

**UCLA**

**UCLA Electronic Theses and Dissertations**

**Title**

Contrast and Word Order in Estonian

**Permalink**

<https://escholarship.org/uc/item/9r3592nq>

**Author**

Kaps, Marju

**Publication Date**

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Contrast and Word Order in Estonian

A thesis submitted in partial satisfaction  
of the requirements for the degree Master of Arts

in Linguistics

by

Marju Kaps

2020

© Copyright by

Marju Kaps

2020

# ABSTRACT OF THE THESIS

Contrast and Word Order in Estonian

by

Marju Kaps

Master of Arts in Linguistics

University of California, Los Angeles, 2020

Professor Jesse A. Harris, Chair

In this thesis, explore the information structure and syntax of declarative matrix V3 and verb-final clauses in Estonian – a discourse-configurational language that typically adheres to V2 word order (Ehala, 2006; Holmberg, 2015). I apply information structural diagnostics like contrastive ellipsis (Repp, 2009) and pronoun choice (Pajusalu, 1997) to propose my central hypothesis that, in Estonian, multiple preverbal constituents are only licensed in clauses containing a left-peripheral Contrastive Topic (CT). Left-peripheral Contrastive Foci (CF), which are not syntactically distinguished from CTs in the closely related language Finnish (Vallduví & Vilkuna, 1998), do not license multiple preverbal elements in Estonian. This hypothesis is supported by a naturalness

rating experiment in which contrastive ellipsis clauses with CT and CF remnants are evaluated in the context of V2 and V3 antecedent clauses. An additional manipulation of discourse contexts in the experiment shows that V3 clauses structurally constrain the comprehenders' interpretation of the information structure of the clause, to the extent that biasing cues in the discourse context have little effect. These findings suggest that, rather than being a stylistic anomaly in an otherwise V2 language, non-V2 matrix clauses warrant careful study and promise to provide deep insight to the syntax of discourse-configurational languages. I set the stage for future research into this topic by challenging a previously assumed multiple Topic approach to Estonian left periphery (Lindström, 2005; Henk, 2009).

The thesis of Marju Kaps is approved.

Hilda Koopman

Timothy Stowell

Jesse A. Harris, Committee Chair

University of California, Los Angeles

2020

## TABLE OF CONTENTS

1. Introduction.....	1
2. Information Structure.....	3
2.1 QUD .....	3
2.2 Topics .....	4
2.3 Foci.....	5
2.4 Contrastive Topics.....	9
3. Estonian Word Order .....	12
3.1 The nature of V2 in Estonian .....	12
3.2 Contrastive ellipsis as a diagnostic for CTs and CFs .....	16
3.3 Pronominal choice as a diagnostic for contrast.....	18
3.4 Violating V2 in Estonian.....	19
4. Naturalness Rating Experiment .....	21
4.1 Design and materials .....	23
4.2 Participants and procedures.....	26
4.3 Results and discussion.....	27
5. The Syntax of Non-V2 Clauses .....	33
6. Conclusion .....	37
Appendix A: Naturalness rating materials.....	39

Experimental items.....	39
Catch items .....	45
Bibliography .....	47



## LIST OF TABLES

<b>Table 1.</b> Predictions by conditions, along with example target sentences .....	23
<b>Table 2.</b> Two types of context sentences used in the target items .....	24
<b>Table 3.</b> Statistical findings for naturalness ratings crossing word order and ellipsis type in a linear mixed effects model. * indicates statistical significance at the $p < .05$ level. ....	28
<b>Table 4.</b> Post-hoc comparisons for a linear mixed effects model of naturalness ratings crossing word order and ellipsis type. * indicates statistical significance at the $p < .05$ level. ....	29

## LIST OF FIGURES

<b>Figure 1.</b> Means and standard errors of naturalness ratings to OVS, SOV and SVO matrix clauses followed by CT ellipsis and CF ellipsis. ....	27
<b>Figure 2.</b> Means and standard errors of naturalness ratings to OVS, SOV and SVO matrix clauses followed by CT ellipsis and CF ellipsis, under two kinds of contexts.....	30

## 1. Introduction

Much like other Finno-Ugric languages, Estonian is a discourse-configurational language in which the possible syntactic positions of a particular linguistic expression are determined by that constituents' information structural status (Henk, 2009). For instance, Topics, or discourse-given elements that the clause is about, usually occur clause-initially. Foci, or elements contributing new information to the clause, usually occur clause-finally. At the same time, Estonian is commonly regarded to be a Verb-Second (V2) language where a single Topic constituent precedes the finite verb (Holmberg, 2015). The requirement for one constituent to move to a preverbal position<sup>1</sup> in declarative matrix clauses appears to operate somewhat independently of information structure. The fact that Estonian has a strong tendency for V2 order can be seen in clauses where no Topic constituent is available. In example (1) the clause-initial 'rain' is clearly not a Topic – the clause does not address the question 'What is going on with the rain?'; indeed, if wasn't raining then there would be no rain to speak of.

- (1) Vihma            sajab  
      *rain.PART*    *rains*  
      'It is raining.'

Despite the fact that Estonian shows V2 tendencies, previous corpus research indicates that multiple constituents may be allowed to occur preverbally if they are sufficiently discourse-given (Lindström, 2005). This is an unexpected finding for a V2 language. Unfortunately, identifying

---

<sup>1</sup> The verb may also occur alone in a clause, e.g. example (a), but when another constituent is present, especially in out-of-the-blue, wide Focus 'What's happening' contexts, V2 order is strongly preferred.

- (a) Sajab  
      *rains*  
      'It's raining'

discourse objects, such as Topics and Foci, can be challenging in a language that does not mark these elements morphologically, particularly in corpus data. The present thesis investigates the conditions under which multiple preverbal elements are allowed in declarative matrix clauses in Estonian, using a more controlled experimental procedure. I report on a novel naturalness rating experiment using contrastive ellipsis to constrain the information structure of full matrix clauses. The experiment provides evidence to support my central hypothesis that multiple preverbal elements may occur in clauses with an initial Contrastive Topic (CT) constituent.<sup>2</sup> Additional linguistic data are considered in order to evaluate possible syntactic analyses that could account for the observations regarding Estonian V2 and non-V2 clauses.

I begin by defining the information structural concepts I rely on in this thesis in Section 2. Section 3 looks at the nature of V2 Estonian, setting forward the hypothesis that it is in clauses containing CTs that V2 order may be deviated from. This hypothesis is tested in the experiment reported in Section 4. Section 5 deals with additional data, arguing against a previously proposed analysis that would treat CT clauses as a slight variation on typical V2 clauses. Instead, it is proposed that the syntactic derivations of CT clauses in Estonian fundamentally differ from the analysis that has been set forward for non-CT clauses which must adhere to V2.

---

<sup>2</sup> As noted, the V2 requirement allows for certain information structural principles in constituent order to be violated; Interestingly, as is the case with CT-initial clauses, it appears that the V2 requirement itself can be violated under specific information structural conditions.

## 2. Information Structure

### 2.1 QUD

In order to define information structurally relevant notions like Topic (T), Focus (F) and CT<sup>3</sup>, I rely on question-answer contexts to disambiguate the information structure of the examples. The Question Under Discussion (QUD) refers to an explicit or implicit question which interlocutors aim to answer at a particular time in a conversation (Roberts, 1996). The QUD is constantly evolving as answers to previous QUDs are added to the common ground or the QUD is broken down into subquestions. I assume that multiple subquestions to a particular QUD can be relevant at the same time, yielding a conjunctive QUD or a sequence of parallel QUDs, as shown in (2). In the present thesis, I treat QUDs involving a conjunction as equivalent to multiple parallel QUDs.<sup>4</sup>

(2) The evolving of a QUD in 3 steps

#### Step 1

Q: How was the food at the party?

A: The people who tried the food at the party said that it was very good.

#### Step 2

Q: Who tried the food?

A: John and Mary tried the food.

#### Step 3

Q: What did John and Mary try? (conjunctive QUD)

Q1: What did John try?      Q2: What did Mary try? (parallel QUDs)

A1: John tried the sweets.      A2: Mary tried everything.

---

<sup>3</sup> In examples, I mark the discourse status of relevant constituents with subscripts

<sup>4</sup> This circumvents issues with joint interpretations of conjunctive questions, e.g. ‘What did both John and Mary try?’, which propose a range of issues outside of the scope of the present thesis

Every informative statement addresses some QUD by being a congruent answer to it. By congruence, I mean that the question and answer match in their information structure – that is, in the entities that are marked as given information, as having Focus alternatives, or being Topical<sup>5</sup>. For instance, example (3) shows a congruent answer to the discourse question, as the Topic and Focus of the answer match those in the question.

- (3) Q: Where<sub>EF</sub> does Mary<sub>T</sub> live?  
A: Mary<sub>T</sub> lives in Paris<sub>F</sub>

In (4), on the other hand, we see an incongruent answer. This is because John has not been previously identified as the Topic by the question, and the answer says nothing about the Topic of the question, Mary.

- (4) Q: Where<sub>EF</sub> does Mary<sub>T</sub> live?  
A: John lives in Paris

This tight pairing between QUDs and their direct answers gives us a shortcut to the kinds of contexts in which a particular sentence would be informative. In Sections 2.2-2.4, I rely on explicit QUDs to set up the appropriate discourse context.

## 2.2 Topics

Topics are defined as linguistic expressions referring to salient entities in the discourse context that a particular sentence is about (Reinhart, 1981). This definition assumes that Topic expressions are referential. Identifying discourse objects can be challenging, so we rely on

---

<sup>5</sup> See Sections 2.2-2.4 for definitions

properties like referentiality to diagnose constituents as Topics. Namely, Topics can be referred back to in a discourse using pronouns (5).

- (5) Q: What happened to **Mary**?  
 A: **Mary**<sub>T</sub> caught the flu. However, **she** is feeling much better now.

In Estonian, Topics canonically occur in the clause-initial position<sup>6</sup>. As shown in examples (6) and (7), topicalizing different elements yields different word orders. The declarative sentences in (6) and (7) do not differ in their truth conditions. They differ only in their information structure, as can be seen from the difference between the questions they answer.

- (6) Q: Who lives in that house nowadays?  
 A: [Selles majas]<sub>T</sub> elab praegu [Mari]<sub>F</sub>  
*that.INE house.INE lives.3S now Mari*  
 ‘MARI lives in that house nowadays.’
- (7) Q: Where does Mari live nowadays?  
 A: Mari<sub>T</sub> elab praegu [selles majas]<sub>F</sub>  
*Mari lives now that.INE house.INE*  
 ‘Mari lives in THAT HOUSE nowadays.’

### 2.3 Foci

Foci are expressions with alternatives that are relevant for their interpretation (Rooth, 1992). Foci can be identified in direct answers to Wh-questions, where the focused constituent in the answer corresponds to the Wh-phrase in the question, as shown in (6) and (7) above. Broadly, a distinction can be drawn between new information Focus and Contrastive Focus (CF) (Büring,

---

<sup>6</sup> I do not say that Topics ‘must occur in the clause-initial position’ because this is not the case for questions, which I wish to also analyze as being able to have Topics.

2016). New information Focus is used in contexts where Focus alternatives are not yet established in the common ground. For instance, in (8) with two speakers A and B, the proposition ‘Mary ate salad’ is added to the common ground following B’s utterance. The alternative set of ‘salad’ could in principle include any number of food items.

- (8) A: What did Mary eat?  
 B: Mary ate salad<sub>F</sub>

CF is used to explicitly exclude a Focus alternative from the common ground (Lee, 2003). In (9) with three speakers A, B and C, the utterance ‘No, Mary ate salad<sub>F</sub>’ both adds ‘Mary ate salad’ and removes ‘Mary ate steak’ from the common ground. The alternative set of ‘salad’ here is the smaller, contextually more salient set {steak, salad}.

- (9) A: What did Mary eat?  
 B: I think that Mary ate steak<sub>F</sub>  
 C: No, Mary ate salad<sub>F</sub>

In Estonian, Foci may occur in a clause-final position (10), although this is not obligatory (11). The focused element receives intonational prominence whether it is in the clause-final position or not (Salveste, 2015).

- (10) Q: Who did Mari give the present to?  
 A: Mari andis kingituse ANNALE<sub>F</sub>  
*Mari gave.3S present.GEN Anna.ALL*  
 ‘Mari gave the present to ANNA’

- (11) Q: Who did Mari give the present to?  
 A: Mari andis ANNALE<sub>F</sub> kingituse  
*Mari gave.3S Anna.ALL present.GEN*  
 ‘Mari gave the present to ANNA’



The ability to occur in the clause-final Focus position does not appear to be dependent on the constituent having a specific kind of Focus. This can be seen in the naturalness of clause-final Focus in both unbiased new information contexts (12) as well as in biased, corrective CF contexts as shown in (13).

- (12) A: What did Mari finish eating?  
 B: Mari sõi vist ära koogi<sub>F</sub>  
*Mari ate.3S probably away cake.GEN*  
 ‘Probably Mari finished eating (‘ate away’) a/the CAKE.’

- (13) A: It appears that Mari finished all the cookies.  
 B: Ei, Mari sõi hoopis ära koogi<sub>F</sub>  
*no Mari ate.3S rather away cake.GEN*  
 ‘No, to the contrary, Mari finished (‘ate away’) a/the CAKE.’

The syntactic position of clause-final Focus can be diagnosed by looking at the position of verbal particles. Particles that form verbal complexes can generally be taken as reliable indicators of the position of the VP as they do not undergo movement out of the VP (Koopman & Szabolcsi, 2000). The ability of a focused element to follow a verbal particle such as *ära* ‘away’ in examples (12) and (13) above suggests that focused elements may occur in a dedicated VP-adjacent Focus position, much like what has been proposed for Hungarian (Horvath, 1995). Unlike examples (12) and (13) above, in out-of-the-blue contexts that do not have narrow Focus on the object, the object must precede the verbal particle (14).

- (14) Q: What happened?  
 A: Mari sõi täna koogi ära  
*Mari ate.3S today cake.GEN away*  
 ‘Today Mari finished eating (‘ate away’) a/the cake’

If the object does not carry narrow Focus, it cannot occur after the verbal particle, as seen in example (15).

- (15) Q: What happened?  
A: # Mari sõi täna ära koogi  
*Mari ate.3S today away cake.GEN*  
'Today Mari finished eating ('ate away') a/the cake'

This suggests that in (12) and (13), it is the narrow Focus on the object 'cake' that allows it to be moved rightwards, out of the VP.

So far, we have seen Focus on lexical constituents in direct answers to explicit or implicit Wh-questions. In the present thesis, I also refer to Polarity Focus, that is, Focus in direct, congruent answers to polar questions such as (16). Here, intonational prominence is placed on the inflected verb.<sup>7</sup>

- (16) Q: Did Mari eat the cake?  
A: Mari ei SÖÖNUD kooki ära  
*Mari NEG eat.PTC cake.PART away*  
'Mari did NOT finish eating the cake.'

A syntactic projection for clausal polarity is typically assumed to occur somewhere above the TP in the syntactic structure (Holmberg, 2016). For this reason, I assume that the inflected verb in Polarity Focus clauses is raised above the TP. An exploration of the nature of a possible left-peripheral Focus position (*cf.* Rizzi, 1997) falls outside the scope of this thesis.

---

<sup>7</sup> The negative particle *ei* functions like a clitic and cannot be separated from the verbal participle form.

## 2.4 Contrastive Topics

Similarly to CFs, CTs have Focus alternatives that are contextually salient (Krifka, 1998). CTs are Topic-like as they must be discourse-given (17) or at least easy to accommodate as given (18). In (18), the Questioner may not know the names of the Addressee’s brothers at the time of asking the question, but it is clear that the answer is about the two brothers – Bill and Bob.

- (17) Q: Where do John and Mary live?  
A: John<sub>CT</sub> lives in Paris<sub>F</sub> and Mary<sub>CT</sub> lives in Rome<sub>F</sub>
- (18) Q: What do your brothers do for a living?  
A: Bill<sub>CT</sub> is an accountant<sub>F</sub> and Bob<sub>CT</sub> is a meteorologist<sub>F</sub>

There are noticeable differences between CTs and Foci. While CFs may act as the sole Focus of a clause or sentence, CTs always occur with an additional Focus in the same sentence. This is because every Wh-question must have a Wh-expression in it.<sup>8</sup> Unlike Foci, CTs cannot correspond to the Wh-expression in the QUD they address. Instead, CTs occur in direct answers to sequential questions. These questions share their information structure and differ only in the CT constituent. For instance, the answer in (18) above most directly corresponds to the sequential questions in (19). This constitutes a direct answer as the CTs are discourse-given. The questioner

---

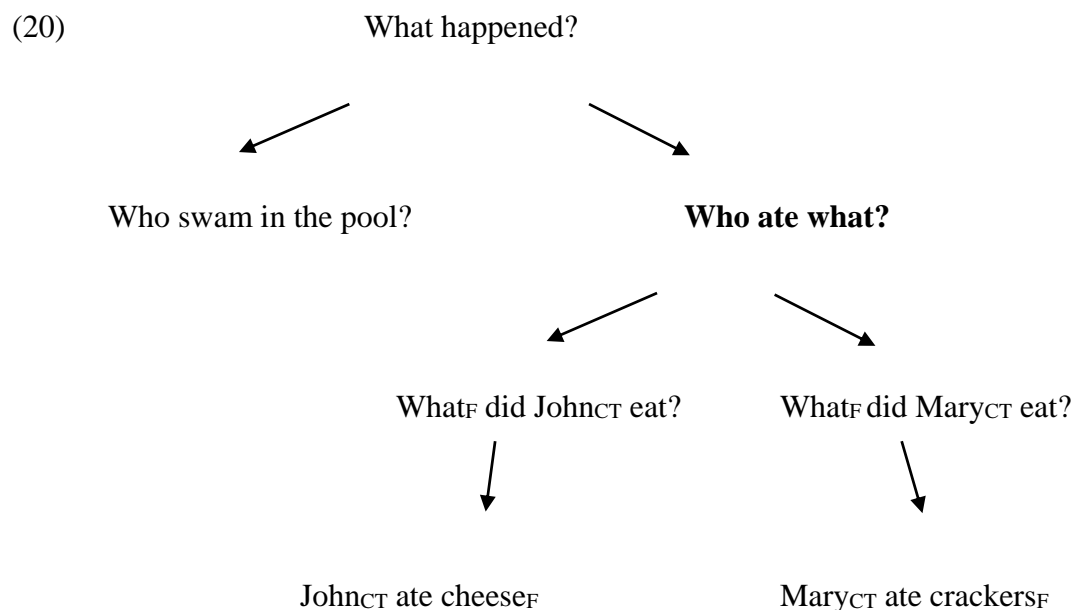
<sup>8</sup> Estonian Polar questions are not that much different from Wh-questions in this regard. In unbiased polar questions, we see the clause-initial interrogative particle *kas* ‘whether’, which can be taken as the constituent that Polarity Focus in answers corresponds to. An example is shown in (a) below.

- (a) Q: Kas<sub>F</sub> Mari kommi sõi?  
*whether Mari candy.PART ate.3S*  
‘Did Mari eat candy?’
- A: Mari [ei söönud]<sub>F</sub> kommi  
*Mari NEG eat.PTC candy.PART*  
‘Mari did NOT eat candy.’

does not have to draw additional inferences, e.g. that the set of the Addressee's brothers consists of Bill and Bob.

- (19) Q: What does Bill<sub>CT</sub> do for a living? What does Bob<sub>CT</sub> do for a living?  
A: Bill<sub>CT</sub> is an accountant<sub>F</sub> and Bob<sub>CT</sub> is a meteorologist<sub>F</sub>

The observation that CTs occur in information-structurally parallel sequential questions is formalized in terms of discourse trees in Büring (2003). In Büring's framework, CTs indicate a strategy of dividing a QUD into further subquestions. An example of a discourse tree is shown below in (20).



Each CT clause (e.g. 'John<sub>CT</sub> ate cheese<sub>F</sub>') offers a partial answer to the wider QUD ('Who ate what?'). Discourse conditions like Distinctiveness (21) explain the usage of CTs.

- (21) Distinctiveness – If two clauses contain CTs that act as each other's alternatives, the clauses must differ in their Focus values (Krifka, 1998).

Returning to the previous example, if both of the Addressee's brothers are accountants, then the QUD 'What do the Addressee's brothers do for a living?' is not divided into subquestions, as shown in example (22).

- (22) Q: What do your brothers do for a living?  
A: They<sub>T</sub> are accountants<sub>F</sub>

If, however, the two brothers have different professions, i.e. the two Topic-like elements are associated with distinct Foci, a sequence of parallel QUDs is appropriate, as seen in (23).

- (23) Q: What does Bill<sub>CT</sub> do for a living? What does Bob<sub>CT</sub> do for a living?  
A: Bill<sub>CT</sub> is an accountant<sub>F</sub> and Bob<sub>CT</sub> is a meteorologist<sub>F</sub>

As condition (21) states, the Foci must differ for a CT structure to be licensed. Namely, example (24) shows that a violation of Distinctiveness results in unnaturalness.

- (24) Q: What does Bill<sub>CT</sub> do for a living? What does Bob<sub>CT</sub> do for a living?  
A: # Bill<sub>CT</sub> is an accountant<sub>F</sub> and Bob<sub>CT</sub> is an accountant<sub>F</sub>

The use of CTs is seen as a discourse strategy when a Single-Focus answer is not felicitous (Büring, 2003). Much like the way that CT structures call for distinct Foci, distinct Foci associated with two Topic-like element call for the use of CTs. This is reflected in the shift from a single QUD ('What do x's brothers do for a living?') to a sequence of more specific sub-QUDs (What does Bill do for a living? What does Bob do for a living?). Note that, as seen in (18), this change

of QUD may be implicit. Once a speaker uses a CT structure in an answer, discourse participants accommodate the current QUD as being broken down into further subquestions.<sup>9</sup>

In Estonian, CTs naturally occur in a clause-initial position (25), much like non-contrastive Topics.

- (25) Villu<sub>CT</sub> on nimelt raamatupidaja<sub>F</sub> ja Robi<sub>CT</sub> on ilmataadlane<sub>F</sub>  
*Villu is namely book.keeper and Robi is weather.scientist*  
‘Villu is namely an accountant and Robi is a meteorologist’

Despite both occurring clause-initially, Topics and CTs have syntactic differences in Estonian as will be argued in Section 5.

### 3. Estonian Word Order

#### 3.1 The nature of V2 in Estonian

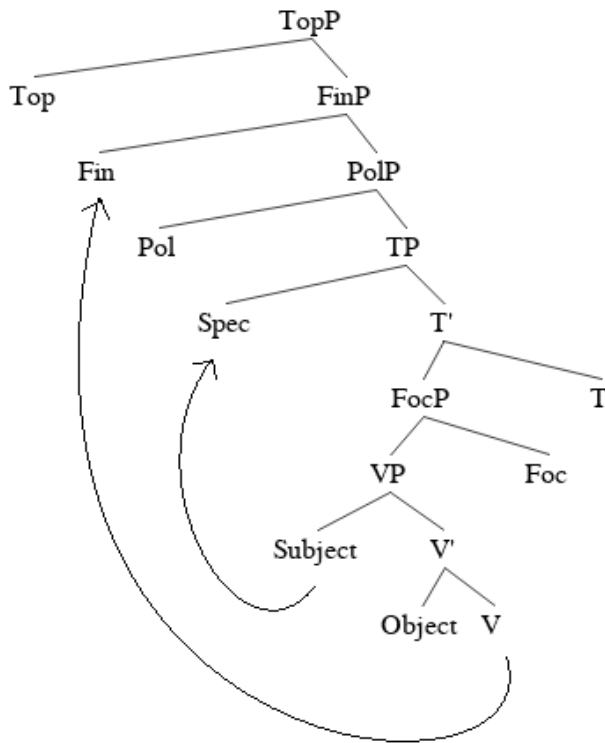
Estonian is a V2 language (Holmberg, 2015). This means that in typical declarative matrix clauses (among other clause types), the inflected verb occurs in the second position after the Topic of the clause (Ehala, 2006). I analyse Estonian V2 clauses as having the inflected verb in a high Finite verb position, with a single Topic position that, in certain kinds of clauses, must be filled by

---

<sup>9</sup> This ability for CT structures to shift the QUD is closely tied to the usage of CTs to shift discourse Topics. In order for a CT structure to be used, the CT must have some contextually salient alternative. For instance, when the question ‘Did you wipe your feet before entering?’ is answered with the English rise-fall-rise CT-intonation (Constant, 2012) ‘I<sub>CT</sub> did<sub>F</sub>’, the asker, worried for her rugs, would probably look around in panic for the contextually salient person who did NOT wipe their feet. The utterance ‘I<sub>CT</sub> did<sub>F</sub>’ is thus not just about the 1<sup>st</sup> person speaker.

a constituent. The simplified structure I assume for Estonian V2 clauses is shown in the diagram in (26).<sup>10</sup>

(26) Estonian V2-clauses



With a subject Topic, the V2 requirement in Estonian yields SVO word order, as can be seen in example (27).

<sup>10</sup> VPs are assumed to be head-final due to objects preceding verbs in infinitival and nominalized VPs. TPs are also assumed to be head-final, due to out-of-the-blue Wh-questions being verb-final, requiring a clause-final verbal position distinct from Focus. Determining the relative position between T and the low Focus position falls outside the scope of the present thesis. In the tree, the Focus is preliminarily placed in a VP-adjacent position, due to the cross-linguistic observation that rightwards movement cannot occur over long distance (T. Stowell, personal communication, September, 22, 2017)

- (27) Q: What did Mari do? SVO  
 A: Mari<sub>T</sub> andis Jaanile pastaka  
*Mari gave.3S Jaan.ALL pen.GEN*  
 ‘Mari gave Jaan a pen.’

OVS order, on the other hand, arises when the object is a Topic, as shown in (28).

- (28) Q: What happened to the pen? OVS  
 A: Pastaka andis Mari Jaanile  
*pen.GEN gave.3S Mari Jaan.ALL*  
 ‘Mari gave the pen to Jaan.’

However, the initial preverbal constituent is not always a Topic. Sometimes, this position may be occupied by a Nominative subject, even when it is a Focus. This is seen in (29), where the subject corresponds to the Wh-word in the question.

- (29) Q: Who fed the cat?  
 A: Mari<sub>F</sub> toit<sub>S</sub> kassi  
*Mari fed.3S cat.PART*  
 ‘MARI fed the cat.’

Object Foci, on the other hand, are not allowed in the clause-initial position, as shown by the infelicity of (30).

- (30) Q: Who did Mari feed?  
 A: # Kass<sub>F</sub> toit<sub>S</sub> Mari  
*cat.PART fed.3S Mari*  
 intended: ‘Mari fed the CAT.’



The special status of subjects as initial constituents in V2 clauses is also attested in Swedish (Holmberg, 2015). This can be explained if the preverbal position has multiple features associated with it, say [Topic] and [Nominative].<sup>11</sup> When an element that acts as a Topic in the discourse context is not Nominative, the syntactically high position of a Nominative subject may allow that constituent instead to be raised to the Topic position during linearization (see Fox & Pesetsky, 2005). Superiority effects and subject/non-subject asymmetries have previously been observed for Wh-movement, another kind of movement to the clausal left periphery (Pesetsky & Torrego, 2001). The ability of the preverbal position to attract elements based on multiple criteria or features ensures that V2 order is maintained even when a Topic is not available. This is exemplified with so-called Topic-drop in example (31). Here, the utterance is clearly about Mari, but there is no linguistic expression in the clause referring to her.

---

<sup>11</sup> Despite non-Nominative subjects of experiencer-type clauses (Erelt & Metslang, 2006) typically acting as Topics and occurring preverbally (a), they are significantly degraded as preverbal Foci when another Nominative element is present in the clause (b). This suggests that it is the Case feature of Nominative subjects that allows them to be raised, rather than subjecthood in a semantic sense.

- (a) Q: What's up with Mari?  
 A: Maril<sub>T</sub> valutab pea  
*Mari.ADE hurts head.NOM*  
 'Mari has a head ache' *lit. 'The head aches on Mari.'*
- (b) Q: Who has a head ache?  
 A: # Maril<sub>F</sub> valutab pea  
*Mari.ADE head.NOM hurts*  
 'MARI has a head ache'

- (31) Q: What's wrong with Mari?  
 A: Jänes on püksis  
*rabbit.NOM is pant.INE*  
 '[She] is scared' *lit.* '[She] has a rabbit in [her] pants.'

In (31), the Topic 'Mari' is dropped and the idiomatic 'rabbit' occurs in the clause-initial position. This yields V2 order in the clause. However, 'rabbit' does not meet our requirements for Topichood, as it is not referential. This is shown by the fact that it cannot be referred back to with a pronoun (32).

- (32) Q: What's wrong with Mari?  
 A: # Jänes on püksis. Sellel on teravad küüned.  
*rabbit.NOM is pant.INE it.ADE is sharp claws*  
 Intended: 'She has a rabbit in her pants. It has sharp claws.'

Examples (29) and (31) indicate that Nominative subjects have a special status when it comes to occurring preverbally. They can occupy this position regardless of their discourse status. The special status of Nominative subjects results in SVO being the so-called canonical word order. When there is no contextually provided Topic, the Nominative subject is raised to the preverbal Topic position. Canonical word order is identified as the least marked word order, occurring in out-of-the-blue contexts such as 'What happened?'

### 3.2 Contrastive ellipsis as a diagnostic for CTs and CFs

As we have seen, a single word order such as SVO may be used with a range of information structures. In written text where prosodic cues are not available, or in dealing with CTs and Foci which both receive intonational prominence, additional diagnostics are needed to disambiguate the status of the clause-initial subject. Gapping with CT and CF remnants to ellipsis (see Repp, 2009)

allows us to identify Foci, CTs and discourse-given elements. Following the terminology of Konietzko and Winkler (2010), I use the term CT ellipsis (CTE) for gapping with a CT remnant (33). In (33), Polarity Focus is expressed on the polarity particle *mitte*.

- (33) Jaan läks randa, **Mari<sub>CT</sub>** **mitter<sub>F</sub>** (CTE)  
*Jaan went beach.ILL Mari NEG*  
 'Jaan went to the beach, Mari didn't.'

Following Winkler (2005) I assume that CTE involves clausal coordination. This is evidenced by the fact that high adverbs are allowed to intervene between the remnant CT and the remnant Focus (34), placing the CT 'Mari' in a high, left-peripheral position. The second, bolded clause in (34) must thus be larger than a TP and the same is assumed for the first clause.

- (34) Jaan läks randa, **Mari<sub>CT</sub>** **kahjuks** **vist** **mitter<sub>F</sub>** (CTE)  
*Jaan went beach.ILL Mari unfortunately maybe NEG*  
 'Jaan went to the beach, unfortunately Mari probably didn't.'

Looking at CTE involving Polarity Focus in Estonian, it appears that all discourse-given material below Pol<sup>12</sup> is elided. The elision of the TP must follow the raising of the CT out of the TP to some left-peripheral, clause-initial position.

Analogously, what I refer to as CF ellipsis is a form gapping with a CF remnant. CF ellipsis is used in order to explicitly exclude a salient Focus alternative, resulting in a corrective reading. Here, *mitte* acts as a Focus particle and immediately precedes the CF it associates with (35). In CF

---

<sup>12</sup> The polarity particle is assumed to check clausal polarity features, placing it in or above the Polarity position shown in (26)

ellipsis, coordination has previously been shown to occur at the vP level (Repp, 2009, Konietzko, 2016) and this analysis is also assumed here.

- (35) Jaan läks randa, **mitte Mari<sub>F</sub>** (CFE)  
*Jaan went beach.ILL NEG Mari*  
 'Jaan went to the beach, not Mari.'

Because CTs and CFs must have contextually salient alternatives of the same type, the remnants in gapping allow us to identify CTs and CFs in the preceding antecedent clause.

### 3.3 Pronominal choice as a diagnostic for contrast

Before turning to instances where multiple lexical elements occur preverbally in declarative matrix clauses, let us devise another diagnostic for identifying the discourse status of constituents. The two varieties of contrastive ellipsis discussed above allow us to distinguish between CTs and CFs. Being a remnant of contrastive ellipsis suggests that the element is contrastive, i.e. has contextually salient focus alternatives. A more direct way to distinguish between contrastive and non-contrastive discourse objects, such as CTs and non-contrastive Topics, is by looking at which pronouns are used to refer to them.

Estonian has two forms of overt animate pronouns. As pointed out by Pajusalu (1997), short form (SF) pronouns like *ma* 'I' are used to refer to the most salient entity in the discourse context without invoking focus alternatives. SF pronouns are not able to bear stress and cannot associate with focus particles like *ka* 'also' as show in example (36).

- (36) \**Ka ma käisin rannas* SF  
*also IS.short went.IS beach.INE*  
 Intended: 'I, too, went to the beach.'

When a speaker wishes to contrast the referent to some other salient referent in the discourse, long form (LF) pronouns like *mina* 'I' are used (see Kaiser & Hiietam, 2004 for a discussion in English). Unlike SF pronouns, LF pronouns can associate with focus particles as shown in example (37).

- (37) Ka    mina            käisin            rannas    LF  
*also*   *IS.long*            *went.1S*            *beach.INE*  
'I, too, went to the beach.'

SF pronouns are thus appropriate Topics while LF pronouns are appropriate CTs (and CFs). In combination with contrastive ellipsis, we obtain in principle a three-way diagnostic distinction between non-contrastive Topics, CTs and CFs. This allows us to examine the discourse status of preverbal elements in non-V2 clauses.

### 3.4 Violating V2 in Estonian

Looking at example (38), the ungrammaticality of two preverbal SF pronouns suggests that only one non-contrastive Topic position is available preverbally, in accordance with the proposed analysis for V2 clauses in Estonian previously seen in (26).

- (38) \* **Ma**            **ta-ga**                    mängisin            jalgpalli    SF+SF  
*1S.short*            *3S-COM.short*            *played.3S*            *football.PART*  
'I played football with him/her'

However, if the clause-initial constituent is a LF pronoun, another SF pronoun is allowed to occur preverbally (39). This results in an unexpected V3 order.

- (39) **Mina ta-ga** mängisin jalgpalli (ja sina korvpalli) LF+SF  
*1S.long 3S-COM.short played.3S football.PART and 2S basketball.PART*  
 'I played FOOTBALL with him/her (and you basketball)'

Only discourse-given, i.e. non-Focus elements may occur between the clause-initial element and the verb.<sup>13</sup> This is seen in the ungrammaticality of (40) where the Focus 'with HIM/HER' is not allowed to be preverbal. In (40), the status of the Comitative form as a Focus is made clear by the second remnant of the CT ellipsis clause.

- (40) \* **Mina tema-ga<sub>F</sub>** mängisin jalgpalli (ja sina nendega<sub>F</sub>) LF+LF  
*1S.long 3S-COM.long played.3S football.PART and 2S.long 3PL-COM.long*  
 'I played football with HIM/HER (and you with THEM)'

The possibility of having additional preverbal material is hypothesized to be linked to the presence of a preverbal CT. The possible alternative that any kind of preverbal contrast licenses non-V2 order is disfavored due to the observation that SOV-type clauses are not congruent direct answers to subject Wh-questions (41).

- (41) Q: Who played football with him? SOV  
 A: # **Mina<sub>F</sub> ta-ga** mängisin jalgpalli  
*1S.long 3S-COM.short played.3S football.PART*  
 'I played football with him'

---

<sup>13</sup> See Lindström's (2005) corpus studies for evidence that the more discourse-given a constituent is, the more likely it is to occur preverbally. For instance, pronouns are more likely to occur preverbally than full DPs, and 1<sup>st</sup> and 2<sup>nd</sup> person pronouns are more likely to occur preverbally than 3<sup>rd</sup> person pronouns.

The unacceptability of (41) is not due to the clause-initial position not allowing Foci. As previously discussed in Section 3.1, subject Foci may occur preverbally in V2 clauses, as evidenced by the SVO version of (31) being felicitous in (42).

- (42) Q: Who played football with him? SVO  
 A: Mina<sub>F</sub> mängisin ta-ga jalgpalli  
*1S.long played.3S 3S-COM.short football.PART*  
 ‘I played football with him’

The hypothesis that it is the presence of clause-initial CTs that licenses multiple preverbal elements was tested in a naturalness rating experiment, reported in the next section.

#### 4. Naturalness Rating Experiment

As outlined above, if the remnant in contrastive ellipsis precedes the polarity particle *mitte*, the remnant is a CT. The matrix clause containing the TP antecedent is expected to contain a preverbal CT alternative. If the presence of a CT licenses multiple preverbal constituents, a verb-final word order such SOV should be possible in the antecedent clause to CTE. If the remnant in contrastive ellipsis follows the polarity particle, then the remnant is a CF. The matrix clause antecedent to CFE is expected to allow the CF alternative in the clause-final Focus position. The naturalness rating experiment presented here tests these predictions about word order in matrix and corresponding contrastive ellipsis clauses. Subject remnants were used in order to obtain matrix clauses where, in the absence of intonational cues, the non-canonical OVS order strongly suggests that the clause-final subject is a CF. As negative matrix clauses have previously been

noted to allow for verb-final order (Erelt, Erelt, & Ross, 1997), only affirmative matrix clauses were tested.

The following predictions were made regarding the relationship between information structure and word order. The contrasted constituent (CT or CF) was always the Nominative subject.

- (43) a. Following SOV matrix clauses, CTE is more natural than CFE
- b. Following OVS matrix clauses, CFE is more natural than CTE

Prediction (43a) follows from the observation in the previous section that preverbal CTs and not preverbal CFs allow multiple preverbal elements. If this is the case then SOV word order identifies the matrix clause subject as a CT. CTs require contextually salient alternatives of the same category, thus a following CTE clause with a CT subject is expected to be more compatible with the matrix clause than a CFE clause is.

Prediction (43b) follows from the proposal that Estonian has a clause-final Focus position. Subjects are typically attracted to the clause-initial position due to their high position in the clause. Being marked as a CF would allow the subject to move rightwards instead, to the clause-final Focus position. A CF subject in the matrix clause would be more compatible with a CF subject in the following ellipsis clause, compared to a CT subject.

These predictions rely on the assumption that CTs and CFs undergo distinct syntactic derivations in Estonian. The alternative hypothesis where a syntactic Contrast position is semantically underspecified holds of Finnish, a Finno-Ugric language closely related to Estonian (Vallduví & Vilks, 1998). In Finnish, both exhaustive CF elements and non-exhaustive CT



elements have been proposed to occur in a clause-initial ‘Kontrast’ position in the left periphery. As both CT subjects and CF subjects may occur clause-initially in Estonian, contrastive subject clauses provide an interesting test case for the reality of the distinction between CT clauses and non-CT clauses proposed in this thesis.

The details of the design and results of the experiment are reported below.

#### 4.1 Design and materials

The 3x2 experimental design crossed matrix Word Order (SVO, SOV or OVS) with Ellipsis Type (CTE or CFE). While SOV clauses were of main interest, the two V2 orders (SVO and OVS)<sup>14</sup>, were included in the experiment in order to validate the assumption that contrastive ellipsis is sensitive to the information structure of its antecedent clause. SVO, being the canonical order, was hypothesized to be compatible with both CT and CF subjects. OVS was predicted to be more compatible with a CF subject than a CT subject. The predictions for each of the three word orders, along with example sentences, are shown in Table 1.

Table 1. Predictions by conditions, along with example target sentences

	<b>SVO</b>	<b>OVS</b>	<b>SOV</b>
<b>CTE</b>	Karl aitas õpetajat, Pia mitte. <i>Karl helped teacher, Pia</i> NEG	Õpetajat aitas Karl, Pia mitte <i>Teacher helped Karl, Pia</i> NEG	Karl õpetajat aitas, Pia mitte. <i>Karl teacher helped, Pia</i> NEG
<b>CFE</b>	Karl aitas õpetajat, mitte Pia <i>Karl helped teacher,</i> NEG Pia	Õpetajat aitas Karl, mitte Pia <i>Teacher helped Karl, NEG</i> Pia	Karl õpetajat aitas, mitte Pia <i>Karl teacher helped,</i> NEG Pia
<b>prediction</b>	<b>CTE ≥ CFE</b>	<b>CFE &gt; CTE</b>	<b>CTE &gt; CFE</b>

<sup>14</sup> Subjects always occurred in Nominative case and objects in Partitive, Accusative or Allative case, so word order in the target sentences was unambiguous.

Every target sentence was preceded by a context sentence mentioning the subject of the matrix clause as well as the subject of the ellipsis clause. This was done in order to make the alternative set of CTs and CFs salient. In addition, the context sentences allowed for an interpretation where the object of the target sentence is discourse-old. This was done to allow the object to occur preverbally, whether in OVS or SOV clauses. To test possible contextual effects on word order and the naturalness of contrastive ellipsis, the context sentences were divided into two categories (see Table 2).

Table 2. Two types of context sentences used in the target items

Type	Equal Salience	Unequal Salience
Example	Kaisa and Peeter entered a busy streetcar with very few seats.	Karl heard that the teacher asked Pia for help with sorting the notebooks
Congruent QUD	Did Kaisa and Peeter get a seat?	Was it Pia who helped the teacher?
Congruent Answer	Peeter got a seat, Kaisa didn't.	It was Karl who helped the teacher, not Pia
Bias	CT	CF

Half of the items included an Equal Salience context and half of the items included an Unequal Salience context. It was hypothesized that conjunctive contexts, where the two subjects were made equally salient and were portrayed as having equally affected thematic roles, would facilitate a conjunctive QUD or a series of parallel QUDs<sup>15</sup>. Unequal salience contexts, on the other hand, were predicted to facilitate a CF reading when the more salient and affected participant is explicitly excluded from the set of Focus alternatives in the next clause, using CF ellipsis. While

---

<sup>15</sup> 'Did Kaisa and Peeter get a seat?' is here used as a shorthand of 'Did Kaisa get a seat? Did Peeter get a seat?', excluding the possible reading 'Did *both* Kaisa and Peeter get a seat?'

the context sentences were expected to be biasing towards either a CT or CF reading, the other reading was not predicted to be ruled out.

In order to further increase variety in the items, the contrastive subjects were either two professions (in 33% of the items) or a male and female proper name. The full paradigm is illustrated with an example quartet from the experiment in (44). The same context sentence preceded the target sentence in each of the six conditions.

- (44) Context: Karl kuulis, et õpetaja palus Pialt vihikute sorteerimiseks abi.  
'Karl heard that the teacher asked Pia for help with sorting the notebooks.'
- a. Karl aitas õpetajat, Pia mitte. (SVO + CTE)  
*Karl helped teacher, Pia NEG*  
'Karl helped the teacher, Pia didn't.'
- b. Karl aitas õpetajat, mitte Pia (SVO + CFE)  
*Karl helped teacher, NEG Pia*  
'Karl helped the teacher, (it was) not Pia.'
- c. Karl õpetajat aitas, Pia mitte. (SOV + CTE)  
*Karl teacher helped, Pia NEG*  
'Karl did help the teacher, Pia didn't.'
- d. Karl õpetajat aitas, mitte Pia (SOV + CFE)  
*Karl teacher helped, NEG Pia*  
'Karl did help the teacher, (it was) not Pia.'
- e. Õpetajat aitas Karl, Pia mitte (OVS + CTE)  
*Teacher helped Karl, Pia NEG*  
'It was Karl who helped the teacher, Pia didn't.'
- f. Õpetajat aitas Karl, mitte Pia (OVS + CFE)  
*Teacher helped Karl, NEG Pia*  
'It was Karl who helped the teacher, (it was) not Pia'

The materials consisted of 30 sextets following the pattern in (44); see Appendix A for a complete list of items.

#### 4.2 Participants and procedures

Native Estonian speaking volunteers (N=41) participated in the experiment over the Internet. The online experiment platform Ibox Farm (Drummond, 2012) was used to present the materials and to collect responses. The experiment took less than 20 min on average to complete. After a short guided practice, participants rated target sentences based on how natural they sounded in the given context, on a 7-point Likert scale as in (45) below.

(45) Merli is happy that Sven bought a new car.

Merli uses the car, Sven doesn't.

How natural does the second sentence sound in the context of the first?  
(Please use entire scale)

(Unacceptable) 1      2      3      4      5      6      7 (Completely natural)

Items were presented in a Latin square design, counterbalanced and randomly interspersed with 30 filler items and 5 nonsensical catch items (see Appendix A), for a total of 65 sentence pairs. The filler items were similar to the experimental items in describing real-life scenarios but did not contain CT or CF subjects in the second sentence. Participants were encouraged to assign ratings at their own pace, with every participant spending an average of at least 6 seconds per sentence pair. After rating each item, participants were presented with a screen asking them to press a button to continue. All materials and instructions were presented in Estonian.

### 4.3 Results and discussion

Data were excluded for participants who assigned a rating of 3 or above to more than one catch item (N=11) or whose mean rating of the filler items was below 5 (N=1), resulting in data from 29 participants being analyzed. Responses faster than 1000 ms were removed from the data. Each experimental item was left with at least 4 ratings per condition. Means and standard errors by conditions are shown in Figure 1.

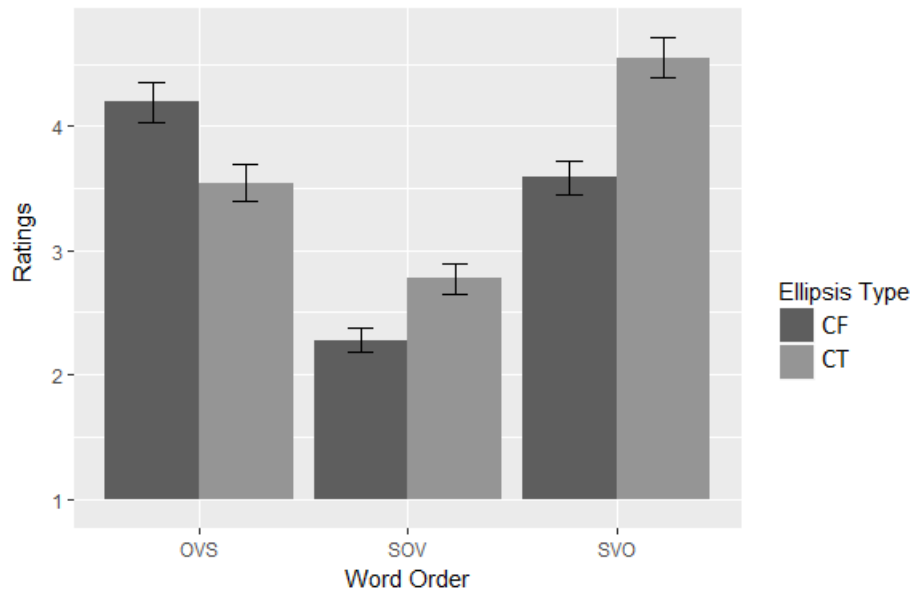


Figure 1. Means and standard errors of naturalness ratings to OVS, SOV and SVO matrix clauses followed by CT ellipsis and CF ellipsis.

As a whole, sentences with SOV matrix word order were judged as less natural than sentences with OVS or SVO order. Yet, the SOV sentences were judged as more natural when the subject of the matrix clause was identified as a CT rather than a CF by the following contrastive ellipsis clause. This suggests that while participants had a strong preference for V2 clauses, SOV clauses were not rejected outright as being ungrammatical.

Linear mixed effects models (Baayen, Davidson, & Bates, 2008) from the *lme4* package (Bates & Maechler, 2009) in R (R Development Core Team, 2016) were used to analyze the data, with sum coded fixed effects contrasts and maximal random effects structures (Barr, Levy, Scheepers, & Tily, 2013). This allowed me to control for as much of individual variation in the items and participants' judgements as possible. The models were compared using the *anova* function in the R base package by forward-fitting random effects and backward-fitting the fixed effects of Word Order and Ellipsis Type. The most complex model that accounted for significantly more variance than any simpler model included Word Order and Ellipsis Type interacting as fixed effects as well as by-subject and by-items random slopes and intercepts for Word Order and random slopes and intercepts by Context Type for Ellipsis Type. This model is spelled out in (46). A summary of the results of the model in (46) is shown in Table 3.

$$(46) \quad \text{rating} \sim \text{WO} * \text{ellipsis} + (1 + \text{WO} | \text{item}) + (1 + \text{WO} | \text{subject}) + (1 + \text{ellipsis} | \text{context}) + \epsilon$$

Table 3. Statistical findings for naturalness ratings crossing word order and ellipsis type in a linear mixed effects model. \* indicates statistical significance at the  $p < .05$  level.

	Estimate	Standard Error	t-value
(Intercept)	3.486	0.188	13.297*
WO(OVS)	0.375	0.094	3.987*
WO(SOV)	-0.959	0.104	-9.231*
Ellipsis(CF)	-0.140	0.163	-0.857
WO(OVS):Ellipsis(CF)	0.442	0.065	6.857*
WO(SOV):Ellipsis(CF)	-0.106	0.064	-1.665

As seen in Table 3, there is a main effect of Word Order in naturalness ratings. Compared to the grand mean, OVS sentences are judged to be more natural by 0.375 points and SOV sentences are rated as less natural by 0.959 points. This can be seen as an overall preference for V2 clauses, whether SVO or OVS. There is no main effect of Ellipsis Type in Naturalness ratings,

suggesting that differences in ratings do not arise from one type of ellipsis being regarded as more natural than the other. Rather, as predicted, there is a significant interaction between Word Order and Ellipsis Type in the two V2 conditions (SVO and OVS). This means that Ellipsis Type affected the subjects' judgements of Naturalness above and beyond more general word order preferences.

In order to further disentangle the interaction between Word Order and Ellipsis Type, *post hoc* comparisons were conducted using t-tests with Tukey's adjustment for multiple comparisons in the R package *lsmeans* (Lenth, 2017). The maximally complex model lending itself to this procedure included Word Order and Ellipsis Type as interactive fixed effects and by-subject and by-item random slopes for Word Order (47).

$$(47) \quad \text{rating} \sim \text{WO} * \text{ellipsis} + (1 + \text{WO} | \text{item}) + (1 + \text{WO} | \text{subject}) + \varepsilon$$

The random slope of Ellipsis Type for Context Type had to be excluded from the model to increase degrees of freedom for the *post hoc* test, despite the more complex model offering a significantly better fit for the data ( $\chi^2=15.35$ ,  $p<.01$ ). The relevant *post hoc* comparisons are reported in Table 4.

Table 4. Post-hoc comparisons for a linear mixed effects model of naturalness ratings crossing word order and ellipsis type. \* indicates statistical significance at the  $p<.05$  level.

	Contrast	Estimate	Standard Error	t-value
i	SVO+CFE – SVO+CTE	-0.953	0.158	-6.025*
ii	SOV+CFE – SOV+CTE	-0.489	0.158	-3.101*
iii	OVS+CFE – OVS+CTE	0.608	0.159	3.822*
iv	OVS+CFE – SVO + CFE	0.575	0.173	3.327*
v	OVS+CTE – SVO + CTE	-0.987	0.173	-5.711*
vi	SOV+CTE – SVO + CTE	-1.775	0.187	-9.491*

As seen in Table 4, SVO word order is judged to be significantly more natural when the remnant

subject is a CT than when it is a CF (i). The opposite holds for OSV word order which is judged to be significantly more natural with a CF subject than with a CT subject (iii), as predicted. Not surprisingly, the SVO word order is more natural than the OSV word order when the remnant subject is a CT (v), showing a preference for clause-initial Topics. With a CF subject, the OSV word order is rated as more natural than the SVO word order (iv), showing a preference for clause-final Focus. Most importantly for the present paper, despite the fact that the canonical SVO word order is preferred over the non-V2 SOV word order (vi), the naturalness of the SOV word order is significantly increased when the subject is a CT rather than a CF (ii).

In addition, it was predicted that Word Order and Ellipsis Type would interact with Context Type, i.e. the relative salience of the two characters in the context sentence. The means and standard errors for each of the 6 conditions are shown by context type in Figure 2.

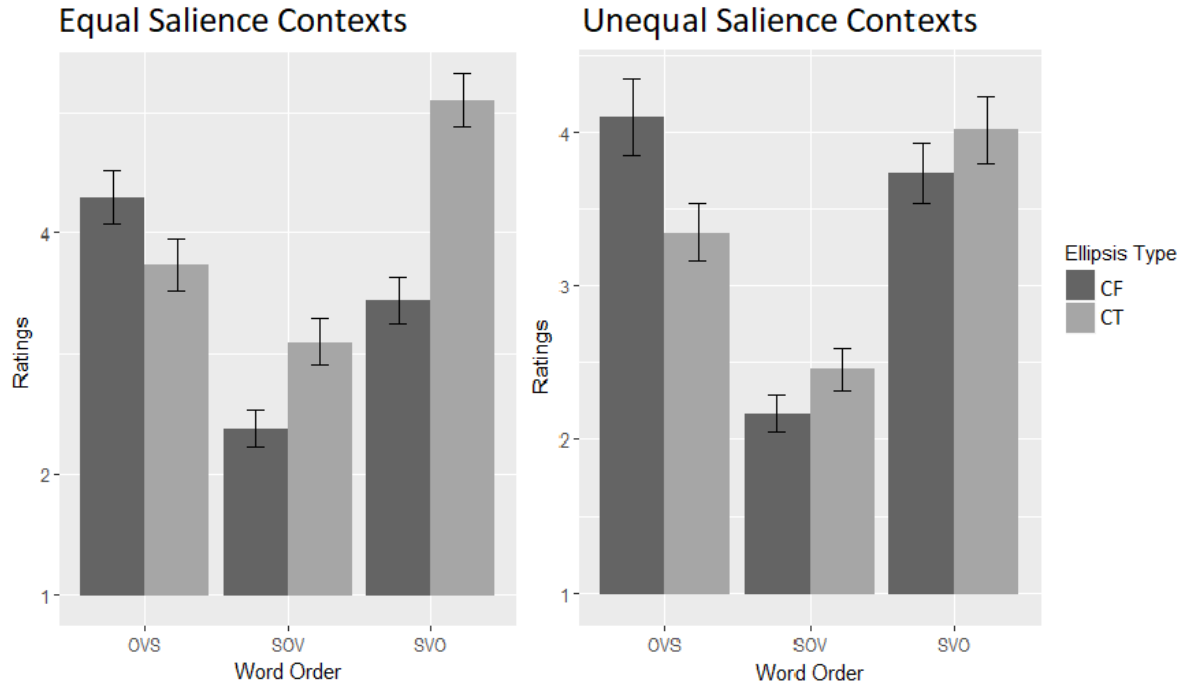


Figure 2. Means and standard errors of naturalness ratings to OVS, SOV and SVO matrix clauses followed by CT ellipsis and CF ellipsis, under two kinds of contexts



To recapitulate, the presentation of two equally salient characters using a coordination in the context sentence was thought to facilitate a CT interpretation for the two subject. This would increase the naturalness of the SOV word order and CTE, both hypothesized to involve a CT constituent. The presence of a salient and a less salient character in the context was predicted to increase the naturalness of the CF reading, wherein expectations set up by the context would be explicitly refuted. This was predicted to facilitate the corrective use of CFE and OVS.

What is most striking in Figure 2 is the extent to which different word orders constrain information structure interpretation. The effect of context appears to be larger with SVO clauses than OVS and SOV clauses. This is compatible with the claim that SVO clauses are the canonical word order – SVO clauses can be used in a wide range of discourse contexts, whereas other orders are limited to contexts that are compatible with the information structure imposed by the context. The predicted contextual effects are not seen in as clearly in the more structurally constraining OVS and SOV clauses. Following OVS clauses, CFE is preferred over CTE regardless of context. A preference for CTE is consistently seen in SOV clauses.

The previous models discussed here have included Context Type as a random effect where possible, controlling for any differences arising from the relative discourse salience of the matrix subject and remnant subject. In order to investigate the effects of Context Type across Word order and Ellipsis Type conditions, additional linear mixed effects models were set up. Here, Context Type was treated as a fixed effect interaction term. With Item and Subject as random intercepts, Context Type was found to interact with Ellipsis Type ( $\chi^2=9.08$ ,  $p<.01$ ). Namely, following equal salience contexts, mean ratings for CTE sentences ( $M=3.96$ ) were higher than mean ratings to CFE

sentences ( $M=3.37$ ), although unequal salience contexts did not have the predicted effect<sup>16</sup>, as mean ratings to CFE sentences ( $M=3.34$ ) were similar to mean ratings to CTE sentences ( $M=3.28$ ). The effect of context is reflected in the random effect structure of the model in (46). Despite an apparently large contextual effect in SVO clauses, Context Type did not significantly interact with Word Order. The lack of an interaction between Context Type and Word Order suggests that the participants mainly relied on the ellipsis clause when determining the intended information structure of the matrix clause, rather than the context. Word order in the ellipsis clause was used as a reliable cue in determining the intended information structure of the clause. This points to CT and CF ellipsis being useful tools in future experiments where context on its own may not be sufficiently constraining in communicating the intended information structure.

Importantly, the naturalness rating study provides experimental evidence that information structure plays an important role in deviations from Estonian V2 word order. SOV clauses are judged to be more natural by native speakers when the subject is a CT rather than a CF. While the information structure in SVO clauses, as judged by the differences in naturalness ratings for CTE and CFE constructions, appears to be largely dictated by discourse context, this is not the case for SOV clauses. This suggests that the availability of multiple preverbal constituents necessarily constrains the comprehenders' interpretation of the information structure of the clause, leaving little work to do to biasing cues in the discourse context. This special status of non-V2 clauses like SOV raises questions about their underlying structure.

---

<sup>16</sup> This could be due to unequal salience contexts being more complex as more thematic roles are involved, yielding lower Naturalness ratings overall.

## 5. The Syntax of Non-V2 Clauses

Looking at word order only takes us so far. A range of syntactic analyses could account for any particular linear ordering of elements. The observation that multiple preverbal elements are allowed in CT clauses could arise through one of the ways listed in (48).<sup>17</sup>

- (48) a. In CT clauses, an additional CT position is available above the Topic position
- b. In CT clauses, discourse-given material that is not moved out of the VP<sup>18</sup> is allowed to pied-pipe with the verb

In this section, I provide evidence against the simpler analysis in (48a), leaving a VP-raising analysis (48b) as a viable account for CT clauses in Estonian to be developed in future research.

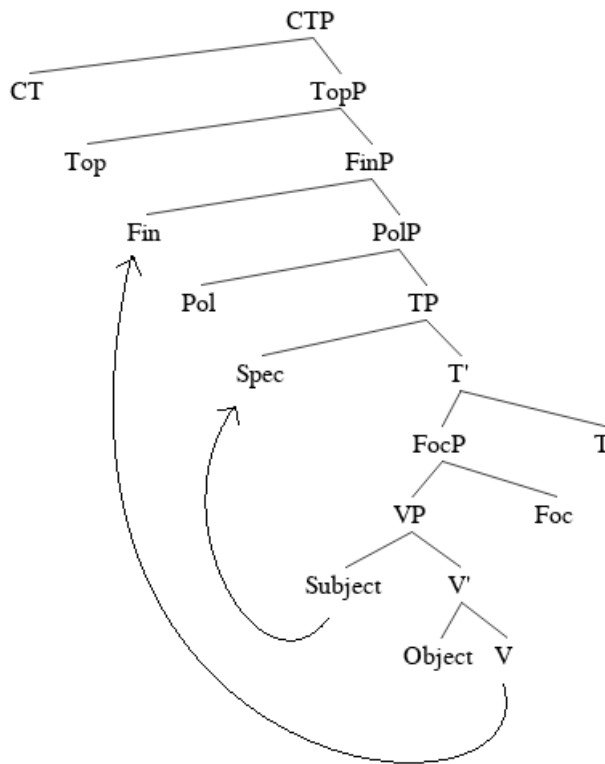
Previous literature on Estonian word order has alluded to the possibility of non-V2 orders arising from the language having multiple preverbal Topics (Lindström, 2005; Henk, 2009). In particular, V3 clauses with a clause-initial CT might be analyzed as containing an additional CT (or more generally, Contrast) position above a single preverbal Topic, as shown in (49). As we will see, this analysis does not provide an adequate account of Estonian CT structures.

---

<sup>17</sup> The possibility that the preverbal material forms a single Topic phrase with obligatory Focus-marking on the first element also remains open. This would constitute a pied-piping analysis where a Focus-marked Topic (i.e. CT) moves as a part of a larger constituent, such as VP, and would require a Focus-like projection at the left edge of the VP in addition to the presently proposed right-hand Focus projection. This analysis does, however, not offer a straightforward account for the differences experimentally observed between subject CF and subject CT clauses.

<sup>18</sup> Either to a Focus position by rightwards movement, or to higher Spec positions through leftwards movement

(49) A multiple (Contrastive) Topic analysis of V3 clauses



A specified clause-initial Contrastive<sup>19</sup> position is not cross-linguistically unattested. Finnish has been proposed to have one (Vallduví & Vilkkuna, 1998). Yet, these are differences between the two languages. The present naturalness rating experiment showed that not any kind of preverbal contrastive element is associated with the possibility of V3 order in Estonian. Even though CF subjects are allowed to occur preverbally in Estonian, SOV clauses with a CF subject are judged to much less natural than SOV clauses with CT subjects.<sup>20</sup> This finding speaks to the special status of CTs in Estonian.

<sup>19</sup> Focus-marked with a set of contextually salient alternatives

<sup>20</sup> This observation provides evidence for the assumption that multiple features (Topic, Nominative) compete for the preverbal Topic position in V2 clauses. If Nominative but non-

An analysis like that shown in (49) is challenged by the observation that the material allowed to intervene between a CT and the finite verb differs from the clause-initial Topic in V2 clauses. Firstly, multiple elements are allowed to intervene between a CT and the verb, as exemplified in (50).

- (50) Koolis<sub>CT</sub>    ma    ta-st            ei    mõelnud,    aga    kodus<sub>CT</sub>    küll  
*school.INE*    *IS*    *3S-ELA*            *NEG*    *think.PTC*    *but*    *home.INE*    *AFF*  
 ‘At school I didn’t think about him/her but at home I did.’

Yet, as seen in (51) below, the same elements cannot occur preverbally when the CT constituent is removed.

- (51) \*Ma    ta-st            ei    mõelnud  
*IS*    *3S-ELA*            *NEG*    *think.PTC*  
 ‘I didn’t think about him/her’

In this particular case, two short form pronouns may not precede the inflected verb. As discussed in Section 3.3, short form pronouns may refer to discourse-old Topic entities but cannot have Focus alternatives. The inability of two non-Contrastive and Topic-like expressions to occur clause-initially suggests that V3 clauses cannot be analysed as V2 clauses with an additional CT position above a single Topic.

Secondly, some non-referential elements may occur between a CT and the inflected verb. This suggests that a preverbal Topic position does not account for all the data as Topics are due to their expression of aboutness necessarily referential. An example of non-Topics intervening

---

Topical CF subjects were raised to a position distinct from Topic, we might expect any Topic-marked element to be able to raise to the preverbal Topic position, yielding V3 word order. This is not the case.

between a CT and the inflected verb is shown with the idiomatic expression ‘rest after a heavy meal’, literally ‘to let bread into bone’ in (52).

- (52) Mar<sub>ICT</sub> leiba luusse lasi, aga Jaan<sub>CT</sub> mitte  
*Mari bread.PART bone.ILL let.3S but Jaan NEG*  
 ‘Mari TOOK a break (to ease digestion) but Jaan didn’t.’

The idiom ‘let bread into bone’ may be used regardless of whether any bread was consumed, so ‘bread’ is neither referential nor a Topic.

These two examples suggest that non-V2 word order in CT clauses cannot be explained by a single intervening position between the CT and the finite verb. If such a position were to exist, the data indicate that it would not require referentiality. While some CT clauses may involve a preverbal Topic, the [CT [Top [Fin]]] analysis cannot account for all the CT clause data.

The alternative analysis for CT structures posed in (48b) would involve the pied-piping (see Ross, 1967) of VP-internal (or TP-internal) material along with the inflected verb. As Estonian VP and TP are both hypothesized to be head-final (see 26), any pied-piped material would, of course, maintain its preverbal position. This would explain the observed V3 and V-final orders.

Some evidence for a pied-piping analysis is provided by the behaviour of verbal particles in CT-clauses and non-CT clauses. The presence of a clause-initial CT appears to license the raising of verbal particles like *ära* ‘away’ (53)

- (53) Mar<sub>ICT</sub> ära sõi oma koogi, aga Jaan<sub>CT</sub> mitte  
*Mari away ate.3S own.GEN cake.GEN but Jaan NEG*  
 ‘Mari ate (all of) her cake but Jaan didn’t (eat his)’

Verbal particles like *ära* are not allowed to undergo raising or pied-piping in the absence of a clause-initial CT. In example (54) below, V2 clauses are shown with Polarity Focus<sup>21</sup> without CT, for a direct comparison with (53). When the particle is stranded in its original low position, the clause is grammatical and felicitous in the context provided (54a). However, example (54b) where the particle precedes the finite verb is infelicitous as conditions for CT are not met in the context.

- (54) Q: Did Mari eat her cake?
- a. Jah, Mari **sõi** oma koogi **ära**  
*yes Mari ate.3S own.GEN cake.GEN away*  
 'Yes, Mari ate (all of) her cake'
- b. # Jah, Mari **ära** **sõi** oma koogi  
*yes Mari away ate.3S own.GEN cake.GEN*  
 'Yes, Mari ate (all of) her cake'

These data point to the syntactic structure of Estonian CT clauses being significantly different from single-Topic V2 clauses. The simple [CT [Top [Fin]]] analysis cannot be maintained. Further research is needed in order to explore the derivational mechanism and motivation behind these non-V2 clauses in Estonian.<sup>22</sup>

## 6. Conclusion

Exploring aspects of V2 and instances where this preferred word order is deviated from, the present thesis has made a case for distinguishing between clauses with and without CTs in Estonian. It was shown that Estonian declarative matrix clauses systematically differ in whether

---

<sup>21</sup> Being congruent answers a polar question

<sup>22</sup> See Kaps (in preparation) for a more detailed syntactic analysis of CT-clauses in Estonian.

they allow multiple preverbal elements or require a single preverbal Topic. Subject CT clauses allow for V3 and verb-final orders while subject CF clauses do not. The distinction between left-peripheral CFs and CTs, as observed in the naturalness rating experiment, raises theoretical questions about the mechanisms underlying Topicalization, or movement to a left-peripheral Topic position. The present findings suggest that we cannot syntactically treat CTs as Focus-marked elements that have undergone Topicalization because similar movement is observed in subject CFs in Estonian. The observations about the behavior of left-peripheral contrast in Estonian differ in interesting ways from Finnish, where a clause-initial Kontrast position has been proposed to be semantically underspecified (Vallduví & Vilkuna, 1998). Further comparisons of the role of contrast in the syntax of these two related languages would help to deepen our understanding of the interplay between information structure and syntactic derivation.



## Appendix A: Naturalness rating materials

### Experimental items

Items 2-30 are presented in the *SVO + CTE* condition here. Other conditions were derived by word order permutations of the same target sentences, as seen in Item 1 below.

1. *SVO + CTE*

Turvamees nägi, et vanaproua palus politseinikult suure liiklusega tee ületamiseks abi.

Turvamees aitas vanaprouat, politseinik mitte.

'A security guard<sup>23</sup> saw that an old lady asked a policeman for help with crossing a busy road. The security guard helped the old lady, the policeman didn't.'

*SVO + CFE*

Turvamees nägi, et vanaproua palus politseinikult suure liiklusega tee ületamiseks abi.

Turvamees aitas vanaprouat, mitte politseinik.

'A security guard saw that an old lady asked a policeman for help with crossing a busy road. The security guard helped the old lady, (it was) not the policeman.'

*SOV + CTE*

Turvamees nägi, et vanaproua palus politseinikult suure liiklusega tee ületamiseks abi.

Turvamees vanaprouat aitas, politseinik mitte.

'A security guard saw that an old lady asked a policeman for help with crossing a busy road. The security guard did help the old lady, the policeman didn't.'

*SOV + CFE*

Turvamees nägi, et vanaproua palus politseinikult suure liiklusega tee ületamiseks abi.

---

<sup>23</sup> The English translations indicate definiteness of nominals, which is not grammatically encoded in Estonian.

Turvamees vanaprouat aitas, mitte politseinik.

'A security guard saw that an old lady asked a policeman for help with crossing a busy road. The security guard did help the old lady, (it was) not the policeman.'

*OVS + CTE*

Turvamees nägi, et vanaproua palus politseinikult suure liiklusega tee ületamiseks abi.

Vanaprouat aitas turvamees, politseinik mitte.

'A security guard saw that an old lady asked a policeman for help with crossing a busy road. It was the security guard who helped the old lady, the policeman didn't.'

*OVS + CFE*

Turvamees nägi, et vanaproua palus politseinikult suure liiklusega tee ületamiseks abi.

Vanaprouat aitas turvamees, mitte politseinik.

'A security guard saw that an old lady asked a policeman for help with crossing a busy road. It was the security guard who helped the old lady, (it was) not the policeman.'

2. Annika ütles Erikule, et peab ruttu enne külaliste saabumist köögis korda looma. Erik koristas kööki, Annika mitte.

'Annika told Erik that she has to tidy up the kitchen before the guests arrive. Erik cleaned the kitchen, Annika didn't.'

3. Ain leidis viimasel hetkel oma referaadi kirjutamiseks raamatu, mida ta ka Evale näitas. Eva luges raamatut, Ain mitte.

'In the last moment, Ain found a book for writing his summary on and showed it to Eva too. Eva read the book, Ain didn't.'

4. Sekretär märkas, et koristaja leidis pärast tööpäeva lõppu koridorist ülemuse käekoti, mis tuli kiiresti tagastada. Koristaja helistas ülemusele, sekretär mitte.

'A secretary noticed that a cleaner found the boss's handbag in the hallway at the end of the work day and it had to be returned quickly. The cleaner called the boss, the secretary didn't.'

5. Karl kuulis, et õpetaja palus Pialt vihikute sorteerimiseks abi. Karl aitas õpetajat, Pia mitte.

'Karl heard that the teacher asked Pia for help with sorting the notebooks. Karl helped the teacher, Pia didn't.'

6. Lauri ütles Iidale, et leidis külmikust kahtlase välimusega jogurti, kuid otsustas seda enne ära viskamist proovida. Iida maitstes jogurtit, Lauri mitte.

'Lauri told Iida that he found a suspicious looking yoghurt in the fridge but that he decided to try it before throwing it out. Iida tasted the yoghurt, Lauri didn't.'

7. Ema märkas, et keegi oli moosipurgi kallal käinud ja heitis Liina pilgu all Ivole süüdistava pilgu. Liina sõi moosi, Ivo mitte.

'Mother noticed that someone had gotten their hands on the jam jar and gave Ivo an accusing look in front on Liina. Liina ate the jam, Ivo didn't.'

8. Hans kuulis, et turist proovis Martalt värskaid talumune osta. Marta müüb mune, Hans mitte.

'Hans heard that a tourist was trying to buy fresh eggs from Marta. Marta sells eggs, Hans doesn't.'

9. Kaisa ja Peeter sisenesid trammi, kus oli palju inimesi ja väga vähe isteruumi. Peeter võttis istet, Kaisa mitte.

'Kaisa and Peeter entered a busy streetcar with very few seats. Peeter took a seat, Kaisa didn't.'

10. Müüja ja klient proovisid kassautomaadist apelsinide tootekoodi leida, kuid sisestasid lõpuks vaid hinna. Klient mäletas hinda, müüja mitte.

‘A cashier and a client were trying to find the product code for oranges from the register, but eventually only entered the price. The client remembered the price, the cashier didn’t.’

11. Markus näitas Hellele räpase olemisega kodutut kutsikat, kelle ta tänavalt leidis. Helle paitas kutsikat, Markus mitte.

‘Markus showed Helle a dirty looking homeless puppy that he found on the street. Helle petted the puppy, Markus didn’t.’

12. Toomas ja Sirje ostsid kumbki lotopileti, lootes suuri summasid võita. Sirje võitis raha, Toomas mitte.

‘Toomas and Sirje bought a lottery ticket each, hoping to win a big sum. Sirje won money, Toomas didn’t.’

13. Maanteeametnik hoiatas Tiinat ja Sandrit, sest Sandril tundus kuhugi kiire olevat. Tiina ületas kiirust, Sander mitte.

‘A highway officer warned Tiina and Sander, because Sander appeared to be in a hurry. Tiina was speeding, Sander wasn’t.’

14. Viiu ütles Urmasele, et nägi piletisabas meest, kes ootamatult minestas. Urmas kutsus kiirabi, Viiu mitte.

‘Viiu told Urmas that she saw a man suddenly faint in the ticket line. Urmas called an ambulance, Viiu didn’t.’

15. Urve ja Ville pisikeses korteris elab kolm kassi, mis ei ole just mõlema meelest ideaalne. Ville armastab kasse, Urve mitte.

‘Three cats live in Urve and Ville’s small apartment and they’re not both exactly happy about it. Ville loves the cats, Urve doesn’t.’

16. Merlil on hea meel, et Sven uue auto ostis. Merli kasutab autot, Sven mitte.

‘Merli is happy that Sven bought a new car. Merli uses the car, Sven doesn’t.’

17. Vanaema kiitis Sofit ja Mihklit lillepeenra rohimise eest, kuigi ei näinud, kes kui kõvasti töötas. Sofi rohis peenart, Mihkel mitte.

‘Grandma praised Sofi and Mihkel for weeding the flowerbed, although she didn’t see who put in how much effort. Sofi weeded the flowerbed, Mihkel didn’t.’

18. Kaia kuulis, et maakler näitas Reinule uusi trendikaid kortereid. Kaia ostis korteri, Rein mitte.

‘Kaia heard that a real estate agent showed Rein new trendy apartments. Kaia bought an apartment, Rein didn’t.’

19. Boris ja Reelika alustasid enne suve uut dieeti, kuid magusast loobumine osutus keeruliseks. Boris langetas kaalu, Reelika mitte.

‘Boris and Reelika started a new diet before summer but giving up sugar turned out to be difficult. Boris lost weight, Reelika didn’t.’

20. Põrandalt leiti nimetähega kuldkest, mida pakuti nii Eedule kui Ellenile. Eedu kannab ehteid, Ellen mitte.

‘A gold necklace with an initial was found on the floor and offered to both Eedu and Ellen. Eedu wears jewelry, Ellen doesn’t.’

21. Birgit ja Rauno läksid paarinädalasele puhkusele ning tagasi tulles oli nende postkast ajakirjadest umbes. Rauno tellib ajakirju, Birgit mitte.

‘Birgit and Rauno went on holiday for a couple of weeks and when they came back the mailbox was stuffed with magazines. Rauno orders magazines, Birgit doesn’t.’

22. Rita teadis, et Kaido köögis elas pikemat aega hiir, ning et Kaidot kahtlustati hiirele toidu välja jätmises. Rita toitis hiirt, Kaido mitte.

‘Rita knew that a mouse had been living in Kaido’s kitchen for a long time and that people were suspecting Kaido of leaving food out for the mouse to eat. Rita was feeding the mouse, Kaido wasn’t.’

23. Leod ja Allit kutsuti pärast koosolekut pidulikule õhtusöögile, kuid nad olid mõlemad teksades. Leo vahetas riideid, Alli mitte.

‘Leo and Alli were both invited to a formal dinner after a meeting but they were both wearing jeans. Leo got changed, Alli didn’t.’

24. Lilli ja Andres arutasid, kas koerad on ohtlikud või mitte. Andres kardab koeri, Lilli mitte.

‘Lilli and Andres were discussing whether dogs were dangerous or not. Andres is afraid of dogs, Lilli isn’t.’

25. Kohtunik ja prokurör kuulasid tunnistajat, kes oli närvis ja ajas pidevalt kuupäevi segamini. Kohtunik uskus tunnistajat, prokurör mitte.

‘A judge and a prosecutor were listening to a witness who was nervous and kept getting dates mixed up. The judge believed the witness, the prosecutor didn’t.’

26. Ülikool kutsus matemaatikut ja füüsikut sissejuhatavaid geomeetrialoenguid andma. Füüsik õpetas geomeetriat, matemaatik mitte.

‘A university invited a mathematician and a physicist to give introductory lectures in geometry. The physicist taught geometry, the mathematician didn’t.’

27. Poliitik seletas ajakirjanikule ühe seaduse sisseviimist, mille tagajärjed oleksid võinud olla katastroofilised. Ajakirjanik mõistis tagajärgi, poliitik mitte.

‘A politician was explaining to a journalist the implementation of a law that could have had catastrophic consequences. The journalist understood the consequences, the politician didn’t.’

28. Näitleja ja lavastaja tahtsid peategelase kostüümi muuta, kuid uus kostüüm saabus alles viimasel hetkel. Lavastaja proovis kostüümi, näitleja mitte.

‘An actor and a director wanted to change the costume of the main character but the new costume arrived only at the last minute. The director tried the costume on, the actor didn’t.’

29. Trepikojas juhtunud avariid tulid hindama elektrik ja torumees, kes tegid probleemi uurimisel koostööd. Torumees kontrollis elektriohutust, elektrik mitte.

‘An electrician and a plumber came to assess the damage in the hallway and they worked together to solve the problem. The plumber checked electrical safety, the electrician didn’t.’

30. Arhitekt ja aednik arutasid linnaväljakule purskkaevu ehitamise plaani. Aednik joonistas kavandi, arhitekt mitte.

‘An architect and a gardener discussed the plan of constructing a fountain on the main square. The gardener sketched a plan, the architect didn’t.’

### **Catch items**

1. Teatri uue hoone avamiseks palgati orkester ja tantsutrupp. Vihma ja teatrit kolm päeva järjest sadas.

‘An orchestra and dance troupe were hired for the opening of the theatre’s new building.  
Rain and theater three days in a row rained.’<sup>24</sup>

2. Lapsed käisid loomaaias, kus nägid palju erinevaid loomi. Maksuamet sai jälile ja kaelkirjak samuti.

‘The children went to the zoo and saw many different animals. The IRS found out and so did the giraffe.’

3. Ilmatüdruk ajas otse-eetris kuupäevad segamini ja lubas nädalavahetuseks pilvitut ilma. Soodsamaid pileteid lubas nädalavahetuseks ilmatüdruk.

‘The weather girl got the dates mixed up on live TV and said that the weekend would be sunny. It was the weathergirl who promised cheaper tickets for the weekend.’

4. Kontoritarkvaras avastati viga ning kahtlustati häkkimist. Kartuleid hakkis ülemus erilise hoolega.

‘An error was discovered in the office software and hacking was suspected. The potatoes were hacked by the boss with particular care.’

5. Aknapesija teatas, et jaanuarist alates tuleb kõrgemate korruste eest rohkem maksta. Aknapesija ei ulatunud keldriakendeni ja lõi need rusikaga kildudeks.

‘The window washer announced that starting from January, higher floors would be charged extra. The window washer couldn’t reach the basement windows and shattered them with her fist.’

---

<sup>24</sup> Non-canonical word order in the translation reflects a marked word order in the target sentence.



## Bibliography

- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390–412.
- Barr, D., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278.
- Bates, D., & Maechler, M. (2009). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-31.
- Büring, D. (2003). On D-trees, beans, and B-accent. *Linguistics and philosophy*, 26(5), 511-545. Chicago.
- Büring, D. (2016). *Intonation and Meaning*. Oxford University Press.
- Constant, N. (2012). English rise-fall-rise: A study in the semantics and pragmatics of intonation. *Linguistics and Philosophy*, 407-442.
- Drummond, A. (2012). *Ibex farm*. Retrieved from [http://spellout.net/latest\\_ibex\\_manual.pdf](http://spellout.net/latest_ibex_manual.pdf)
- Ehala, M. (2006). The word order of Estonian: implications to universal language. *Journal of Universal Language*, 7(1), 49-89.
- Erelt, M., Erelt, T., & Ross, K. (1997). *Eesti keele käsiraamat*. Tallinn: Eesti Keele Sihtasutus.
- Erelt, M., & Metslang, H. (2006). Estonian clause patterns—from Finno-Ugric to standard average European. *Linguistica Uralica*, 42(4), 254-266.
- Fox, D., & Pesetsky, D. (2005). Cyclic linearization of syntactic structure. *Theoretical linguistics*, 31(1-2), 1-45.
- Henk, P. (2009). Information Structure of Estonian-Compared to Finnish and Hungarian. *Cognitive Science*.
- Holmberg, A. (2015). Verb second. Theory and analysis. *An international handbook*. Berlin: Mouton de Gruyter, 242-283.
- Holmberg, A. (2016). *The syntax of yes and no*. Oxford University Press.
- Horvath, J. (1995). Structural focus, structural case, and the notion of feature-assignment. In *Discourse Configurational Languages* (pp. 28-64). Oxford University press.
- Kaiser, E., & Hiietam, K. (2004). A comparison of the referential properties of third person pronouns in Finnish and Estonian. *Nordlyd*, 31(4).

- Kaps, M. (in preparation). *Processing information structure: A case study of Contrastive Topics in Estonian*. (Doctoral dissertation, UCLA).
- Konietzko, A., & Winkler, S. (2010). Contrastive ellipsis: Mapping between syntax and information structure. *Lingua*, 120(6), 1436-1457.
- Koopman, H. J., & Szabolcsi, A. (2000). *Verbal complexes* (No. 34). MIT Press.
- Krifka, M. (1998). Additive particles under stress. In *Semantics and Linguistic Theory* (Vol. 8, pp. 111-128).
- Lee, C. (2003). Contrastive topic and/or contrastive focus. *Japanese/Korean Linguistics*, 12, 352-364.
- Lenth, R. (2017). *lsmeans: Least-Squares Means*. R package version 2.26-3.
- Lindström, L. (2005). *Finiitverbi asend lauses. Sõnajärg ja seda mõjutavad tegurid suulises eesti keeles*. Tartu.
- Pajusalu, Renate. (1997). 'Eesti pronomeneid I. Ühiskeele see, too ja tema/ta.' *Keel ja Kirjandus*. 24-30, 106-115.
- Pesetsky, D., & Torrego, E. (2001). T-to-C movement: Causes and consequences. *Current Studies in Linguistics Series*, 36, 355-426.
- R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Reinhart, T. (1981). Pragmatics and Linguistics: An Analysis of Sentence Topics in Pragmatics and Philosophy I. *Philosophica anc Studia Philosophica Gandensia Gent*, 27(1), 53-94.
- Repp, S. (2009). *Negation in gapping*. Oxford University Press on Demand.
- Rizzi, L. (1997). The fine structure of the left periphery. In *Elements of grammar* (pp. 281-337). Springer Netherlands.
- Roberts, C. (1996). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Working Papers in Linguistics-Ohio State University Department of Linguistics*, 91-136.
- Rooth, M. (1992). A theory of focus interpretation. *Natural language semantics*, 1(1), 75-116. Chicago.
- Ross, J. R. (1967). *Constraints on variables in syntax*. (Doctoral dissertation, MIT)
- Salveste, N. (2015). *On the pragmatic and semantic functions of Estonian sentence prosody* (Doctoral dissertation, LMU).

Vallduví, E., Vilkkuna, M., (1998). On rheme and kontrast. In: Culicover, P., McNally, L. (Eds.), *The Limits of Syntax, vol. 29*. Academic Press, New York, pp. 79–108

Winkler, S. (2005). *Ellipsis and focus in generative grammar*. Walter de Gruyter.