ii) (fieldworker-driven) paradigmatic substitution elicitation tasks.
- What to test: intervention effects, i.e. the order of expressions having scope and the wh-phrase.
- Results of the test:

#### 1 The ‘only’ focus

- (12) \**shkola-ři-xina xība jewej-m? ŋǎwor-ca?*  
 school-FOC-LOC who soup-ACC eat-INTERR.3SG  
 ‘Who ate the soup ONLY IN THE SCHOOL?’ [Subject, \**in situ*]
- (13) ??*Jewej-ři-m? xība ŋǎwor-ca?*  
 soup-FOC-ACC who eat-INTERR.3SG  
 intended: ‘Who ate ONLY THE SOUP?’ [Subject, \**ex situ*]
- (14) \**Igor-ři ŋamge-m? ŋǎwor-ca?*  
 Igor-FOC what-ACC eat-INTERR.3SG  
 intended: ‘What did ONLY IGOR eat?’ [Object, \**in situ*]
- (15) \**tuku jǎlja-ři xība-m? Pavel mǎne?-ca?*  
 this day-FOC who-ACC Pavel see-INTERR.3SG  
 intended: ‘Whom did Pavel see ONLY TODAY?’ [Object, \**ex situ*]

#### 2 The ‘even’ focus

- (16) ?/\**weńeko-xowa-m? xība mǎńije?*  
 dog-AFF-ACC who see.3SG  
 intended: ‘Who sees EVEN THE DOG?’ [Object, \**ex situ*]

#### 3 Negative polarity item

- (17) \**xība-xǎrt xurka jewej-m? ní-sa ŋawor-ʔ?*  
 who-FOC what.kind soup-ACC NEG.AUX-INTERR.3SG eat-CNG  
 intended: ‘What kind of soup did not eat anyone?’ [Object, \**in situ*]

- **Generalization[3]:** focus cannot precede the wh-phrase in genuine content questions (\*[foc...wh]).

#### 4 Universal quantifier

- (18) \**xusuwej ně-m? xība mane?-ća?*  
 every woman-ACC who see-INTERR.3SG  
 intended: ‘Who saw every woman?’ [Subject, \**ex situ*]

- **Generalization[3’]:** expressions that have scope cannot precede the wh-phrase in genuine content questions (\*[op...wh]).

(Note that **Generalization[1]** is not relevant here.)

- In many of the above cases, the *ex situ* is motivated, i.e. to rescue the construction from intervention effect.

## 4 Conclusions

- To sum up the main points of the talk:

**Generalization[1]:** SOV is the most frequent interrogative order attested in the corpus. *Ex situ* position is possibly due to discourse-pragmatic factors.

⇒ according to the corpus-based approach

**Generalization[2']:** the information focus cannot appear after the finite verb (\*[Vfoc]).

⇒ according to the (functional) typological approach

**Generalization[3']:** expressions that have scope cannot precede the wh-phrase in genuine content questions (\*[op...wh]).

⇒ according to the formal syntactic approach

## Acknowledgments

The support of the research project “Theoretical and experimental approaches to dialectal variation and contact-induced change: a case study of Tundra Nenets” (FK\_129235) is gratefully acknowledged.