

KAIS ALLKIVI (Tallinn)

PRONOMINAL DOUBLING IN ESTONIAN COMPLEX *WH*-QUESTIONS

Abstract. In recent years, the use of pronoun-doubling constructions as alternatives to standard long-distance *wh*-questions, where the *wh*-phrase is spelled out only in the matrix clause, has received substantial attention in research of syntactic variation. However, the doubling phenomenon has scarcely been studied in the Uralic languages. This paper concentrates on Estonian complex, i.e. bi-clausal, *wh*-questions that contain the bridge verb *arvama* 'think' or *üttelema* 'say' and where the subject or object, either animate or inanimate, is being questioned. An acceptability judgement test and corpus analysis were applied to determine which pronominal patterns and to what extent are attested in such interrogative sentences. Both identical and non-identical doubling appear to be common in Estonian, although identical doubling is restricted to inanimate subject/object questions (the pronoun *mis* 'what' introduces both clauses), while non-identical doubling is preferred if a person is questioned (the matrix clause is introduced by the pronoun *mis* 'what' and the subordinate clause by the pronoun *kes* 'who'). Proposing a tentative syntactic analysis for the different doubling patterns, I argue that partial *wh*-movement involving two independent *wh*-chains is generally favoured and it also underlies identical doubling. Long-distance *wh*-movement, on the other hand, turns out not to be freely permitted in Estonian.

Keywords: Estonian, syntactic variation, syntactic doubling, long-distance *wh*-movement, partial *wh*-movement, indirect dependency approach, interrogative pronouns.

Introduction

Wh-fronting in content questions is common to most European languages, including the Finno-Ugric languages such as Estonian (see Metslang 1981), Finnish (see Vainikka 1989; Huhmarniemi 2012), Hungarian (see Toft 2001), Udmurt (see Грамматика современного удмуртского языка 1970), Komi and the Saami languages (see The Uralic Languages 1998). Example (1a) is an Estonian neutral declarative sentence with the SVO word order, whereas in (1b) the interrogative pronoun questioning the object of the verb occupies the left periphery and leaves behind an empty position.

- (1) a. *Liisa loeb ajalehte*
 Liisa-NOM read-3SG newspaper-PART
 'Liisa is reading a newspaper'

- (1) b. *Mida Liisa loeb mida?*
 ↑—————|
 what-PART Liisa-NOM read-3SG
 'What does Liisa read?'

In contrast, there are languages where *wh*-phrases mostly remain *in situ*. One of such Uralic languages is Tundra Nenets (see Salminen 2012), which has an SOV neutral word order. The *wh*-pronoun functioning as the object thus follows the subject, as in (2), although complex and topicalized interrogative phrases may behave differently (Mus 2015 : 130–134, 153–160).

- (2) *Pida ŋamge-m? xeta?* (Ненянг 2005 : 48)
 3SG what-ACC say-PST-3SG
 'What did he say?'

The movement of the *wh*-phrase to the front of the clause is a case of movement to a non-argument position, which has been referred to as "A'-movement" in the transformational theories of syntax, of which the minimalist program is the most recent (see Chomsky 1995; 2000). The *wh*-phrase may also move out of the embedded clause to the front of the matrix clause, passing a possible final landing site at the lower clause-initial position. Such movement, often called "long-distance *wh*-movement", is illustrated in (3) and (4) for Hungarian and (5) for Finnish. In Hungarian, the distribution of long-distance *wh*-movement is subject to dialectal variation (Marác 1991 : 153–154). However, if it is considered acceptable by speakers, then subjects, objects as well as adverbials may each move out of the CP within which they originate (Toft 2001 : 194–195). Finnish, on the other hand, only allows long-distance object and adjunct extraction, while subject extraction may be possible in colloquial speech to a limited extent (Huhmarniemi 2012 : 96–98).

- (3) [*Kit gondolsz, [hogy ~~kit~~ látta Jánost ~~kit~~]*?]
 ↑—————| ↑—————|
 who-ACC think-2SG that see-PST-3SG János-ACC
 'Who do you think saw Janos?' (Toft 2001 : 195)

- (4) [*Kinek hallottad, [hogy ~~kinek~~ János kölcsönadott 2000 dollárt ~~kinek~~]*?]
 ↑—————| ↑—————|
 whom-DAT hear-PST-2SG that János-NOM loan-PST-3SG 2000 dollar-ACC
 'To whom did you hear János had loaned 2000 dollars?' (Horvath 2007 : 115)

- (5) [*Kenet Pekka luuli [että ~~kenet~~ Merja oli tavannut ~~kenet~~]*?]
 ↑—————| ↑—————|
 who-ACC Pekka-NOM think-PST-3SG that Merja-NOM be-PST-3SG meet-PP
 'Who did Pekka think Merja had met?' (Huhmarniemi 2012 : 96)

In a number of languages, alternative strategies are used to form such long *wh*-questions, namely identical and non-identical pronominal doubling (see Fanselow 2006; Rett 2006; Schippers 2010a). In the former case, commonly referred to as "*wh*-copying", the higher and the lower clause are introduced by the same interrogative pronoun, as exemplified for German in (6). In the latter case, referred to as "partial *wh*-movement" or "*wh*-scope marking", the two *wh*-

pronouns differ in phonological shape, as can be seen in the Hungarian example (7). The matrix clause is then introduced by a pronoun equivalent to *what* (or *how*, as in Polish; see Lubanska 2004), which provides no evident semantic content; its sole function appears to be to extend the scope of the lower *wh*-phrase, i.e. to signal that it is a direct question (Dayal 1994 : 138). The complementizer may either become silent in the doubling constructions, like in German (normally, *dass* would introduce the embedded CP), or it may remain in its position at the front of the CP, like *hogy* in Hungarian.

- (6) *W e n g l a u b s t d u w e n s i e l i e b t ?*
 who-ACC believe-2SG you-NOM who-ACC she-NOM love-3SG
 'Who do you think she loves?' (Pankau 2013 : 1)
- (7) *M i t g o n d o l s z , h o g y k i t l á t o t t J á n o s ?*
 what-ACC think-2SG that who-ACC see-PST-3SG János-NOM
 'Who do you think János saw?' (Horvath 1997 : 510)

Long-distance dependencies and syntactic doubling have hitherto not been studied in Estonian, even though both identical and non-identical doubling are used in long *wh*-questions (likewise, e.g. Romani, Frisian and some varieties of German and Dutch have been listed among such languages). There is, however, an interesting asymmetry.

Identical doubling rather occurs if an inanimate object is questioned, as in (8), where the matrix clause and the subordinate clause are mutually introduced by the interrogative pronoun *mis* 'what'. If a person is questioned, non-identical doubling is preferred, as shown in (9), where the pronoun *kes* 'who' referring to a human remains in the subordinate clause, while the *wh*-scope marker in the matrix clause is *mis* – the least specified *wh*-phrase in Estonian. Whether these superficially different cases of pronoun-doubling actually entail different syntactic structures or not, is a question in itself, which is also tackled in this paper.

- (8) *M i s s a a r v a d , m i s j u h t u s ?*
 what-NOM you-NOM think-2SG what-NOM happen-PST-3SG
 'What do you think happened?'
- (9) *M i s s a a r v a d , k e s s e d a t e g i ?*
 what-NOM you-NOM think-2SG who-NOM it-PART do-PST-3SG
 'Who do you think did it?'

In case of regular long-distance *wh*-movement, the *wh*-phrase moving successive-cyclically to the higher left periphery is spelled out once. The subordinate clause is then introduced by a declarative complementizer *et*, which is always obligatory. Note that in (10) *kes* surfaces in the partitive form *keda*, partitive being the default object-marking case in Estonian. In case of subject extraction, long-distance *wh*-movement seems to be limited similarly to Finnish.

- (10) *K e d a t a ü t t l e s , e t t a t u n n e b ?*
 who-PART (s)he-NOM say-PST-3SG that (s)he-NOM know-3SG
 'Who did (s)he say (s)he knew?'

Verbs that permit long *wh*-extraction from a finite CP have been referred to as "bridge verbs" (Felser 2004 : 549) and include epistemic verbs, such as *arvama* 'think', *ütlema* 'say', *väitma* 'claim', *uskuma* 'believe', *lootma* 'hope' and *otsustama* 'decide' in Estonian. Nevertheless, Featherston (2004 : 205) concludes, based on German, that the bridge feature should be seen as a continuum, not a categorical distinction, meaning that there is no absolute group of bridge verbs; language users perceive some verbs as more natural than others in long-distance *wh*-questions.

The purpose of the study presented in this article was to determine to what extent different pronominal patterns are attested in Estonian complex *wh*-questions, and in which way their acceptability is related to (a) the animacy/inanimacy of the object being questioned, (b) subject and object question, and (c) lexical factors. The object of study is restricted to biclausal *wh*-questions that contain the verb *arvama* 'think' or *ütlema* 'say' in the matrix clause, which presumably entail somewhat different preferences for the long *wh*-questions. This study concerns subject and partial object questions, thereby excluding extraction of total object and other arguments.

In order to test the acceptability of different pronominal patterns, I conducted a survey among 89 non-philologist native speakers of Estonian. In addition, I used corpus analysis to see what evidence large online corpora provide for long *wh*-questions in both standard written Estonian and rather informal written language use. The methodology is described in Section 2, the results are reported and discussed in Section 3. The following section gives an overview of different pronominal combinations in Estonian complex *wh*-questions and proposes a syntactic analysis for each pattern, which is a further important aim of this paper.

1. Possible pronominal patterns and their syntactic structure

1.1 Interrogative pronouns and nominative-partitive case alternation

The Estonian interrogative/relative pronoun *kes* 'who' primarily refers to a human object, whereas the pronoun *mis* 'what' refers to an inanimate object, action or situation. *Kes* can also refer to a group of people, an institution or an animal if they are considered an actor in the situation. If they are treated as an undergoer, *mis* is preferred. At the same time, the singular-plural opposition is neutralized. *Kes* and *mis* usually preserve their singular form even when the NP they refer to is in plural (Metslang 1981 : 63–72; Erelt, Erelt, Ross 2007 : 560–561).

Based on Dutch, Boef (2013) has shown that the patterns of pronominal doubling in long-distance dependencies are best accounted for by the feature specifications of the relevant pronouns, namely that the distribution of the pronouns depends on which features they can spell out. In Estonian, these features are semantic animacy (roughly human/non-human) and the type of the argument questioned (subject/object), the latter being derived from the case form of the pronoun.

When functioning as a subject, *kes* and *mis* constantly appear in the default nominative case form. In the partial object function, *kes* is always assigned the partitive case and surfaces as *ked*, e.g. *K e d a sa ootad?* 'Who are you waiting for?'. *Mis*, on the other hand, can preserve its nomi-

native form, so that both case forms are possible, e.g. *M i d a ~ M i s sa teed?* 'What are you doing?'; *M i d a ~ M i s sa poest tõi?* 'What did you bring back from shopping?'.¹

This kind of nominative-partitive case alternation also applies in complex *wh*-questions. This is what gives rise to the non-identical doubling patterns in long *wh*-questions with a non-human referent. Firstly, the matrix clause can either contain a nominative or a partitive *what*-phrase, independently of whether we have a subject or object question, whereas the subordinate clause can be introduced by the partitive form only if an object is questioned. Therefore, the identical pattern *mida-mida* (11a) and the non-identical pattern *mis—mida* (11b) cannot occur in a complex subject question.

- (11) a. *M i d a ta ütles, mida ta teha otsustas?*
 what-PART (s)he-NOM say-3SG what-PART (s)he-NOM to-do decide-PST-3SG
 'What did (s)he say (s)he decided to do?' [OBJ question]
- b. *M i s sa arvad, mida Laura Peetrile kinkis?*
 what-NOM you-NOM think-2SG what-PART Laura-NOM Peter-ALL give-PST-3SG
 'What do you think Laura gave to Peter?' [OBJ question]

The reverse *mida—mis* pattern and the identical *mis—mis* pattern are compatible with both subject (12a, 13a) and object questions (12b, 13b).

- (12) a. *M i d a nad ütlesid, mis probleeme tekitab?*
 what-PART they-NOM say-PST-3PL what-NOM problems-PART cause-3SG
 'What did they say causes problems?' [SUB] question]
- b. *M i d a sa arvad, mis vanaema seekord küpsetab?*
 what-PART you-NOM think-2SG what-NOM grandmother-NOM this time bake-3SG
 'What do you think grandmother is baking this time?' [OBJ question]
- (13) a. *M i s te arvate, mis see on?*
 what-NOM you-NOM think-2PL what-NOM it-NOM be-3SG
 'What do you think it is?' [SUB] question]
- b. *M i s sa arvad, mis nad vastasid?*
 what-NOM you-NOM think-2SG what-NOM they-NOM reply-PST-3PL
 'What do you think they replied?' [OBJ question]

In the non-doubling construction, *mis* can also be either nominative or partitive if extracted from an object position, as in (14).

- (14) *M i d a ~ M i s sa arvad, et ta ütles?*
 what-PART / what-NOM you-NOM think-2SG that (s)he-NOM say-PST-3SG
 'What do you think (s)he said?' [OBJ question]

¹ The partitive *mida* may demand a more specified answer, like in case of the question phrase *mida teha* 'what to do' that requires an answer in the *da*-infinitive form equivalent to *to do* in English. E.g. to the question *M i d a sul on plaanis teha?* 'What are you planning to do?' one has to reply *Lage värvida* 'To paint the ceiling', whereas one should not reply using an NP, such as *Lae värvimine* 'The painting of the ceiling' or *Töö* 'Work'. The question *M i s sul on plaanis teha?*, however, allows both types of answers (Metslang 1981 : 71).

The case of the *what*-phrase can have a similar alternation in the matrix clause of long *wh*-questions with a human referent, independently of the case and function of the pronoun *kes* in the lower clause. An example of a subject question is given in (15).

- (15) *Mida ~ Mis sa arvad, kes Harriga tantsis?*
 what-PART / what-NOM you think-2SG who Harry-COM dance-PST-3SG
 'Who do you think danced with Harry?'

Following Schoorlemmer (2009 : 126–128) and Boef (2013 : 48–49), I take the syntactic feature specifications to be dependent on the morphological realization of these features, meaning that the lack of a value (underspecification) for an attribute corresponds to a morphologically unrealized feature. The nominative form of the *wh*-pronoun *mis* 'what' can then be considered completely underspecified, for it is underspecified for definiteness, number, animacy as well as case (since the nominative case is morphologically unrealized). It can function both as a subject and an object, whereas the nominative form of *kes*, which is sensitive to animacy, can solely function as a subject. Partitive forms of the pronouns are specified for case and are compatible with objects only. The referential possibilities of the interrogative pronouns *mis* and *kes* and their partitive forms *mida* and *keda* are presented in Table 1.

Potential pronominal combinations in Estonian long-distance *wh*-questions are given in Table 2. These possible patterns are theory-neutrally categorized as non-doubling, identical doubling and non-identical doubling patterns, judging by the surface form of the pronouns and leaving syntactic interpretations aside.

Table 1

Nominal referents of the interrogative pronouns *kes*, *mis*, *keda*, *mida*

nominal referent	<i>kes</i> 'who'	<i>keda</i> 'who-PART'	<i>mis</i> 'what'	<i>mida</i> 'what-PART'
[SG, HUMAN, SUBJ] e.g. <i>poiss</i> 'boy-NOM'	+	–	–	–
[SG, HUMAN, OBJ] e.g. <i>poissi</i> 'boy-PART'	–	+	–	–
[SG, NON-HUMAN, SUBJ] e.g. <i>raamat</i> 'book-NOM'	–	–	+	–
[SG, NON-HUMAN, OBJ] e.g. <i>raamatut</i> 'book-PART'	–	–	+	+
[PL, HUMAN, SUBJ] e.g. <i>poisid</i> 'boys-NOM'	+	–	–	–
[PL, HUMAN, OBJ] e.g. <i>poisse</i> 'boys-PART'	–	+	–	–
[PL, NON-HUMAN, SUBJ] e.g. <i>raamatud</i> 'books-NOM'	–	–	+	–
[PL, NON-HUMAN, OBJ] e.g. <i>raamatuid</i> 'books-PART'	–	–	+	+

version of (19a), as exemplified by Fanselow (2006 : 451) for German; a similar analysis seems to apply for the Estonian sentences (20a) and (20b). According to Dayal (2000 : 187) and Fanselow (2006 : 451), the partial *wh*-movement construction (20b) can also be analyzed as a monosentential counterpart of sequential questions shown in (20c).

- (19) a. *Was denkst du?*
 what-ACC think-2SG you-NOM
 'What do you think?'
 b. *Was denkst du wer gekommen ist?*
 what-ACC think-2SG you-NOM who-NOM come-PP be-3SG
 'Who do you think has come?'
- (20) a. *Mis ~ Mida sa arvad?*
 what-NOM ~ what-PART you-NOM think-2SG
 'What do you think?'
 b. *Mis ~ Mida sa arvad, kes tuli?*
 what-NOM ~ what-PART you-NOM think-2SG who-NOM come-PST-3SG
 'Who do you think has come?'
 c. *Mis ~ Mida sa arvad? Kes tuli?*
 what-NOM ~ what-PART you-NOM think-2SG who-NOM come-PST-3SG
 'What do you think? Who has come?'

Another argument pro IDA is that a clausal correlative equivalent to 'it' can be found in Estonian declarative sentences, such as (21). Dayal (1994 : 149–151), based on Hindi, and Stepanov and Stateva (2006 : 2144–2115), based on Slavic languages, assume the *what*-phrase to be a similar correlative of the subordinate clause, with which it forms a constituent.

- (21) [*Ma arvan seda*, [_{CPi} *et Anne suudles Martinit*]]
 I-NOM think-1SG it-PART that Anne kiss-PST-3SG Martin-PART
 'I think that Anne kissed Martin'

The question remains whether the ambiguous patterns *mis—mis* and *mida—mida* represent partial *wh*-movement, or is *wh*-copying restricted by the animacy of the referent. In the former case, those patterns should occur and be accepted to a similar extent as the partial *wh*-movement pattern *mis—kes* and its variants. This would mean that the partial *wh*-movement construction could be used for both human and non-human referents, as well as both subject and object questions, leaving no obvious need for other constructions fulfilling the same purpose.

A tendency has been noted that long-distance *wh*-movement and partial *wh*-movement are generally in complementary distribution: The latter shows up in languages where the former is not freely permitted; either dominates the other. *Wh*-copying, on the other hand, co-occurs with long-distance *wh*-movement, compared to which it is always of a secondary nature (Stepanov, Stateva 2006; Schippers 2010a). If partial *wh*-movement is the preferred construction for forming complex *wh*-questions in Estonian, then one expects the non-doubling construction to have a limited use and the copy construction to be even rarer. The matter will be readdressed in Section 3 in the light of survey and corpus data.

2. Methodology

2.1. Corpus analysis

In order to observe the usage of complex *wh*-questions in written Estonian, I carried out corpus queries, using language data collected from fiction and newspapers on the one hand and from websites on the other hand. Firstly, I ran a query in three subcorpora of the "Eesti keele koondkorpus" (Estonian Reference Corpus; <http://www.cl.ut.ee/korpused/segakorpus/>), namely, Fiction from the year 1990 onwards (5.6 million words), Daily "Postimees" from the years 1995–2000 (32.9 million words) and Daily "Eesti Päevaleht" from the years 1995–2007 (87.9 million words), which all represent standard written Estonian. Then I repeated the query in the Estonian web corpus etTenTen (<http://www2.keeleeveeb.ee/dict/corpus/ettenten/about.html>), which is based on 686,000 Estonian websites downloaded from the Internet (270 million words), assuming that Internet language use reflects colloquial language phenomena and therefore shows more variation in the pronominal patterns of long *wh*-questions.

I restricted the search phrase to the possible matrix clause sequences containing a second-person pronoun (singular *sa* or plural/polite *te*), e.g. *mis sa arvad* 'what do you-SG think' or *keda te ütlete* 'who do you-PL say'. With the more frequent verb *arvama*, I formed search strings in the present tense only. With *ütleva*, I expanded the query by including the imperfect verb forms, e.g. *mida sa ütlesid* 'what did you say'. Out of the results, I took into account all bi-clausal subject/object questions, regardless of whether they formed a sentence by themselves or a part of a sentence that contains further subordinate clauses, such as (22).

- (22) *M i s s a a r v a d , m i d a s a n ä e d ,*
 what-NOM you-NOM think-2SG what-PART you-NOM see-2SG
k u i s a ü m b e r p ö ö r a d ? (Estonian web corpus etTenTen)
 if you-NOM around turn-2SG
 'What do you think you will see if you turn around?'

2.2. Acceptability judgment test

Since complex *wh*-questions are not very frequent and some pronominal patterns that do not occur in written language may be used in colloquial language, I conducted an acceptability judgement test among 89 native Estonian speakers.

The study involved four factors: 1) type of the pronominal pattern (identical doubling, non-identical doubling, non-doubling), 2) type of the argument questioned (subject, object), 3) animacy of the object being questioned (human, non-human), and 4) nominative-partitive case alternation.

I took all potentially possible pronominal combinations given in Table 2 into consideration when constructing the test sentences, which were each bi-clausal, i.e. long root *wh*-questions with the matrix verbs *arvama* 'think' and *ütleva* 'say'. The test also included alternative constructions containing either a non-finite adverbial clause with a gerund (*sinu/teie arvates* 'in your-SG/your-PL opinion', literally 'in your thinking'), a postpositional

phrase (*Martini sõnul* 'according to Martin') or an adverb (*kuuldavasti* 'as they say').

All in all, I constructed 17 sentences with each of the bridge verbs and 4 alternative sentences, resulting in 38 test sentences. In doing so, I only employed the vocabulary represented among the 10,000 most frequent lemmas in Kaalep, Muischnek 2002, whilst avoiding unusual names. The bridge verbs appeared in second- or third-person forms, the verb *arvama* in the simple present and *ütleva* in the simple past tense, e.g. *Mida sa arvad, keda ta peole kutsub?* 'Who do you think (s)he will invite to the party?', *Mis ta ütles, mis seda müra tekitab?* 'What did (s)he say causes this noise?'

To prevent the informants from ruling out variants that they might find acceptable in spoken language but not in written language, the test sentences were presented orally, not on the questionnaire. The recorded sentences were played in a mixed order with 10 second breaks, during which the participants had to evaluate them using a five-point Likert-type scale, with the scale points defined as follows: 5 — perfectly acceptable, 4 — rather acceptable, 3 — neutral, 2 — rather unacceptable, 1 — absolutely unacceptable. The sentences were not repeated because it was important that the judgements be based on immediate reaction.

The test was performed in three groups. The informants included 29 upper secondary school students (15 male, 14 female, aged 16–17), 37 university students (20 male, 17 female, aged 21–40), 23 volunteering employees of a private company (13 male, 10 female, aged 23–62). The mean age of the participants was 25.8 (SD = 11.1). It was required that the university students and employees do not have linguistic background. The secondary school students filled in a printed questionnaire, the other two groups gave their responses online. I preferred a conveniently accessible sample in order to carry the test out in controlled conditions, so that the respondents did not have a possibility to rewind the recording and rethink their judgements.

For every test sentence, I calculated the median (Mdn) to measure the central tendency of the judgements, which indicates the likeliest response, and the interquartile range (IQR), i.e. the difference between the first and third quartiles of the distribution, to measure dispersion. Due to the ordinal nature of the Likert type data where the scale values cannot be presumed equal, these measurements have been recommended in place of means and standard deviations, which could lead to a misinterpretation of the findings of a survey (see, e.g., Blalock 1979; Jamieson 2004; Allen, Seaman 2007; Kostoulas 2014).

Admittedly, the test was exploratory in nature. The absence of filler sentences and low number of test sentences may have affected the results. Designing the test, I took into account that, despite the abundance of control conditions, it should not become too long and elaborate for the participants to remain focused. The idea was to offer first insight for further research.

3. Results and discussion

3.1. Corpus analysis

Nearly all examples found in each of the corpora involved the verb *arvama* 'think'. The query in the reference corpus (126.4 million words) returned 37 relevant examples, while 272 relevant examples were found in the web

corpus (270 million words). In either case, only one example contained the verb *üttelema* 'say'. Both examples were in the simple past tense.

In the corpus of standard written Estonian compiled of fiction and newspaper texts, most examples of long *wh*-questions with the verb *arvama* questioned a non-human object, as outlined in (23). In this case, all 17 examples of subject questions involved the identical doubling pattern *mis—mis*. Out of the 12 object questions, most involved the nominative-partitive non-identical pattern *mis—mida*, while the identical pattern *mis—mis* was used once. In addition, seven examples were found of a long *wh*-question with a human referent. All of them questioned a subject and involved the non-identical doubling pattern *mis—kes*. The only example with the verb *üttelema*, which was a subject question, also involved a human referent and the *mis—kes*-pattern. There were no examples of the non-doubling construction with the complementizer *et*, nor of identical doubling in case of a human referent.

(23) Results of the corpus query: fiction and newspapers

36 examples with *arvama* 'think':

HUMAN, SUBJ question — 7 examples of non-identical doubling (*mis—kes*)

NON-HUMAN, SUBJ question — 17 examples of identical doubling (*mis—mis*)

NON-HUMAN, OBJ question — 11 examples of non-identical doubling (*mis—mida*), and one (1) example of identical doubling (*mis—mis*)

one example with *üttelema* 'say':

HUMAN, SUBJ question: non-identical doubling (*mis—kes*)

In the etTenTen web corpus, where the data largely originates from blog and forum posts as well as from commentaries on news websites, significantly more examples were found, affirming the colloquiality of the phenomenon. Again, the majority of questions had a non-human referent, in which case all three types of pronominal patterns were attested this time. Table 3 presents the frequencies and percentages of the pronominal patterns represented in long *wh*-questions with the verb *arvama*.

Table 3

Results of the corpus query: *arvama* 'think' (etTenTen web corpus)

HUMAN		Frequency	Percentage
Non-identical doubling		47	17.3%
SUBJ question	<i>mis—kes</i>	41	15.1%
OBJ question	<i>mis—keda</i>	6	2.2%
NON-HUMAN			
Non-identical doubling		64	23.6%
SUBJ question	<i>mida—mis</i>	10	3.7%
OBJ question	<i>mis—mida</i>	54	19.9%
Identical doubling		151	55.7%
SUBJ question	<i>mis—mis</i>	132	48.7%
OBJ question	<i>mis—mis</i>	12	4.4%
	<i>mida—mida</i>	7	2.6%
Non-doubling		9	3.3%
SUBJ question	<i>mis—et</i>	4	1.5%
OBJ question	<i>mida—et</i>	5	1.8%

In case of a human referent, again only non-identical doubling emerged with the nominative *mis* in the matrix clause. In *wh*-questions with a non-human referent, identical doubling (*mis—mis*) dominated if a subject was questioned, which was much more common than questions regarding an object, just like in the reference corpus. If an object was questioned, the non-identical doubling pattern *mis—mida* was preferred. All possible combinations of nominative and partitive forms of *mis* occurred, however, the patterns with *mis* in the higher clause were preferred to the patterns with the partitive form *mida* in the higher clause, probably for reasons of economy.

Non-doubling was observed as well, although much less frequently than the doubling patterns. It occurred both in case of subject (*mis—et*) and object extraction (*mida—et*). In the only example with the verb *üttelema*, which involved a non-human referent and object extraction and where the nominative pattern *mis—et* was used, standard long-distance movement was employed.

In conclusion, the corpus data analyzed indicates that if a person is questioned, non-identical doubling is strongly preferred in Estonian complex *wh*-questions, while identical and non-identical doubling are similarly common in case of a non-human referent. One can further conclude that the identical doubling pattern *mis—mis* is best compatible with subject questions, while the nominative-partitive non-identical pattern *mis—mida* is best compatible with object questions. All in all, doubling patterns with a nominative *mis* in the matrix clause appear to be preferred; long *wh*-questions without pronominal doubling look rare.

Moreover, *arvama* appears to be a substantially more frequent bridge verb than *üttelema*. This is in line with Featherston's (2004 : 182) claim that movement restrictions are not merely syntactic but must be related to lexical factors. Based on German, he has pointed out that the verb *meinen* 'think' allows long-distance *wh*-extraction, whereas *hoffen* 'hope' seems less natural and *bezweifeln* 'doubt' feels entirely inappropriate in the construction. 'Think' has also been found to be by far the most frequent bridge verb in, e.g. Dutch and English (Verhagen 2006 : 334–335).

The corpus data further reveals that certain verbs tend to occur in the lower clause more often than others. These are *olema* 'be' (30.9% of all cases, e.g. *Mis sa arvad, kes need inimesed on?* 'Who do you think these people are?'), *tegema* 'do' (16.4%, e.g. *Mis te arvate, mida mees tegei?* 'What do you think the man has done?'), *juhtuma* 'happen' (9.6%, e.g. *Mis te arvate, mis siis juhtus?* 'What do you-PL think happened then?') and *saama* 'become, happen to' (7.4%, e.g. *Mis sa arvad, mis riigist saab?* 'What do you think will happen to the country?').

3.2. Acceptability judgement test

Similarly to the corpus data, the informants of the acceptability judgement test strongly preferred non-identical doubling in long *wh*-questions with a human referent. With the verb *arvama* 'think', all non-identical doubling patterns were found to be rather or perfectly acceptable by most of the informants (73.1%–83% depending on the pattern). In object questions, *mis—keda* (Mdn = 5, IQR = 1) was preferred to *mida—keda* (Mdn = 4, IQR = 2), whereas in subject questions there was little difference in the perception of the patterns *mis—kes* and *mida—kes* (Mdn = 4, IQR = 1 in both cases).

If the object was extracted, regular long-distance *wh*-movement (*keda—et*: Mdn = 4, IQR = 2.5) was also rated acceptable by more respondents (52.8%) than those who rated it unacceptable (29.2%). In case of subject extraction, the responses to the non-doubling pattern *kes—et* nonetheless vary to a greater extent (Mdn = 3, IQR = 2), being rather considered acceptable by 40.5% and unacceptable by 41.5% of informants.

Identical doubling patterns *kes—kes* (Mdn = 3, IQR = 3) and *keda—keda* (Mdn = 2, IQR = 2) were most often seen as rather or absolutely unacceptable (by 49.5% and 64.1% of respondents, respectively), even though the fact that some informants (32.6% and 16.9%, resp.) still found them acceptable implies that they may occur in colloquial language.

The distribution of judgements for pronominal patterns in long *wh*-questions that contain the verb *arvama* and question a person is presented in Table 4. The number of informants who considered a pattern rather or perfectly acceptable is in the light grey column, while the number of informants who considered a pattern rather or absolutely unacceptable is in the dark grey column, and the number of informants who remained neutral is in the white column (the same holds for subsequent tables). The modes, i.e. most frequent responses, are marked by giving the number of the respective responses in bold.

Table 4

Acceptability judgements: *arvama* 'think', HUMAN

Type of pronominal pattern		5 — perfectly acceptable	4 — rather acceptable	3 — neutral	2 — rather unacceptable	1 — absolutely unacceptable	Median	IQR
Non-identical doubling								
SUBJ question	<i>mis—kes</i>	42 (47.2%)	30 (33.7%)	8 (9.0%)	7 (7.9%)	2 (2.2%)	4	1
	<i>mida—kes</i>	39 (43.8%)	35 (39.3%)	7 (7.9%)	3 (3.4%)	4 (4.5%)	4	1
OBJ question	<i>mis—keda</i>	48 (53.9%)	23 (25.8%)	9 (10.1%)	8 (9.0%)	1 (1.1%)	5	1
	<i>mida—keda</i>	33 (37.1%)	32 (36.0%)	13 (14.6%)	6 (6.7%)	5 (5.6%)	4	2
Identical doubling								
SUBJ question	<i>kes—kes</i>	15 (16.9%)	14 (15.7%)	16 (18.0%)	20 (22.5%)	24 (27.0%)	3	3
OBJ question	<i>keda—keda</i>	7 (7.9%)	8 (9.0%)	17 (19.1%)	33 (37.1%)	24 (27.0%)	2	2
Non-doubling								
SUBJ question	<i>kes—et</i>	12 (13.5%)	24 (27.0%)	16 (18.0%)	19 (21.3%)	18 (20.2%)	3	2
OBJ question	<i>keda—et</i>	22 (24.7%)	25 (28.1%)	16 (18.0%)	17 (19.1%)	9 (10.1%)	4	2.5

With the bridge verb *üttelema* 'say', non-identical doubling was again notably preferred in case of a human referent. However, as shown in Table 5, only the patterns *mis—kes* (Mdn = 5, IQR = 1) and *mis—keda* (Mdn = 4, IQR = 2) seem

Table 5

Acceptability judgements: *üttelema* 'say', HUMAN

Type of pronominal pattern		5 – perfectly acceptable	4 – rather acceptable	3 – neutral	2 – rather unacceptable	1 – absolutely unacceptable	Median	IQR
Non-identical doubling								
SUBJ question	<i>mis—kes</i>	45 (50.6%)	28 (31.5%)	10 (11.2%)	5 (5.6%)	1 (1.1%)	5	1
	<i>mida—kes</i>	33 (37.1%)	15 (16.9%)	22 (24.7%)	12 (13.5%)	7 (7.9%)	4	2
OBJ question	<i>mis—keda</i>	35 (39.3%)	23 (25.8%)	16 (18.0%)	10 (11.2%)	4 (4.5%)	4	2
	<i>mida—keda</i>	12 (13.5%)	28 (31.5%)	19 (21.3%)	22 (24.7%)	8 (9.0%)	3	2
Identical doubling								
SUBJ question	<i>kes—kes</i>	2 (2.2%)	5 (5.6%)	11 (12.4%)	37 (41.6%)	34 (38.2%)	2	1
OBJ question	<i>keda—keda</i>	6 (6.7%)	7 (7.9%)	12 (13.5%)	26 (29.2%)	38 (42.7%)	2	2
Non-doubling								
SUBJ question	<i>kes—et</i>	10 (11.2%)	18 (20.2%)	18 (20.2%)	28 (31.5%)	15 (16.9%)	3	2
OBJ question	<i>keda—et</i>	14 (15.7%)	18 (20.2%)	22 (24.7%)	20 (22.5%)	15 (16.9%)	3	2

to be generally accepted (by 82.1% and 65.1% of respondents, resp.). The patterns *mida—kes* (Mdn = 4, IQR = 2) and *mida—keda* (Mdn = 3, IQR = 2), which were found acceptable in 54% and 45% of the cases, resp., and received a neutral response in more than 20% of the cases, rather left the respondents indecisive.

The identical doubling patterns *kes—kes* (Mdn = 2, IQR = 1) and *keda—keda* (Mdn = 2, IQR = 2) were mainly rated unacceptable by the informants (79.8% and 71.9%, resp.). The non-doubling patterns *kes—et* and *keda—et* (Mdn = 3, IQR = 2 in both cases) were also more often rejected (by 48.4% and 39.4% of respondents, resp.) than approved (by 31.4% and 35.9%, resp.).

As soon as one turns to *wh*-questions with a non-human referent, identical doubling suddenly becomes perfectly acceptable, as illustrated by Table 6, which presents the judgements for sentences with the matrix verb *arvama*. The partitive pattern *mida—mida* used in an object question (Mdn = 4, IQR = 2) was, however, rated considerably lower than the nominative pattern *mis—mis*, which can be used both in subject (Mdn = 5, IQR = 0.5) and object questions (Mdn = 5, IQR = 1); the first pattern was perceived acceptable by 64.1% of informants, whereas the latter was accepted by nearly all of them, regardless of the argument being questioned.

Likewise, the non-identical pattern with the partitive form in the higher clause, *mida—mis* (SUBJ question: Mdn = 4, IQR = 2; OBJ question: Mdn = 4, IQR = 1), was found acceptable by fewer informants (65.2% and 78.6%, resp.) than *mis—mida* (Mdn = 5, IQR = 1) used in an object question, which was accepted by 93.2% of respondents.

Table 6

Acceptability judgements: *arvama* 'think', NON-HUMAN

Type of pronominal pattern		5 – perfectly acceptable	4 – rather acceptable	3 – neutral	2 – rather unacceptable	1 – absolutely unacceptable	Median	IQR
Non-identical doubling								
SUBJ question	<i>mida—mis</i>	28 (31.5%)	30 (33.7%)	16 (18.0%)	10 (11.2%)	5 (5.6%)	4	2
OBJ question	<i>mis—mida</i>	64 (71.9%)	19 (21.3%)	4 (4.5%)	1 (1.1%)	1 (1.1%)	5	1
	<i>mida—mis</i>	44 (49.4%)	26 (29.2%)	13 (14.6%)	4 (4.5%)	2 (2.2%)	4	1
Identical doubling								
SUBJ question	<i>mis—mis</i>	67 (75.3%)	13 (14.6%)	7 (7.9%)	1 (1.1%)	1 (1.1%)	5	0.5
OBJ question	<i>mis—mis</i>	61 (68.5%)	14 (15.7%)	8 (9.0%)	5 (5.6%)	1 (1.1%)	5	1
	<i>mida—mida</i>	28 (31.5%)	29 (32.6%)	22 (24.7%)	6 (6.7%)	4 (4.5%)	4	2
Non-doubling								
SUBJ question	<i>mis—et</i>	15 (16.9%)	16 (18.0%)	24 (27.0%)	24 (27.0%)	10 (11.2%)	3	2
OBJ question	<i>mis—et</i>	44 (49.4%)	21 (23.6%)	13 (14.6%)	8 (9.0%)	3 (3.4%)	4	2
	<i>mida—et</i>	22 (24.7%)	25 (28.1%)	16 (18.0%)	17 (19.1%)	9 (10.1%)	4	2.5

In case of object extraction, the non-doubling patterns *mis—et* (Mdn = 4, IQR = 2) and *mida—et* (Mdn = 5, IQR = 1) were also mainly considered acceptable (by 73% and 79.7% of informants, resp.). In case of subject extraction, however, *mis—et* (Mdn = 3, IQR = 2) was found to be unacceptable slightly more often (in 38.2% of the cases) than acceptable (in 34% of the cases).

With the verb *üttelema*, most of the patterns were again considered to be rather acceptable with the exceptions of the non-doubling *mis—et* if a subject is extracted (Mdn = 3, IQR = 2.5), just like in the case of *arvama*, and also the identical object extraction pattern *mida—mida* (Mdn = 3, IQR = 1). Nevertheless, even in these cases there were more informants who found the pattern acceptable (49.4% and 45%, resp.) than those who rated it unacceptable (29.2% and 32.6%, resp.), as can be seen in Table 7.

The non-identical doubling patterns were, in general, all perceived similarly (Mdn = 4, IQR = 2); *mida—mis* was considered acceptable by 66.3% of respondents, both in case of subject and object extraction, and the object extraction pattern *mis—mida* by 70.8%. The identical doubling pattern *mis—mis* (SUBJ question: Mdn = 4, IQR = 2; OBJ question: Mdn = 4, IQR = 1) was considered acceptable by more informants (66.3%) in case of subject extraction than in case of object extraction (52.8%). Similarly, the non-doubling object extraction patterns *mis—et* (Mdn = 4, IQR = 3) and *mida—et* (Mdn = 4, IQR = 2) were accepted by more than half of the informants (53.9% and 55%, resp.).

Table 7

Acceptability judgements: *üttelema* 'say', NON-HUMAN

Type of pronominal pattern		5 — perfectly acceptable	4 — rather acceptable	3 — neutral	2 — rather unacceptable	1 — absolutely unacceptable	Median	IQR
Non-identical doubling								
SUBJ question	<i>mida—mis</i>	37 (41.6%)	22 (24.7%)	20 (22.5%)	8 (9.0%)	2 (2.2%)	4	2
OBJ question	<i>mis—mida</i>	34 (38.2%)	29 (32.6%)	14 (15.7%)	8 (9.0%)	4 (4.5%)	4	2
	<i>mida—mis</i>	26 (29.2%)	33 (37.1%)	14 (15.7%)	14 (15.7%)	2 (2.2%)	4	2
Identical doubling								
SUBJ question	<i>mis—mis</i>	29 (32.6%)	30 (33.7%)	15 (16.9%)	9 (10.1%)	6 (6.7%)	4	2
OBJ question	<i>mis—mis</i>	19 (21.3%)	28 (31.5%)	21 (21.3%)	14 (15.7%)	7 (7.9%)	4	1
	<i>mida—mida</i>	20 (22.5%)	20 (22.5%)	19 (24.7%)	21 (23.6%)	8 (9.0%)	3	1
Non-doubling								
SUBJ question	<i>mis—et</i>	22 (24.7%)	22 (24.7%)	19 (21.3%)	19 (21.3%)	7 (7.9%)	3	2.5
OBJ question	<i>mis—et</i>	27 (30.3%)	21 (23.6%)	14 (15.7%)	20 (22.5%)	7 (7.9%)	4	3
	<i>mida—et</i>	27 (30.3%)	22 (24.7%)	21 (23.6%)	13 (14.6%)	6 (6.7%)	4	2

The test sentences containing the gerund *arvates* (*Kes seda sinu arvates tegi?* 'Who did it, in your opinion?', *Mida ma teie arvates tegema peaks?* 'What should I do, in your-PL opinion?'; Mdn = 5, IQR = 1 in both cases) were considered acceptable by 92.1% and 94.3% of informants, resp., thereby predominantly perfectly acceptable. This refers that such a non-finite construction may often be used as an alternative to long *wh*-questions with the verb *arvama*. Nevertheless, the biclausal subject question with the doubling pattern *mis-mis* and the object question with the pattern *mis-mida* were rated as highly as their uniclausal counterparts.

Out of the uniclausal alternatives to complex *wh*-questions with the verb *üttelema*, the one featuring the postposition *sõnul* 'according to' (*Kes selle auto Martini sõnul korda tegi?* 'Who fixed this car, according to Martin?'; Mdn = 5, IQR = 1) was preferred to the one featuring the adverb *kuuldavasti* 'as they say' (*Mida Anne kuuldavasti õppida kavatseb?* 'What do they say Anne is planning to study?'; Mdn = 4, IQR = 2). While the former was found acceptable by 85.3% of respondents, the latter was accepted by just 50.6% and cannot be considered a common alternative.

Overall, the sentences with the matrix verb *arvama* received higher ratings, which is in accordance with the results of corpus analysis. However, while the corpus queries returned next to no examples of complex *wh*-questions with the verb *üttelema*, the participants of the acceptability judgement test

found such sentences to be largely acceptable if the non-identical doubling patterns *mis—kes*, *mis—keda*, *mis—mida*, *mida—mis* or the identical doubling pattern *mis—mis* (in SUBJ question) were used. This indicates that *üttelema* colloquially figures as a bridge verb, although it has a more limited bridge feature than *arvama*, which allows more variation in pronominal patterns.

Complex object questions, especially those with a non-human referent, also give rise to greater variation. Firstly, subject questions do not permit the nominative-partitive case alternation of the pronoun *mis* in the embedded clause. More importantly, regular long-distance *wh*-questions rather seem to be permitted if an object is extracted from the subordinate clause. The difference turned out to be more significant if the matrix verb was *arvama*. While inanimate object extraction was found to be acceptable by most informants, subject extraction was rather considered unacceptable. The human object extraction was only accepted by slightly more than half of the respondents but was still rated considerably higher than the subject extraction. With the verb *üttelema*, the difference was quite subtle, and if a person was questioned, object extraction was considered rather unacceptable, same as subject extraction.

Thus, long-distance *wh*-movement appears to be limited in Estonian, rather occurring with a non-human object and depending on the bridge verb. The only construction that can account for all accepted patterns of pronominal doubling observed in this study is partial *wh*-movement. That also includes the seemingly identical patterns *mis—mis* and *mida—mida*. In fact, analyzing them as *wh*-copying would contradict the crosslinguistic observation that in most languages partial *wh*-movement is not equally likely to co-occur with long-distance *wh*-movement nor its structural variant *wh*-copying. Namely, it would entail that in case of a non-human referent *wh*-copying would somehow be as acceptable as partial *wh*-movement and more acceptable than ordinary long-distance *wh*-extraction. Taking into account that non-identical doubling representing partial *wh*-movement was considered highly acceptable by native Estonian speakers both with a human and a non-human referent, the inanimate identical doubling patterns should also be analyzed as partial *wh*-movement. Adopting the IDA for Estonian would mean that, despite their similar form, the pronouns introducing the matrix and the subordinate clause are not part of the same movement chain and hence not coreferential.

The conclusion that partial *wh*-movement is the overall preferred construction in Estonian for forming complex *wh*-questions allows to place Estonian among languages as diverse as Hungarian, Russian, Serbian-Croatian, Czech, Romani, Polish, Albanian, Frisian, but also Hindi, Bangla, Kashmiri, Marathi, Iraqi Arabic, Warlpiri, and Passamaquoddy (Fanselow 2006 : 442–443). Stepanov and Stateva (2006) claim that long-distance *wh*-movement and partial *wh*-movement share the same derivational history, and the difference between languages with and without partial *wh*-movement lies in the lexical matter, namely, in whether the scope marker, i.e. the *what*-phrase is overt or silent in that particular language. A mixed example (24) found in the etTenTen web corpus, where both the complementizer and the interrogative pronoun emerge in the SpecCP of the subordinate clause, seems to support the assumption that the *what*-element is there even if no doubling is employed. Apparently, it may get spelled out even if not necessary, preventing the lower pronoun from moving out of the subordinate clause.

conducted among native speakers. The results suggest that a similar syntactic structure underlies non-identical doubling and identical doubling with an inanimate referent, since the patterns occur with a similar frequency and are perceived as equally acceptable. Although identical doubling has generally been analyzed as *wh*-copying, i.e. a structural variant of long-distance *wh*-movement where the intermediate link of the movement chain gets spelled out, it is suggested that the Estonian inanimate doubling patterns *mis—mis* and *mida—mida* should be seen as cases of partial *wh*-movement, which is usually associated with non-identical doubling only.

All in all, partial *wh*-movement appears to be the construction employed in most cases when a complex *wh*-question is formed in Estonian. I further argue that each clause contains an independent *wh*-movement chain and the higher pronoun is coindexed to the entire embedded clause, thereby adopting the indirect dependency approach. In case of a human referent, identical doubling was accepted to a minor extent by the tested informants, leading to a conclusion that *wh*-copying is not intrinsic to Estonian. Regular long-distance *wh*-movement does not seem to be freely permitted either, receiving higher ratings with the matrix verb *arvama* 'think' and in case of object extraction.

In line with the crosslinguistic observation that the verb equivalent to 'think' tends to be the most common bridge verb, the verb *arvama* proved to be generally more accepted and to entail a greater extent of syntactic variation than *üttelema* 'say'. The overall preference for partial *wh*-movement and restricted use of long-distance *wh*-movement in Estonian matches with the tendency that these constructions are in complementary distribution.

Acknowledgements

I am grateful to Professors Sjef Barbiers and Craig Thiersch, who were teaching the Micro- and Macrovariation course during my exchange period at Utrecht University in 2014, both for inspiring me to tackle this topic and for helpful advice regarding the pilot study. I would also like to thank Heete Sahkai from the Institute of the Estonian Language and Pille Eslon from Tallinn University for advising me throughout the research.

Address

Kais Allkivi
kais.allkivi@gmail.com

Abbreviations

ACC — accusative; ALL — allative; **comp** — complementizer; **CP** — complementizer; DAT — dative; **DDA** — direct dependency approach; **IDA** — indirect dependency approach; **IQR** — interquartile range; **Mdn** — median; **NOM** — nominative; **NP** — noun phrase; **OBJ** — object; **PART** — partitive; **PL** — plural; **2PL** — 2nd person plural; **3PL** — 3rd person plural; **POL** — polite register; **PP** — past participle; **PST** — past tense; **SG** — singular; **2SG** — 2nd person singular; **3SG** — 3rd person singular; **SUBJ** — subject.

REFERENCES

- Allen, I. E., Seaman, C. A. 2007, Likert Scales and Data Analysis. — Quality Progress. The Official Publication of American Society for Quality. <http://asq.org/quality-progress/2007/07/statistics/likert-scales-and-data-analyses.html>.
- Barbiers, S., Koenenman, O., Lekakou, M. 2008, Syntactic Doubling and the Structure of Chains. — Proceedings of the 26th West Coast Conference on Formal Linguistics, Somerville, MA, 77–86. <http://www.lingref.com/cpp/wccfl/26/paper1658.pdf>.
- Blalock, H. M. Jr. 1979, Social Statistics, New York.
- Boef, E. 2013, Doubling in Relative Clauses. Aspects of Morphosyntactic Microvariation in Dutch, Utrecht. <https://dspace.library.uu.nl/handle/1874/261909>.
- Chen, L. 2000, Moving Just the Feature. — Wh-Scope Marking, Amsterdam, 77–100.
- Chomsky, N. 1995, The Minimalist Program, Cambridge, MA.
— 2000, Minimalist Inquiries. The Framework. — Step by Step. Essays on Minimalist Syntax in Honor of Howard Lasnik, Cambridge, MA, 89–155.
- Davison, A. L. 1984, Syntactic Constraints on Wh-in-situ. Wh-Questions in Hindi-Urdu. Paper Presented at the Annual Meeting of the Linguistic Society of America.
- Dayal, V. S. 1994, Scope Marking as Indirect Wh-Dependency. — Natural Language Semantics 2, 137–170.
- Erelt, M., Erelt, T., Ross, K. 2007, Eesti keele käsiraamat, Tallinn.
- Fanselow, G. 2006, Partial Wh-movement. — The Blackwell Companion To Syntax, Malden, MA, 437–492.
- Fanselow, G., Mahajan, A. 2000, Towards a Minimalist Theory of Wh-Expletives, Wh-Copying, and Successive Cyclicity. — Wh-Scope Marking, 195–230, Amsterdam.
- Featherston, S. 2004, Bridge Verbs and V2 Verbs. The Same Thing in Spades? — Zeitschrift für Sprachwissenschaft 23 (2), 181–209.
- Felser, C. 2001, Wh-Expletives and Secondary Predication. German Partial Wh-Movement Reconsidered. — Journal of Germanic Linguistics 13, 5–38.
— 2004, Wh-Copying, Phases, and Successive Cyclicity. — Lingua 114 (5), 543–574.
- Horvath, J. 1997, The Status of “Wh-Expletives” and the Partial Wh-Movement Construction of Hungarian. — Natural Language and Linguistic Theory 15, 509–572.
— 2000, On the Syntax of “Wh-Scope Marker” Constructions. — Wh-Scope Marking, Amsterdam, 271–316.
— 2007, Separating “Focus Movement” from Focus. — Phrasal and Clausal Architecture. Syntactic Derivation and Interpretation, Amsterdam—Philadelphia, 108–145.
- Huhmarniemi, S. 2012, Finnish A'-Movement. Edges and Islands, Helsinki (Studies in Cognitive Science 2: 2012).
- Jamieson, S. 2004, Likert Scales. How to (Ab)use them. — Medical Education 38 (12), 1217–1218.
- Kaalep, H-J., Muischnek, K. 2002, Eesti kirjakeele sagedussõnastik, Tartu. <http://www.cl.ut.ee/ressursid/sagedused/index.php?lang=en>.
- Klepp, M. 2001, Partial Wh-Movement in German, Dublin.
- Kostoulas, A. 2014, How to Interpret Ordinal Data. <https://achilleaskostoulas.com/2014/02/23/how-to-interpret-ordinal-data/>.
- Marác, L. K. 1991, Asymmetries in Hungarian, San Sebastian.
- Lubanska, M. 2004, Wh-Scope Marking in Polish. — Poznan Studies in Contemporary Linguistics 39, 73–88.
- McDaniel, D. 1989, Partial and Multiple Wh-Movement. — Natural Language and Linguistic Theory 7, 565–604. <http://ling.umd.edu/~stacey/McDaniel.pdf>.
- Metslang, H. 1981, Kõsilause eesti keeles, Tallinn.

- M u s, N. 2015, Interrogative Words and Content Questions in Tundra Nenets, Szeged. http://doktori.bibl.u-szeged.hu/2764/1/Content_questions_in_TN_Nikolett_Mus_2015_06.pdf
- P a n k a u, A. 2013, Replacing Copies. The Syntax of Wh-Copying in German, Utrecht. <https://dspace.library.uu.nl/handle/1874/288519>
- R e t t, J. 2006, Pronominal vs. Determiner Wh-Words. Evidence from the Copy Construction. — Empirical Issues in Syntax and Semantics 6, 355–374. <http://www.cssp.cnrs.fr/eiss6/rett-eiss6.pdf>
- R i e m s d i j k, H. C. van 1982, Correspondence Effects and the Empty Category Principle, Tilburg (Tilburg Papers in Language and Literature 12).
- R i z z i, L. 1992, Argument/Adjunct (A)symmetries. — Proceedings of North East Linguistic Society 22, 365–381.
- S a b e l, J. 1998, Principles and Parameters of Wh-Movement. Habilitationsschrift, Frankfurt/Main.
- S a l m i n e n, T. 2012, Tundra Nenets. <http://www.helsinki.fi/~tasalmin/sketch.html>.
- S c h i p p e r s, A. 2010a, Partial Wh-Movement and Wh-Copying in Dutch. Evidence for an Indirect Dependency Approach. — Proceedings of the Thirty Sixth Annual Meeting of the Berkeley Linguistics Society. February 6–7, 2010, Berkeley, CA, 338–352.
- 2010b, On the (Un)availability of Long-Distance Movement. — Movement and Clitics. Adult and Child Grammar, Newcastle upon Tyne, 39–62.
- S c h o o r l e m m e r, E. 2009, Agreement, Dominance and Doubling. The morphosyntax of DP, Utrecht. https://openaccess.leidenuniv.nl/bitstream/handle/1887/13952/thesis_schoorlemmer_finaal.pdf?sequence=1.
- S t e p a n o v, A., S t a t e v a, P. 2006, Successive Cyclicity as Residual Wh-Scope Marking. — *Lingua* 116, 2107–2153.
- The Uralic Languages, London—New York 1998.
- T o f t, Z. 2001, Is There Ever Multiple Wh-Movement? Evidence from Superiority Effects and Focus in Hungarian. — Durham Working Papers in Linguistics 7, 126–144.
- V a i n i k k a, A. 1989, Defining Syntactic Representations in Finnish. PhD Dissertation, Amherst.
- V e r h a g e n, A. 2006, On Subjectivity and Long-Distance Wh-Movement. — Subjectification. Various Paths to Subjectivity, Berlin—New York, 323–346. http://www.academia.edu/8847362/On_subjectivity_and_long_distance_Wh-movement_.
- Грамматика современного удмуртского языка. Синтаксис простого предложения, Ижевск 1970.
- Н е н я н г М. А. 2005, Русско-ненецкий разговорник, Санкт-Петербург.

КАЙС АЛЛКИВИ (Таллинн)

**СДВОЕННОЕ УПОТРЕБЛЕНИЕ
ВОПРОСИТЕЛЬНО-ОТНОСИТЕЛЬНЫХ МЕСТОИМЕНИЙ
В СЛОЖНОПОДЧИНЕННЫХ ВОПРОСИТЕЛЬНЫХ ПРЕДЛОЖЕНИЯХ
ЭСТОНСКОГО ЯЗЫКА**

Исследования синтаксического варьирования показали, что наряду с длинным сдвигом вопросительного слова, в связи с которым вопросительно-относительное местоимение из позиции аргумента в придаточном предложении перемещается в начало главного предложения, а придаточное предложение вводится с помощью связующего слова, в сложноподчиненных предложениях встречается также сдвоенное употребление местоимений, в таком случае обе составные части сложного предложения начинаются местоимениями. Такое дублирование вопросительных слов анализировалось прежде всего в индоев-

ропейских языках, при исследовании уральских языков ему уделялось мало внимания. В данной статье рассматриваются эстонские вопросительные предложения, состоящие из двух частей, при этом главное предложение содержит глагол-мостик *arvama* 'считать, полагать' или *üttelema* 'сказать, говорить'. Автор поставила перед собой цель выяснить, какие местоимения и как часто встречаются в сложноподчиненных вопросительных предложениях, а также как восприятие их пользователем языка зависит от того, касается вопрос субъекта или объекта, одушевленный или неодушевленный референт местоимения, каков падеж местоимения(-й) и какой глагол употреблен в главном предложении.

В эстонском языке использование сдвига вопросительного слова ограничено. Однако встречается как идентичное, так и неидентичное сдвоенное употребление вопросительно-относительных местоимений, для которого характерна асимметрия: идентичное употребляется при неодушевленном референте (обе части предложения начинаются с местоимения *mis* 'что'), тогда как при одушевленном предпочитается неидентичное (в начале главного предложения — местоимение *mis*, в начале придаточного — местоимение *kes* 'кто'). Автор приходит к выводу, что в обоих случаях между местоимениями главного и придаточного предложений имеется косвенная, а не прямая зависимость, т. е. в обоих случаях вопросительное слово перемещается из позиции аргумента в начало предложения, и местоимение в начале главного предложения соотносится с придаточным предложением в целом.