Carl Eric Simmul*

Information structure of converb constructions: Estonian *-des*, *-mata* and *-maks* constructions

Abstract: This paper describes the variation in information structure of Estonian *-des*, *-mata* and *-maks* constructions, and analyzes the factors influencing this variation. The paper describes information structure via the categories of information status and information role. Information status, which refers to the general pragmatic status of a linguistic unit, has two possible values: an information unit or an element of an information unit. Information role refers to the pragmatic status of an element in the information structure of a larger information unit. Information role has five possible values: focus, background of the comment, topic, frame or prominent element. Through qualitative and quantitative analysis, this article gives an account of the variation in the information status and information role of Estonian converb constructions. In addition, this paper discusses the way in which the relevant explanatory variables relate to the information status and role of converb constructions. This analysis gives an overview of how the information status of converb constructions relates to the presence of punctuation, the position of the converb construction relative to the main clause, the number of words in the converb construction, the semantic function of the converb construction, the position of the converb within the construction, the morphological form of the converb, and the number of modifiers of the converb. This analysis also discusses how the information role of converb constructions relates to the position and semantic function of the construction and the presence of a lexicogrammatical prominence marker.

Keywords: information structure; converb construction; complex sentence; syntax; semantics; Estonian

1 Introduction

This article describes and explains the variation in information structure in written Estonian *-des*, *-mata* and *-maks* converb constructions.

*Corresponding author: Carl Eric Simmul**, University of Tartu, Tartu, Estonia,
E-mail: carl.eric.simmul@ut.ee

Information structure is the pragmatic status of a linguistic unit in relation to its context (see Lambrecht 1994: 5). Information structure depends on the state of mind of the speaker and the speaker's assessment of the state of mind of the addressee (Lindström 2017: 537; Matić and Nikolaeva 2018: 2; Nikolaeva 2001: 3). It is described using various categories: theme and rheme, topic and comment, background and focus, presupposition and assertion, old/known/given and new info/referent, predictable and unpredictable info; active, semi-active and inactive referent; accessible and non-accessible, identifiable and unidentifiable referent (e.g. Adamou et al. 2018; Chafe 1987; Krifka 2007; Lambrecht 1994; Leino 2013; Zimmermann and Féry 2010). Information structure may be encoded by various linguistic strategies and devices, e.g., prosody, word order, lexicogrammatical markers and pronouns.

The main domain of information structure is the information unit (Halliday 1967, 1985): a linguistic unit which consists of elements that express a complete idea in the context. The categories of information structure have been developed primarily on the basis of simple sentences. A typical information unit is a simple sentence and a typical information element is a phrasal constituent of a simple sentence (see Adamou et al. 2018; Ebert 2009; Gundel and Fretheim 2004; Halliday 1967; Krifka 2007; Lambrecht 1994; Matić and Wedgwood 2013; Zimmermann and Féry 2010).

Recently, however, increasing attention has been given to complex sentences, i.e. sentences consisting of multiple clauses (see van Gijn et al. 2014). The complex sentences of several individual languages have been discussed on the level of information structure (van Putten 2014; Reesink 2014). More specifically, the information structure of complement clauses (Ibarluzea 2014), adverbial clauses (Komagata 2003; van der Wal 2014) and relative clauses (Komen 2014; Lindström 2006; Storto 2014) has been described. Some thought has also been given to the information structure of infinitive constructions (see Matić et al. 2014: 6–7, 18–19).

The information structure of subordinate clauses can be viewed from two angles (Matić et al. 2014: 9–10). On the one hand, one can examine the clause's external information structure, i.e. the information status of the subordinate clause as a whole in the complex sentence. For instance, it has been found that an adverbial clause can function as the topic or a part of the comment of a complex sentence (see van Gijn et al. 2014: 12; Kinberg 2001; Thompson et al. 2007; van der Wal 2014: 62). On the other hand, it is also possible to analyze the clause's internal information structure, i.e. to determine which information elements make up the subordinate clause.

Converb constructions, i.e. non-finite clauses functioning as adverbial adjuncts (e.g. *He sleeps **wearing pajamas***), have been quite thoroughly discussed on the morphosyntactic (Bisang 2020; Haspelmath 1995; Nedjalkov 1995; Ylikoski 2003) and semantic (e.g. Creissels 2010; Croft 2012: 320; König 1995; Nedjalkov 1998) levels, but

not yet on the level of information structure. Information structure of a converb construction is a novel problem because a converb construction combines a different set of morphosyntactic and semantic features than a typical information unit (sentence) and a typical information element (phrase).

Broadly speaking, the morphosyntactic and semantic features of a converb construction lie between the features of a sentence and the features of a phrase. Like a sentence, a converb construction is a clause, expresses an event and has a verb form as its head. Similar to a phrase, however, a converb construction functions as a constituent, expresses a part of an event and is syntactically subordinate, typically to the predicate. In a closer perspective, a converb construction has similarities with other non-finite clauses, with adverbial phrases and adverbial clauses. Converb constructions are similar to other non-finite clauses (Shagal et al. 2022) in that they have a non-finite verb form as the head, but unlike other non-finite clauses, they function as adjuncts (Ylikoski 2003: 191). As an adjunct, a converb construction is similar to an adverbial phrase and adverbial clause, but unlike them, it has a non-finite verb form as its head.

Furthermore, Estonian converb constructions have a highly variable adverbial function[1] and word order that contribute to the challenge (Erelt 2017c: 815–818; Simmul 2020; see Section 2). This article aims to broaden the understanding of information structure and converbs, putting information-structural categories to the test on the basis of empirical observation of Estonian converb constructions.

This study proceeds from a previous qualitative description (Simmul 2021), according to which some Estonian converb constructions function as information units and others as elements within an information unit, having the role of frame, focus, or the background part of the comment (Simmul 2021: 327). In the present study I further develop the description of the information structure of converb constructions, employing among other things the category of prominent element. I also use quantitative methods to identify the relevant variables which predict the variation of the information structure of converb constructions.

The article begins with an overview of Estonian converb constructions (Section 2), Estonian information structure (Section 3), and the research questions, data, and the categories used in the present study (Section 4). This is followed by a qualitative (Section 5) and quantitative (Section 6) analysis of the information structure of the converb construction. The article concludes with a discussion (Section 7) of how the relevant explanatory features are related to the information structure of the converb construction and a summary (Section 8) of the research findings.

---

1 Except the *-maks* construction that always carries the meaning of PURPOSE (see Section 2.1).

2 Estonian converb constructions

2.1 The semantics of Estonian converb constructions

The tense of Estonian converb constructions is relative to the event of the main clause. There are three unmarked relative tense converb constructions in Estonian: the *-des*, *-mata* and *-maks* constructions.

The *-des* construction (1) expresses an occurring event, while the *-mata* construction expresses an event that does not take place (2). The event of *-des* and *-mata* constructions is usually simultaneous with the event of the main clause. However, depending on the context, the event of *-des* and *-mata* constructions may also be understood as (immediately) preceding or following the event of the main clause (Erelt 2017c: 814).

(1)   *Tädi Olga täidab kohe   **sisse astudes** kogu   korteri*
    aunt Olga fill:3sg immediately inside step:DES whole apartment.GEN
    elu   ja   liikumisega.
    life.GEN and motion:COM
    'Aunt Olga fills the whole apartment with life and motion immediately **as she enters**'
    (FIC).[2]

(2)   ***Lauset   lõpetamata*** *tõusis   ta   ähvardavalt   püsti.*
    sentence:PRT finish:MATA rise:PST.3sg s/he threateningly up
    '**Without finishing the sentence**, s/he stood up threateningly'
    (FIC).

The *-maks* construction (3) is a converb construction serving as an adverbial of PURPOSE which was intentionally devised in the early 20th century. It is generally regarded as stylistically marked and it is subject to syntactical limitations and primarily found within specific registers (Uuspõld 1980). The *-maks* construction serves primarily as a (non-obligatory) adjunct (Erelt 2017c: 799). In Estonian, the primary obligatory non-finite purpose clauses are the *-da* infinitive and *-ma* infinitive constructions (Erelt 2017c: 780–782, 792–799). In contrast to the *-des* and *-mata* construction, the *-maks* converb construction always expresses a posterior event, specifically, an event that (hypothetically) follows the event of the main clause. Given that posteriority is the sole temporal relationship with respect to the meaning of PURPOSE, I also consider the *-maks* construction as an unmarked tense construction.

---

2 The corpus references are at the end of the article.

(3)   *Sestap   katkestasin   teda   poolelt   sõnalt,*
      so   interrupt:PST:1SG s/he:PRT half:ABL word:ABL
      alustamaks   solvamist.
      begin:MAKS   insulting:PRT
      'So I interrupted her/him in mid-word, **to begin insulting her/him**'
      (BC).

In addition to the three unmarked tense converb constructions, Estonian features two marked tense converb constructions, the *-nud* and *-tud* constructions, which express an event that precedes the event of the main clause. The *-nud* and *-tud* constructions have been considered as anterior tense counterparts to the *-des* construction (Erelt 2017c: 807). For the purposes of this study, I exclusively concentrate on the unmarked tense (*-des*, *-mata* and *-maks*) converb constructions, excluding the marked tense (*-nud* and *-tud*) constructions. This approach allows for a more focused examination of the information structural effects associated with the diverse adverbial semantics (and polarity) found in converb constructions.

Converb constructions express an event that serves as a circumstance of the event of the main clause. The semantic function of the converb construction depends on the (e.g. temporal, causal) nature of this circumstance. The semantic function of converb constructions has been described using a variety of categories both in typological studies (Kortmann 1991; König 1995; Nedjalkov 1995) and in studies of converb constructions of one language family or one language (Killie and Swan 2009; Nedjalkov 1998), including Estonian (Erelt 2017c; Plado 2015a, 2015b; Simmul 2018; Valijärvi 2003; Veismann et al. 2017; Uuspõld 1966).

The semantic polysemy of *-des* and *-mata* constructions is extensive (Erelt 2017c: 815–818). Estonian converb constructions can have meanings characteristic of both clauses (e.g. RESULT, SPECIFICATION) and phrases (e.g. MEANS, MANNER). In this article I makes use of typologically known semantic categories to describe the polysemy of converb constructions (e.g. TIME, CONCOMITANCE, MANNER, CAUSE, CONDITION; see Kortmann 1991; König 1995; Erelt 2017c: 815–818). Among these categories, I distinguish between basic and complementary functions, based on the previous description of the polysemy of *-des* and *-mata* constructions (see Simmul 2018). For the sake of explanatory power, I have somewhat generalized this system for this study. Next, I describe the semantic categories used in this study.

Converb constructions always carry their basic function, which is either TIME or CONCOMITANCE. A construction carrying the basic function of TIME expresses the temporal context of the event described in the main clause (1, 4), while a construction carrying the basic function of CONCOMITANCE is itself placed within the context of the event described in the main clause (2, 3, 5). The *-des* construction, having the most extensive polysemy of Estonian converb constructions (Simmul 2018: 863–865), can carry the basic function of either TIME or CONCOMITANCE, the *-mata* and *-maks*

construction always carry the function of CONCOMITANCE as they do not express the temporal context of the event of the main clause.

The difference and opposition between TIME and CONCOMITANCE is illustrated by examples 4 and 5. In example 4, the converb construction with the function TIME serves as a temporal context of the event of the main clause. In example 5, the event of the construction with the function CONCOMITANCE is placed within the temporal context of the event of the main clause.

(4) ***Kodu*** 　 ***poole*** 　 ***kõndides*** *vilistas* 　 *mees* 　 *tuttavat*
　　home.GEN 　 towards 　 walk.DES 　 whistle.PST.3SG 　 man 　 familiar.PRT
　　viisijuppi.
　　tune.PRT
　　'**Walking towards home**, the man whistled a familiar tune'
　　(constructed).

(5) *Mees* 　 *kõndis* 　 *kodu* 　 *poole,* 　 ***vilistades*** 　 ***tuttavat***
　　man 　 walk.PST.3SG 　 home.GEN 　 towards 　 whistle.DES 　 familiar.PRT
　　viisijuppi.
　　tune.PRT
　　'The man walked towards home, **whistling a familiar tune**'
　　(constructed).

In addition to the basic function, a converb construction can also carry a complementary function, based on a possible additional causal, descriptive or contrastive relationship between the two events. In this article I describe the semantic function according to categories based on combinations of basic and complementary functions. A total of 15 combinations occur in the data: TIME-none[3] (4), CONCOMITANCE-none (5), TIME-MEANS (6), CONCOMITANCE-MEANS (7), CONCOMITANCE-CAUSE (8), CONCOMITANCE-CONDITION (9), CONCOMITANCE-MANNER (21), CONCOMITANCE-CONCESSION (22), CONCOMITANCE-PURPOSE (32), CONCOMITANCE-RESULT (33), CONCOMITANCE-CONTRAST (34), TIME-CAUSE (40), TIME-CONCESSION (44), TIME-CONDITION (47), CONCOMITANCE-SPECIFICATION (48).

(6) ***Parajalt*** 　 ***pingutades*** 　 *ja* 　 *olukordi*
　　enough 　 make_effort.DES 　 and 　 situation.PL.PRT
　　ära kasutades 　 *suudetakse* 　 *oma* 　 *sihis* 　 *edasi*
　　take_advantage_of.DES 　 be_able.IPS 　 own.GEN 　 goal.INE 　 forward
　　liikuda.
　　move.INF
　　'**Making just enough effort** and **taking advantage of situations**, one is able to move forward toward one's goal' (NEWS).

---

3 *none* indicates that the construction has no complementary function, only a basic function.

(7) *Hakkasin*   **kotis**   **sobrades**   *taskurätikut*   *otsima.*
     start:PST.1SG   bag.INE   rummage.DES   handkerchief.PRT   look.SUP
     'I started to look for the handkerchief **by rummaging through the bag**' (FIC).

(8) *Kohe*   *lendasid*   *ta*   *ette*   *asfaldile*   *tuvid,*
     immediately   fly.PST.3PL   s/he.GEN   to_the_front   asphalt.ALL   dove.PL
     lootes   **söögipoolist**.
     hope.DES   something_to_eat.PRT
     'Immediately, doves flew down onto the asphalt in front of her/him, **hoping for something to eat**' (FIC)

(9) **Ilma**   **õppimata**   *ei*   *saavutata*   *kunagi*   *midagi.*
     without   study.MATA   NEG   achieve.IPS.CNG   never   anything.PRT
     '**Without studying**, one can never achieve anything' (FIC).

2.2 The word order of Estonian converb constructions

There is no clear unmarked word order position for Estonian converb constructions. As an adjunct, the converb construction would be expected to be positioned inside the main clause, typically as the third constituent, after the subject and predicate (example 39; Sahkai 1999: 29). As a non-finite clause, however, converb constructions tend to be located further in the clause than would be expected on the basis of syntactic constituency (Sahkai 1999: 32). Most frequently, Estonian converb constructions are indeed postposed to the main clause (Simmul 2020: 228), but preposed and interposed constructions are also regular, so the word order of converb constructions in Estonian is highly variable (Lindström 2017: 560–561; Remmel 1963: 292–298; Simmul 2020). The variation of the position of converb constructions is related to the semantic function, the position of the converb, the length of the construction (Simmul 2020: 237) and the information structure (Simmul 2021: 327). In addition to this, the position of the converb also varies. The converb can appear at the beginning or the end of the construction, can function as a construction by itself, or, more rarely, can occur in the middle of the construction.

3 Information structure

3.1 Information unit

This article's treatment of information units is rooted in Halliday's concept of information units (1967, 1985) and Chafe's concept of idea units (1979, 1985, 1992).

An information unit is a linguistic unit that connects a (new) piece of information to a (known) piece of information, effectively asserting something novel and contextually relevant (see Grice 1975: 45–46). The new information within a unit aligns with what Lambrecht (1994: 52) defines as assertion. Meanwhile, the known or readily accessible information, as described by Chafe (1987), corresponds to the concept of presupposition according to Lambrecht.

The amount of asserted and presupposed information within an information unit is organized according to the categories of quantity, relation, and manner (see Grice 1975: 45–46; Halliday 1985: 275). Assertion renders an information unit effective and worthy of expression, but in excess it impedes comprehension, thus contradicting the category of manner. On the other hand, presupposition ensures that the information unit remains contextually relevant and intelligible, yet excessive presupposition goes against the category of quantity (see Grice 1975: 46).

An information unit comprises information elements, which are smaller linguistic units that express distinct pieces of information, collectively forming a complete idea within a given context. Although an information element can be intricate both semantically and grammatically,[4] it operates as a cohesive conceptual entity within the information unit. I differentiate five types of information elements: focus, background part of the comment, topic, frame, and prominent element. In the following sections, I will delve into these categories, drawing insights from the framework of Estonian information structure descriptions.

3.2 Information structure of Estonian

The structure of Estonian information units has been mainly described along three interacting dimensions, i.e. topic-comment, background-focus, and known-new. These information-structural dimensions have in turn been associated with the prosodic, syntactic and semantic domains (e.g. Asu et al. 2016; Lindström 2005; 2017; Tael 1988).

The topic-comment dimension is the aboutness relation of the information unit: the comment gives information about the topic which, in turn, states the subject matter of the comment and thereby makes the comment relevant by linking it to some accessible piece of information (Klumpp and Skribnik 2022: 1019; Lindström 2017: 537–539; Reinhart 1981).

The focus is the element that expresses the most unpredictable piece of information in the information unit. In the description of the information structure of

---

4 An information element may in its turn consist of smaller pieces of information. In this article I do not analyze the internal structure of information elements.

Estonian, the focus has been mainly treated as the *one* most unpredictable element, i.e. narrow focus (e.g. Valin and LaPolla 1997), argument focus (Lambrecht 1994; 2000) or constituent focus (Lindström 2017: 545). In description of Estonian, the so-called wide focus or predicate-focus (Lambrecht 1994: 228) has been mainly treated as the comment: the part of the information unit that provides new information about the topic (Lindström 2017: 537, 545).

The dimension of known-new is about how familiar and accessible a piece of information is in the speaker's assessment to the addressee (cf. Prince 1981, 1992; Chafe 1987; Lindström 2017: 542–544).

In description of Estonian, these information-structural dimensions have been associated with word order, lexicogrammatical markers, length of linguistic units and the number of modifiers (e.g. Lindström 2017; Remmel 1963; Tael 1988). The order of syntactic constituents, i.e. word order in Estonian is syntactically flexible and information-structurally functional. A typical information unit begins with a known topic and ends with a new focus. Between them is the background (i.e. unfocused) part of the comment, which relates the focus to the topic (Lindström 2017: 549–551). Alongside information structure, Estonian word order is regulated by a syntactic tendency, the so-called V2 rule, according to which the predicate in simple declarative sentences and main clauses tends to appear in the second position (see Lindström 2017: 547, 549, 551; Sahkai and Tamm 2019; Tael 1988: 40). The combined effect of information-structural considerations and the V2 rule is that a typical Estonian information unit features a known subject in the topic role in initial position, followed in the second position by a predicate in the background of the comment, and then concludes with a new focused modifier (Asu et al. 2016: 181–183; Lindström 2017: 545–549; Tael 1988). Thus, the main constituent order of Estonian is SVX on the syntactic level, and Top-BGoC-Foc on the level of information structure (10).

(10)   [^{S}[*Nad*]$_{Top}$ *armastavad* ^{O}[*vanu*   *filme*]$_{Foc}$]$_{IU}$.
         they   love:3PL   old.PL.PRT movie.PL.PRT
         'They love old movies'
         (constructed).

In Estonian, lexicogrammatical markers mainly contribute to indicate those information-structural properties of a linguistic unit that are unusual for its word order position (vt Lindström 2017: 544). If the topic expresses a new piece of information or if the comment contains known pieces of information, the word order usually adheres to the topic-comment structure (topic before comment). The somewhat atypical knownness/newness is thereby encoded lexicogrammatically (Lindström 2017: 543). If the referent of the topic is new (example 11) or if the referent of the comment is known (12), it is often marked by a pronoun expressing the (in)

definiteness of the referent. Also, lexicogrammatical markers are used to highlight a focus that is not information unit-final (13).

(11) [***Üks mees** näitas* meile oma uut kutsikat]~IU~.
one/a man show:PST.3SG us.ALL own.GEN new.PRT puppy.PRT
'**A man** showed us his new puppy'
(constructed).

(12) [*Jaan armastab **seda koera**]~IU~.
Jaan love:3SG this.PRT dog.PRT
'Jaan loves **that dog**'
(constructed).

(13) [*Jaan käis* [***juba eelmisel reedel**]~Foc~ *siin*]~IU.~
Jaan go:PST.3SG already last.ADE Friday.ADE here
'Jaan was here **already last Friday**'
(constructed).

Knownness and newness are also related to the number of words and modifiers of linguistic units. Linguistic units expressing new information tend to contain more words, to be more often modified and to have more modifiers than linguistic units expressing known information (Lindström 2017: 544, 547, 550).

In addition to the main information element order (Top-BGoC-Foc), Estonian has another regular element order, which starts with the so-called "alternative element" usually followed by predicate, topic and focus (Lindström 2017: 539, 551–552; Tael 1988: 6). In the description of Estonian, this alternative element has been called a secondary topic (Lindström 2017: 551–552). In this study, I avoid the term *secondary topic* to distinguish the alternative element from the aboutness-topic. The reason for this is that, even in the case of the alternative element order, the comment is not mainly about the alternative element, but still about the aboutness-topic, despite it being not positioned at the beginning of the information unit (Lindström 2017: 539).

Among those "alternative elements", I distinguish between frame and prominent element. The frame is an information unit-initial element that contextualizes the rest of the information unit (see Chafe 1976: 50–51; Jacobs 2001: 655–657; Krifka 2007: 35–37; Krifka and Féry 2008: 129; Tael 1988: 31). A typical frame is a localizing adverbial: a locative place or state adverbial, an adverbial of condition, or a localizing time adverbial, that places the rest of the information unit in a spatial, causal or temporal context. The information unit in (14) is temporally contextualized by the frame *Reedel*, 'on Friday'. At the same time, the comment is still understood to give information about the topic *ta*, 'he'.

(14)   (*Jaan käis eelmisel nädalal Tartus*).
         [[***Reedel***]_{Fr} *näitas*   [*ta*]_{Top} *meile   oma   uut   kutsikat*]_{IU}.
         Friday.ADE   show.PST.3SG he   we.ALL own.GEN new.PRT puppy.PRT
         '(Jaan came to Tartu last week.) **On Friday**, he showed us his new puppy'
         (constructed).

In Estonian, the frame does not have the aboutness relation with the comment and does not interfere with the aboutness relation between topic and comment (Lindström 2017: 539). With this in mind, unlike some approaches (e.g. Chafe 1976; Erteschik-Shir 2007), I do not treat a frame as a type of topic, but as a distinct information role.

According to the description of Estonian information structure, the alternative element, including the frame, is always the first element of the information unit (Lindström 2017: 539). However, the information unit-initial position in itself is not a sufficient criterion by which to define or identify the frame nor any of the information roles. Depending on the information unit-internal relations, the initial element may have any information role: topic (11), background part of comment (15), focus (16) or prominent element (17).

(15)   (*Mees jõudis koju ja*)
         [***pani   lilled   kapi   peale***]_{IU}.
         put.PST.3SG   flower.PL   cupboard.GEN   on
         '(The man arrived home and) **put the flowers** on the cupboard'
         (constructed).

(16)   [[***Kuhu***]_{Foc} *ta   lilled   pani*]_{IU}?
         Where   s/he   flower.PL   put.PST.3SG
         '**Where** did he put the flowers?'
         (constructed).

In addition to the frame, an alternative element may have the information role of prominent element (Tael 1988: 38–39): an unfocused emphatic element that does not express the most unpredictable piece of information in the information unit but is emphasized relative to another element in the information unit (usually the background part of the comment). Unlike a frame, a prominent element is placed at the beginning of the information unit not for the sake of contextualization, but for the sake of emphasis. Prominent element as a distinct information-structural category makes it possible to analyze an information unit with multiple emphasized elements as including one main, i.e. the most unpredictable focus while the other emphasized elements function as prominent elements. In (17) the prominent

element is *muretult*, 'carelessly' which is emphasized by its information unit-initial position. The focused piece of information is expressed by the information unit-final element *kingadele*, 'on shoes'.

(17) [[**Muretult**]$_{PrE}$ *kulutas* [*Jaan*]$_{Top}$ *kogu* *raha*
      Carelessly   spend:PST.3SG Jaan   all   money.GEN
      [*kingadele*]$_{Foc}$]$_{IU}$.
      shoe:PL:ALL
      '**Carelessly**, Jaan spent all the money on shoes'
      (constructed).

The information structure of Estonian information units exhibits variation. Information units can feature different combinations of information elements and are not required to include all of the elements. In an information unit that contains a focus, background part of the comment, topic, and frame, the division of roles can be described as follows.

The focus conveys the most unpredictable piece of information. The background of the comment offers an immediate context for the focus and, in conjunction with the focus, forms the comment. The topic contextualizes the comment according to the aboutness relation. The frame contextualizes the rest of the information unit, encompassing the entire topic-comment structure. An example of such an information unit is presented in (14), and the structure of this type of information unit can be depicted as follows:

(18)   [Frame [Topic [Background part of comment [Focus]]$_C$]]$_{IU}$

Furthermore, the background of the comment, the topic, and the frame collectively constitute the background of the information unit, representing the portion of the information unit that is backgrounded relative to the focus.

The elements of the same information unit can be interconnected through three types of information unit-internal relations. Namely, these relations are the comment-forming relation between the focus and the background part of the comment, the aboutness relation between the comment and the topic, and the contextualizing relation between the frame and the rest of the information unit. These relations serve to connect the elements of an information unit in such a way that they collectively express a coherent and contextually relevant idea.

4 Research questions and data

I aim to answer the following questions:

- What are the information status[5] and roles of Estonian converb constructions?
- What features are relevant to the variation in the information status and role of Estonian converb constructions?
- How do these features relate to information status and role?

Next, I will outline the data collection process (Section 4.1) and provide an overview of the response variables, explanatory variables, and their respective values (Section 4.2). The first research question is addressed through qualitative analysis (Section 5), while the second question is approached using a quantitative methodology (Section 6). The third question is answered by interpreting the findings from the quantitative analysis (Section 7).

4.1 Data

This paper is based on a corpus data set consisting of sentences that contain *-des*, *-mata* or *-maks* constructions. The sentences with *-des* and *-mata* constructions come from the 1990s print media (NEWS; 865,000 words) and fiction (FIC; 602,000 words) subcorpora of the Corpus of Written Estonian. I started collecting the data from the *-mata* constructions, the amount of which is the most difficult to estimate based on the initial results of corpus inquiry. This is primarily because the *-mata* form is multifunctional. It functions not only as a converb but also as a participle and part of compound predicates. Furthermore, the *-mata* form is relatively frequent in lexicalized and grammaticalized converb constructions. I collected all the *-mata* converb constructions occurring in the corpora (1,255). Then, I removed the lexicalized and grammaticalized converb constructions (e.g. *kogemata*, 'accidentally; lit. without experiencing'; *kahtlemata*, 'undoubtedly; lit. without doubting'; *tingimata*, 'necessarily; lit. without bargaining'), resulting in 476 *-mata* constructions in the data. Example (19) illustrates a concessive postpositional phrase that has been grammaticalized from the *-mata* converb construction.

(19) **Võlgnevusest hoolimata**   *tuleb*   *dokumendid* *kliendile*
      debt.ELA   notwithstanding must.3SG document.PL   customer.ALL
      tagastada.
      return.INF
      '**Notwithstanding the debt**, the documents must be returned to the customer'
      (www).

Collecting *-des* constructions, I had in mind that data should reflect the fact that *-des* converb constructions are more frequent than *-mata* converb constructions. I also wanted the data to represent the rarer semantic functions of *-des* construction. On the other hand, I wanted a comparable number of *-des* and *-mata* constructions. With that in mind, I collected approximately twice as many *-des* constructions as *-mata* constructions, collecting *-des* constructions from a fifth of the text from which I had collected *-mata* constructions. There were 1,025 *-des* forms in the selected text. After removing the lexicalized and grammaticalized constructions (e.g. *alates*, 'since; lit. beginning'; *möödaminnes* 'passingly; lit. passing by'; *ausalt öeldes*, 'honestly; lit. honestly saying') 825 *-des* constructions remained.

Given the rarity of the stylistically marked *-maks* construction in fiction, I collected *-maks* constructions from the larger Balanced Corpus (BC), which comprises 15 million words encompassing various registers, including scientific, print media, and fiction. I collected 500 *-maks* constructions, resulting in a dataset containing a total of 1,803 converb constructions.

The corpora also provided the opportunity to consider the context of the sentences containing the converb constructions. Assessing the converb constructions and their relationship with the main clause in their context, I systematically annotated the values of the 11 information structural, semantic, morphosyntactic and orthographic variables of this study.

4.2 Variables and values

In the data, I annotated 2 response variables and 9 explanatory variables. The response variables are information status and information role – two information structural categories. The explanatory variables are various semantic, morphosyntactic and orthographic categories (see Table 3). Each variable's values were annotated independently, separate from the values of other variables. It's important to note that the values of the explanatory variables are not considered the defining or identifying criteria for the values of the response variables. The primary focus of this study lies in exploring the relationships between the response variables and the explanatory variables.

The values of the information structural response variables have been operationalized in terms of contextual relations. Specifically, the values of the response variables were annotated based on the informational relations between a linguistic unit and its surrounding context.

The response variable of information status has two values: information unit and information element. These two categories are distinguished by the involvement

of a linguistic unit in at least one of the three information unit-internal relations: 1) the comment-forming relation, 2) the aboutness relation, or 3) the contextualizing relation between the frame and the rest of the information unit.

A linguistic unit directly linked to another linguistic unit through any of these relations is categorized as an information element. Conversely, a linguistic unit not directly linked to another linguistic unit through these relations is designated as an information unit. From another perspective, an information element has one or more information elements of the same information unit as its context, while an information unit has other information units as its context.

In terms of informational completeness, an information element expresses a part of an idea, whereas an information unit expresses a complete idea. In terms of internal information structure, an information unit possesses its own information structure, while an information element is a component of the information structure of a larger linguistic unit.

The response variable of information role concerns information elements. Information role has five values in this study: focus, background part of the comment, topic, frame and prominent element. The values of information role were annotated based on the way in which an information element relates to the other information elements within the information unit.

In this study, focus is understood as the most unpredictable information element in the information unit with respect to the context of the information unit. Such a definition combines Lambrecht's (1994: 207) notion of pragmatic unpredictability with the notion of narrow focus (Valin and LaPolla 1997). According to this definition, every information unit has its own focus. The information structure of an information unit is pragmatically centered on the focus that is distinct from the foci of other information units. If an information unit consists of a single element, then this is the focus. In addition to the focus, an information unit may have other, unfocused information elements, that contextualize the focus and that are pragmatically backgrounded to the focus.

Usually, the focus can be questioned in relation to the other elements of the information unit. However, when dealing with subordinate clauses, determining the implicit question that the information unit addresses is not always straightforward. Since both the converb construction and its main clause can be subordinated clauses, I have not relied solely on questioning as the method for identifying the focus. Instead, I have relied on a number of interrelated and overlapping interpretative focus effects (see Matic and Wedgwood 2013). Specifically, I have identified the focus as the information element that most strongly evokes the interpretative effects of salience, remarkability, newsworthiness,

Table 1: Characteristics of information roles.

	The most unpredictable piece of information	About topic	Subject matter of comment	Contextualizes the topic-comment structure	Emphatic
Focus	+	+	−	−	+/−
Background part of the comment	−	+	−	−	−
Topic	−	−	+	−	−
Frame	−	−	−	+	−
Prominent element	−	+/−	+/−	+/−	+

and/or irreplaceability within the context of this information unit (see Matic and Wedgwood 2013: 137, 141, 157–158).

The unfocused information roles have been annotated based on their relationships within the information unit. The background part of the comment is an information element that provides immediate context to the focus. Combined with the focus, it offers information about the topic, whereas the topic itself is an information element that contextualizes the comment by indicating its subject matter. The frame is an information element that contextualizes the topic-comment structure, presenting an intermediary context that elaborates on the broader, information unit-external context of the information unit. The prominent element is any unfocused information element that is emphasized in relation to another information element of the same information unit. A summary of the characteristics of these information roles is provided in Table 1.

Table 2 gives an overview of the information status and information role categories used in this study.

In addition to the response variables, I have also annotated 9 explanatory variables (cf. Table 3) identified on the basis of previous studies (Lindström 2017; Remmel 1963; Simmul 2021; Tael 1988). Two of the explanatory variables – the number of words in the converb construction and the number of modifiers of the construction – are quantitative, while the rest are qualitative. Next, I will explain 3 explanatory variables that may not be self-evident.

As values of the morphological form (Form), I distinguish the three unmarked tense converbs of Estonian: the -*des*, -*mata*, and -*maks* forms. This variable makes it possible to find out how important the form of the converb is with respect to the information structure of the converb construction.

Table 2: Information structure categories.

Type	Category value	Definition	Example
Information status	Information unit (IU)	A linguistic unit expressing a complete idea. Has an information structure of its own. Has a focus of its own	***Eile kirjutasin ma*** yesterday write:PST.1SG I ***luuletuse:*** poem.GEN '**Yesterday I wrote a poem**'
	Information element (IE)	A linguistic unit that expresses a part of an idea. Is a part of an information structure of a larger linguistic unit. Does not have a focus of its own	***Eile*** *kirjutasin ma luuletuse.* '**Yesterday** I wrote a poem'
Information role	Focus (Foc)	The most unpredictable element in the information unit	*Eile kirjutasin ma **luuletuse**.* 'Yesterday I wrote **a poem**'
	Background part of the comment (BGoC)	An element that, together with the focus, gives information about the topic	*Eile **kirjutasin** ma luuletuse.* 'Yesterday I **wrote** a poem'
	Topic (Top)	An element that states the subject matter of the comment	*Eile kirjutasin **ma** luuletuse.* 'Yesterday **I** wrote a poem'
	Frame (Fr)	An element that contextualizes the topic-comment structure	***Eile*** *kirjutasin ma luuletuse.* '**Yesterday** I wrote a poem'
	Prominent element (PrE)	An emphasized but unfocused element	***Alles eile*** *kirjutasin ma luuletuse.* '**Only yesterday** I wrote a poem'

By the modifiers of the converb (NoM) I mean the direct dependents of converb: the object, the predicative and adjuncts of the converb. In (20) the converb has four modifiers.

(20) *Marvi   sobras   neis,   valis,   ᴹise*
      Marvi   rummage.PST.3SG   they.INE   choose.PST.3SG   herself
      ᴹ*sealjuures   ᴹrõõmsameelselt ᴼjuhtumeid   kirjeldades*.
      at_the_same_time   happily   incident.PL.PRT   describe.DES
      'Marvi rummaged through them, picked them out, **at the same time happily describing the incidents**'
      (FIC).

Table 3: Explanatory variables and their values.

Abbr	Variable	Values
Form	Morphological form of the converb	*-des, -mata, -maks*
Punct	Presence of punctuation separating the converb construction from the main clause	yes, no
PrM	Presence of a lexicogrammatical prominence marker marking the converb construction	yes, no
SF	Semantic function of converb construction	CONCOMITANCE-CAUSE (**C_CA**), CONCOMITANCE-CONCESSION (**C_CONC**), CONCOMITANCE-CONDITION (**C_COND**), CONCOMITANCE-CONTRAST (**C_CONT**), CONCOMITANCE-MANNER (**C_MA**), CONCOMITANCE-MEANS (**C_ME**), CONCOMITANCE-NONE (**C_none**), CONCOMITANCE-PURPOSE (**C_PU**), CONCOMITANCE-RESULT (**C_RE**), CONCOMITANCE-SPECIFIACATION (**C_SP**), TIME-CAUSE (**T_CA**), TIME-CONCESSION (**T_CONC**), TIME-CONDITION (**T_COND**), TIME-MEANS (**T_ME**), TIME-NONE (**T_none**)
PoCv	Position of the converb within the construction	first, last, alone, other
PoCC	Position of the converb construction relative to the main clause	pre, intra, post
Deix	Presence of a deictic in the converb construction	yes, no
NoW	Number of words in the converb construction	1–29
NoM	Number of modifiers of the converb	0–4

By lexicogrammatical prominence markers (PrM)[6] I mean non-constituent markers (21) and *-gi/-ki* clitics (22) that emphasize the converb or the construction. I also treat as a prominence marker the preposition *ilma* 'without', which emphasizes the absence of the event of the *-mata* construction.

(21)   (.. *ta lööb pea vastu lage ja*)
    *saab   edasi   liikuda   **vaid**   kummardudes.*
    can:3SG   on   move.INF   only   bow_down.DES
    '(S/he hits her/his head on the ceiling and) can move on **only** by bowing down'
    (FIC).

(22)   *Elada   võis   maad   kaevamata=**gi***
    live.INF   can.PST.3SG   ground.PRT   dig.MATA=CL
    'It was possible to live without digging the ground, **too**'
    (FIC).

---

6 I term these items as *prominence* markers, not focus markers, because they directly relate to the prominence (i.e. being emphasized) of a linguistic unit and not necessarily focus, i.e. the pragmatical status of expressing the most unpredictable piece of information in the information unit.

(23)   *Me oleme selle   asja   lõpule   viinud*
        we   be.1PL this.GEN thing.GEN end.ALL take.PST.PTCP
        ilma   põhioperatsiooni   alustamata.
        without   main_operation.PRT   start.MATA
        'We have finished it **without** starting the main operation'
        (NEWS).

Table 3 below gives an overview of all the explanatory variables and their values. The abbreviations used in the table are used in figures, tables, and examples throughout this article.

5 Information status and information roles of converb constructions: a qualitative analysis

In this section I provide an overview of how Estonian converb constructions function both as information units and as information elements.

A converb construction that contains a focus that is not the focus of any other linguistic unit functions as an information unit. In (24), the converb construction contains a distinct focus *uute riiulitega*, 'with the new shelves' and has the status of information unit. The main clause functions as a different information unit with the focus *õues*, 'outside'.

(24)   [*Veetsime   kogu   päeva   [õues]*_{Foc}]_{IU}, [*jännates   mitu*
        spend.PST.1PL whole   day.GEN outside   struggle.DES several
        *tundi   [uute   riiulitega]*_{Foc}]_{IU.}
        hour.PRT   new.PL.GEN   shelf.PL.COM
        'We spent the whole day **outside**, struggling several hours **with the new shelves**'
        (constructed).

A converb construction that shares a focus with another linguistic unit functions as an information element. An information element may relate to another linguistic unit as its focus or function itself as the focus of another linguistic unit. In (25), the converb construction functions as the focus of the rest of the main clause, which constitutes the background part of the comment.

(25)   [*Veetsime   kogu   päeva   [aias   töötades]*_{Foc}]_{IU.}
        spend.PST.1PL whole   day.GEN garden.INE work.DES
        'We spent the whole day **working in the garden**'
        (constructed).

A converb construction may also contain several information units. If a converb construction contains a subordinate embedded clause that has a focus that differs from the focus of the rest of the converb construction, then this subordinate clause functions as a distinct information unit. In (26), such a subordinate clause has a focus that differs from the focus of the rest of the converb construction.

(26)　　(.. *keda ta oma esseedes ellu äratas,*)
　　　　[*küsimaks　uue　　　pingestusega* [**_noidsamu　suuri_**
　　　　Ask.MAKS　　new.GEN　tension.COM　　same.PL.PRT　big.PL.PRT
　　　　küsimusi]~Foc~]~IU~, [*mida　　ta　　esitas　　　endale*
　　　　question.PL.PRT　　what.PRT　s/he　put.PST.3SG　her/himself.ALL
　　　　[**_alatasa_**]~Foc~]~IU~.
　　　　perpetually
　　　　'(.. whom he revived in her/his essays), to ask with a new tension **those same big questions** that s/he put to herself/himself **perpetually**'
　　　　(constructed).

On the other hand, a converb construction can also function as a part of an information element. In (27), the converb construction *tagasi pöördumata* 'without returning' forms an integral focus with the adverbs *jäljetult, jäägitult* 'without a trace, completely'.

(27)　　[*Kaduda　　　* [*jäljetult,　　　jäägitult,　　**_tagasi　pöördumata_**]~Foc~]~IU~ –
　　　　Disappear.INF　without_a_trace　completely　back　　return.MATA
　　　　[*milline　õndsus*]~IU~.
　　　　what　　bliss
　　　　'To disappear without a trace, completely, **without returning** – what bliss'
　　　　(FIC).

In this study, I focus on the distinction between the two main information statuses of the converb construction, that of the information unit and information element. If a syntactic converb construction contains several information units, I consider only the information unit that contains the converb (in [26], **_küsimaks_** *uue pingestusega noidsamu suuri küsimusi*). If the converb construction functions as a part of an information element, I consider it an information element (e.g. a focus in [27]).

5.1　Information unit

A converb construction can function as an information unit, which expresses a complete idea, contains new information and relates to the main clause as its

information unit-external context. It contains a focus of its own and may be preposed (28) or postposed (29) to the main clause or interposed between the parts of the main clause (30).

(28)   **[BGoC [Foc]]**_{IU} + [Top [BGoC [Foc]]]_{IU}

(29)   [Top [BgoC [Foc]]] + **[BGoC [Foc]]]**_{IU}

(30)   [Top [BgoC **[BGoC [Foc]]**_{IU} [Foc]]]_{IU}

A preposed information unit presents a new idea, which in turn becomes the starting point of the idea of the main clause (31). A postposed information unit adds a new idea about the topic of the preceding main clause (32). An interposed information unit presents relevant information that stands relatively separate from the adjacent information structure of the main clause (33).

(31)   **[*Saades* [*positiivsed ja ammendavad vastused*]_{Foc}]_{IU}**,
         get.DES   positive.PL   and comprehensive.PL answer.PL
         [*otsustati   tehing   sooritada*]_{IU}.
         decide.IPS.PST transaction perform.INF
         '**After getting positive and comprehensive answers**, they decided to perform the transaction'
         (NEWS).

(32)   [*Rüütlid olid   aga   kummardunud üle*
         knight.PL be.PST.3PL however lean.PST.PTCP   over
         reelingu]_{IU}, [**nägemaks** [*sogastes   rohelistes   lainetes*
         railing.GEN see.MAKS   muddy.PL.INE green.PL.INE wave.PL.INE
         hulpivaid   jäätükke]_{Foc}]_{IU}.
         swaying.PL.PRT piece_of_ice.PL.PRT
         'The knights, however, had leaned over the railing **to see the pieces of ice bobbing up and down in the muddy green waves**'
         (BC).

(33)   [*Ja   tõlkis   seda suurepärase meisterlikkuse,*
         and translate.PST.3SG it.PRT excellent.GEN mastery.GEN
         elegantsiga – **[*saavutades külalistega [väga hea*
         elegance:COM achieve.DES   guest.PL.COM very good.GEN
         ***kontakti*]_{Foc}]_{IU}** – *Oleg Mutt*]_{IU}.
         contact.GEN   Oleg Mutt
         'And he translated it with outstanding mastery, elegance – **making a very good connection with the guests**'
         (NEWS).

A typical converb construction functioning as an information unit contains a focus and a background part of the comment. The topic of the converb construction is implicitly understood from the context, usually from the main clause (examples 32–33). There are, however, instances (example 34) in which the converb construction arguably contains an explicit topic that reactivates the topic of the main clause. A word that often performs the role of reactivated topic is the adverbial pronoun *ise* 'self', which usually appears in constructions that contrast with the main clause.

(34) *"Kuidas jalad on?"* *päris* *Uber,* [*tundes* **[*ise*]**$_{\text{Top}}$
      how leg.PL be.3PL ask.PST.3SG Uber feel.DES self
      [*arusaamatut* *piinlikkust*]$_{\text{Foc}}$]$_{\text{IU}}$.
      Incomprehensible.PRT awkwardness.PRT
      '"How are your legs?", asked Uber, **himself** feeling incomprehensible
      awkwardness'
      (FIC).

A converb construction functioning as an information unit has an information structure of its own organized around a distinct focus. The converb construction expresses a complete idea and an event that is related to the event of the main clause, but is considered form a different perspective. Although syntactically the converb construction is embedded within the main clause, on the information level it stands apart from the other constituents of the main clause and relates to the rest of the main clause as a relative whole to a whole. The converb construction functioning as an information unit expresses sentential, eventive meaning and, on the information level, resembles a typical simple sentence, apart from usually not having an explicit topic of its own.

5.2 Element of an information unit

A converb construction functioning as an information element forms an information unit together with the main clause, wherein the converb construction has the role of frame, focus, background of the comment, or prominent element. In this data, none of the converb constructions function as a topic, i.e. there is no information unit whose comment is mainly about the event of the converb construction. This can be attributed to the fact that the dynamism of the event of the converb construction is not characteristic to the role of topic.[7]

---

7 At least not in Estonian.

5.2.1 Frame

A converb construction functioning as a frame contextualizes the rest of the information unit by (re)activating a situation accessible in the context (35; see Chafe 1987: 25, 1992: 21).

(35)　　[**Fr** [Top [BGoC [Foc]]]]$_{IU}$

In (36), the frame *Seal seistes* 'standing there' reactivates the situation presented two sentences earlier, to which a new event (*tuli mul mõte* 'I had an idea') is related.

(36)　　(*Seisime tundide viisi nende häbematute tegelaste akende all, pisarad silmis.*
　　　　[—])
　　　　[[***Seal***   ***seistes***]$_{Fr}$ *tuli*   *mul*   *mõte*]$_{IU}$ –   *korraldada*
　　　　There   stand.DES   come.PST.3SG   me.ADE   idea   organize.INF
　　　　Sahharovi   *auks*   *kontsert*
　　　　Sakharov.GEN   honor.TRSL   concert.
　　　　('We stood for hours beneath the windows of those shameless characters,
　　　　with tears in our eyes. [—]) **Standing there** I had an idea: to organize a
　　　　concert in Sakharov's honor.'
　　　　(NEWS)

The frame typically carries a localizing meaning, which in the case of converb constructions is expressed most clearly by the basic function TIME (example 36) or the complementary function CONDITION (37). Such converb constructions express the temporal or conditional context of the event of the main clause (see Diessel 2005).

(37)　　[[***Sõnagi***   ***keelt***   ***oskamata***]$_{Fr}$ *ei*   *saa*
　　　　word.PRT=CL   language.PRT   know.MATA   NEG   can.CNG
　　　　võõra   *rahva*   *keskel*   *elada*]$_{IU}$.
　　　　foreign.GEN   people.GEN   among   live.INF
　　　　'**Without knowing a word of the language**, you can't live among a foreign
　　　　people'
　　　　(FIC).

5.2.2 Background part of the comment

A converb construction functioning as the background part of the comment complements the focused piece of information with an additional circumstance (38).

(38)　　[Top [**BGoC** [Foc]]]$_{IU}$

Typically, the backgrounded part of the comment expresses a new event of secondary importance (example 39).

(39) [*Nad võivad **aega raiskamata** pühenduda*
they can:3PL time.PRT waste:MATA devote:INF
[*turumajandusele üleminekule*]Foc]IU.
market_economy:ALL transition:ALL
'They can devote themselves to the transition to a market economy **without wasting time**'
(NEWS).

The background part of the comment can also express a known, contextualizing event (40). Compared to the frame, the background part of the comment is more closely related to the rest of the comment and topic and does not stand out in the unit.

(40) *Oleksite te näinud, Isabel,* [*kuidas ema **mind***
be:COND:2PL you see:PST.PTCP Isabel how mother me:PRT
nähes [*rõõmustas*]Foc]IU.
see:DES rejoice:PST.3SG
'If only you had seen, Isabel, how my mother rejoiced **when she saw me**'
(FIC).

5.2.3 Focus

A focused converb construction expresses a circumstance which is the most unpredictable piece of information in the information unit (41–43). Such an information unit expresses a complex situation, the center of which is an eventive circumstance.

(41) [Top [BGoC [**Foc**]]]IU

(42) [*Seisime minutikese* **[üksteise ümbert kinni hoides**
stand.PST.1PL minute.GEN each_other.GEN around fixed hold.DES
ja nuttes]Foc]IU.
and cry:DES
'We stood for a minute **holding each other and crying**'
(FIC).

(43) [*Roosi teadis seda* **[*pärimata=gi*]**Foc]IU.
Roosi know.PST.3SG this.PRT ask.MATA=CL
'Roosi knew that **even without asking**'
(FIC).

5.2.4 Prominent element

A converb construction functioning as a prominent element expresses an emphatic but unfocused circumstance of a complex event. Emphasis can be expressed in many ways, of which the most common in the case of Estonian converb constructions are information unit-initial position and lexicogrammatical prominence markers.

A contextualizing preposed construction functions as a prominent element when it's emphasized by a lexicogrammatical prominence marker (example 44).

(44)   [*Kas* **[*isegi minuga magades*]**~PrE~ *mõtles*   *ta*   [*oma*
         Q   even   me.COM   sleep.DES   think.PST.3SG she   own.GEN
         kõutsist]~Foc~]~IU~?
         cat.ELA
         'Was she thinking about her cat **even while sleeping with me**?'
         (FIC)

A construction without a contextualizing function can be emphasized purely by its information unit-initial position (Tael 1988: 38–39). This kind of emphasis is typical of converb constructions performing the function of MANNER (example 45). The prominence of information unit-initial position has received some attention in the description of Estonian MANNER adverbials (see Erelt et al. 2020: 472; Tael 1988: 43).

(45)   [[***Naerdes ja lauldes***]~PrE~ *olid*   *nad* [*hullutanud*
         laugh.DES and sing.DES   be.PST.3PL they drive_crazy.PST.PTCP
         küla]~Foc~]~IU~, *kuid* [[***naerdes ja lauldes***]~PrE~ *olid*   *nad*
         village.PRT but laugh.DES and sing.DES   be.PST.3PL they
         püstitanud [*ka kõige nõudlikumad   hooned*]~Foc~]~IU~.
         erect.PST.PTCP also most demanding.CMP.PL building.PL
         '**Laughing and singing** they had driven the village crazy, but **laughing and singing** they had also erected the most demanding buildings'
         (FIC).

An interposed information element may be emphasized by a lexicogrammatical prominence marker (46).

(46)   (*Kes vene usku läheb, tollele antakse maad ja toda ei võeta soldatiks ja*)
         [*too*   *võib* **[*ilma   leeritamata=gi*]**~PrE~ [*naist   võtta*]~Foc~]~IU~.
         that   can.3SG without confirm.MATA=CL   wife.PRT take.INF
         ('He who joins the Russian faith is given land and is not conscripted as a soldier and) he can marry **even without going through confirmation**'
         (FIC).

5.2.5 The information roles of converb constructions

A converb construction functioning as an information element forms a common information unit with other constituents of the main clause, functions as a part of the information structure of the main clause like a typical phrase and expresses an event that serves as e circumstance of a larger event. The information role of the converb construction depends on how this circumstance relates to the other circumstances of the event of the information unit. If this circumstance is the most unpredictable piece of information in the unit, the converb construction has the information role of focus. If this circumstance is emphatic, but not the most unpredictable piece of information, the converb construction has the information role of prominent element. If this circumstance is not emphatic nor the most unpredictable piece of information but is about the topic, the converb construction has the role of background part of the comment. If this circumstance contextualizes the rest of the event and is not emphatic, the converb construction has the role of frame.

6 Quantitative analysis of variables relevant to the information status and information role of converb constructions

47 % of the converb constructions in the data set function as information units, while the other 53 % function as information elements, having the information role of focus, background of the comment, frame or prominent element. Table 4 summarizes the relative frequencies of the information status and role of converb constructions.

I will identify variables relevant to the information status and role of converb constructions by using a classification tree and random forest method, which is based on the repeated grouping of observations (Breiman 2001; Strobl et al. 2009). The tree and forest method is suitable for sorting out natural language data, because interaction between the explanatory variables does not hamper the effectiveness of

Table 4: The information status and role of the converb constructions.

IU	IE			
	Foc	**BGoC**	**Fr**	**PrE**
842 (47 %)	399 (22 %)	269 (15 %)	191 (11 %)	102 (5 %)

the method (Baayen et al. 2013; Klavan et al. 2015; Levshina 2015; Strobl et al. 2008; Tagliamonte and Baayen 2012).

The tree and forest method measures how strongly the explanatory variables correlate with the response variable, i.e. how much of the variation in the response variable is predicted by the variation in the explanatory variables. The tree and forest method treats explanatory variables as predictors and the response variables as predicted variables. Such a point of view does not say anything about the direction in which the variables are linked in the real-life language use. The tree and forest method does not determine whether the value of any variable causes the value of another variable as its effect. Even if the explanatory variable and response variable are strongly correlated, neither is necessarily the cause of the other.

Sections 6.1 and 6.2 give a general overview of the variables that predict the variation in the information status and information role according to the tree and forest method. Further interpretation of this variation is provided in Section 7.

6.1 Variables relevant to information status

The random forest, developed by Breiman (2001), uses trial and error to determine the significance of each individual explanatory variable to the response variable (Baayen et al. 2013: 265, 267; Levshina 2015: 292; Tagliamonte and Baayen 2012: 157–158). The random forest measures how strongly each explanatory variable is correlated with the response variable. The interaction between the explanatory variables does not affect their relationship with the response variable. This is important in this study because some explanatory variables are closely related to each other. There is interaction between, for example, 1) the morphological form of the converb and the semantic function of the converb construction, 2) the position of the converb and the presence of punctuation, and 3) the number of modifiers of the converb and the number of words in the converb construction. Neither these nor other interactions increase or decrease the predictive power of any of the explanatory variables relative to the response variable. The random forest does not give any information about the relationships among the explanatory variables themselves; these relationships can be analyzed with the help of a classification tree (Baayen et al. 2013: 265, 273; cf. Figure 2).

Figure 1, based on a random forest analysis, arranges the explanatory variables from top to bottom in order of how strongly correlated they are to the information status of Estonian converb constructions. Seven relevant variables are shown by points located to the right of the vertical line on the diagram: the presence of punctuation (Punct; see Section 7.3), the position of the converb within the construction (PoCv), the semantic function of the construction (SF), the position of the
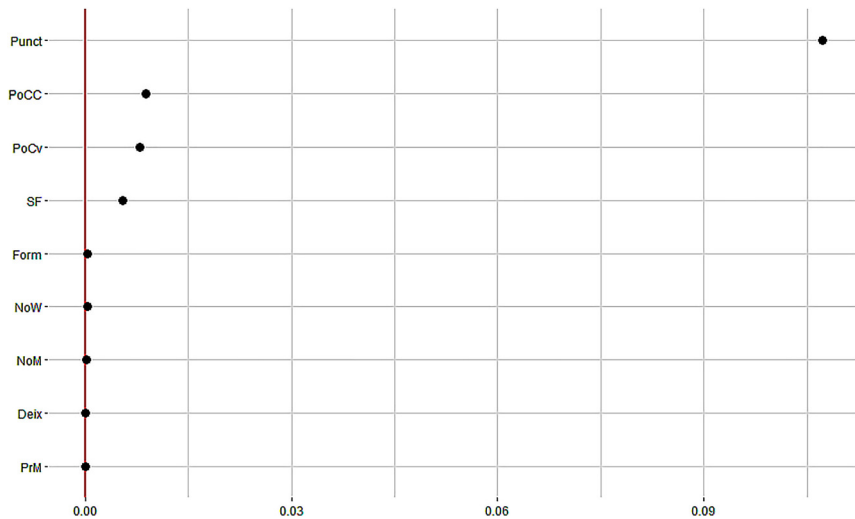
Figure 1: Variables relevant to the information status of the converb construction (random forest).

construction (PoCC), the morphological form of the converb (Form), the number of words in the construction (NoW) and the number of modifiers of the converb (NoM).

A classification tree gives an overview of the relationships of different explanatory variables to each other as well as to the response variable (Baayen et al. 2013: 265, 273). It identifies an explanatory variable closely tied to the response variable and splits observations into two sets on the basis of this explanatory variable (Baayen et al. 2013: 265; Levshina 2015: 291; Strobl et al. 2009: 325). This is then repeated in the subsets created in the previous iteration until either a) no more statistically significant variables can be found or b) the number of observations in any one subset reaches a given lower limit (Baayen et al. 2013: 265; Strobl et al. 2009: 327).

Although the principle behind the classification tree is simple, interpreting the diagram may be difficult if it contains a high number of nodes (Baayen et al. 2013: 265, 287). Thus, to optimize the explanatory power of a classification tree, a minimum number of observations required to form a subset is often established. In this study, I use a minimum of 7 % of observations, which means that each subset used in the analysis of information status contains at least 120 observations (cf. Figure 2) and each subset used in the analysis of information role contains at least 65 observations (cf. Section 6.3, Figure 4).

The classification tree (cf. Figure 2) predicts the variation in information status based on the presence of punctuation, the position of the converb and the semantic function of the construction. In the diagram showing the classification tree, divisions,
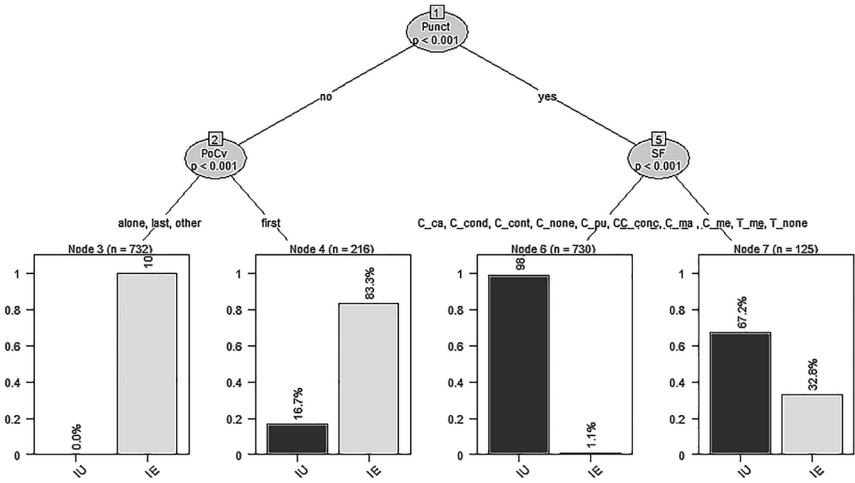
Figure 2: Variables relevant to variation in information status (classification tree).

i.e. nodes, are arranged from top to bottom: starting from the uppermost node (1), which divides all observations in the data set, each node (1, 2, 5) divides all observations that reach it into two groups. The groups (3, 4, 6, 7) at the bottom of the diagram show the summary totals of observations as bar charts, which show the proportions of different values of the response variable belonging to a particular group. The number of observations in each group is shown above the bar charts.

According to Figure 2, constructions separated by punctuation (node 1) typically function as information units (nodes 6 and 7). However, constructions that carry the functions CONCESSION, MANNER, MEANS or TIME also function as information elements in almost one third of cases (node 7). Constructions not separated by punctuation usually function as information elements (nodes 3 and 4), but if the construction is converb-initial, it can also function as an information unit (node 4). In general, the presence of punctuation and the placement of the converb at the beginning of construction predict the status of information unit, while the absence of punctuation and the functions of CONCESSION, MANNER, MEANS and TIME predict the status of information element.

6.2 Variables relevant to information role

Figure 3, based on a random forest analysis, arranges the explanatory variables from top to bottom in order of how strongly they correlate with information role. Three relevant variables emerge, shown by points located to the right of the vertical line on
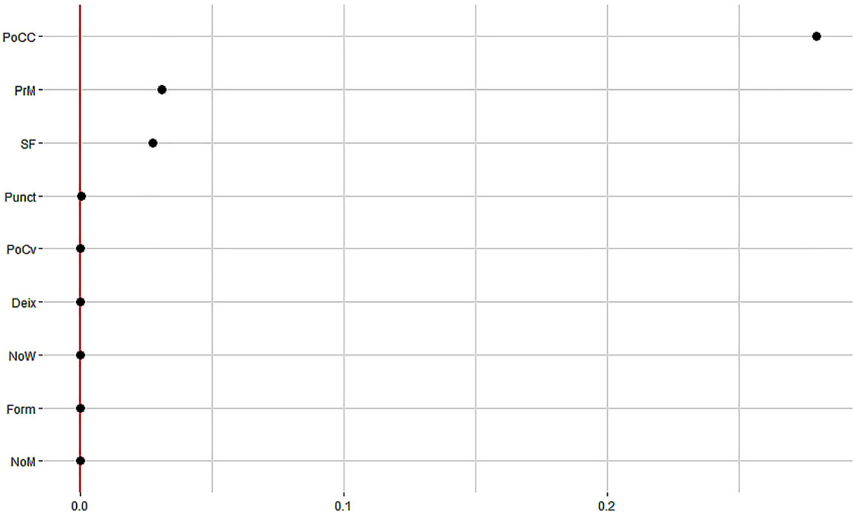
Figure 3: Variables relevant to information role (random forest).

the diagram: the position of the converb construction (PoCC), the presence of a prominence marker (PrM) and the semantic function of the converb construction (SF).

The classification tree (cf. Figure 4) predicts the variation in information role based on the position and semantic function of the construction. The information role of converb constructions generally, i.e. except for the role of topic, follows the
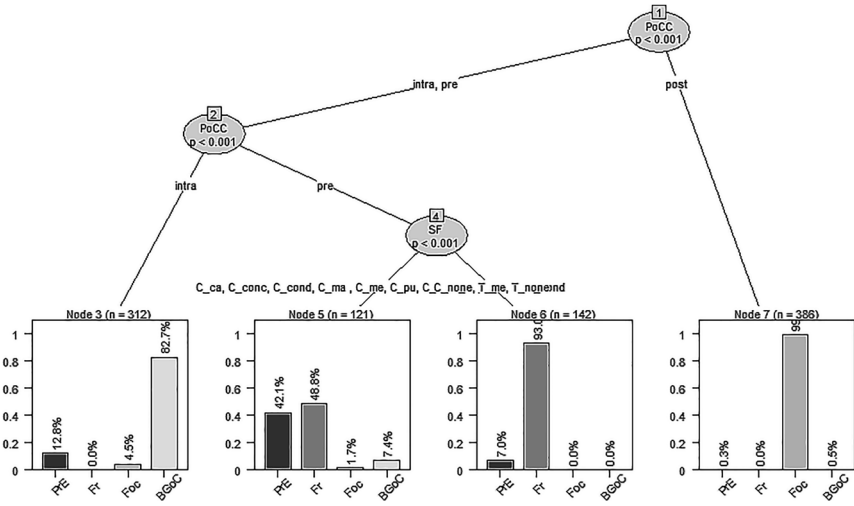

Figure 4: Variables relevant to information role (classification tree).

same word order patterns as the information role of a typical Estonian phrase (see Lindström 2017: 549–555). An information unit-final (postposed) converb construction typically acts as a focus (node 7), while an interposed construction typically functions as background of the comment (node 3). An information unit-initial converb construction functions as frame (node 5) or prominent element (node 6), depending mainly on its semantic function.

The only coded variable that shows no significant relationship to the variation in either information status or information role is the presence of a deictic in the construction.

7 Interpretation of quantitative analysis: the relationship of explanatory variables to the information status and role of the converb construction

In this section I interpret the results of the quantitative analysis, also using descriptive statistics to support the interpretations.

7.1 Morphological form of the converb

The distribution of information status and role of *-des* and *-mata* constructions differs from that of *-maks* constructions. Over half of *-des* and *-mata* constructions function as elements of an information unit, whereas almost two thirds of *-maks* constructions function as information units. Table 5 provides a breakdown of information status and role for the three morphological forms of the converb.

Table 5: Distribution of information status and role by the morphological form of the converb. The numbers indicate the percentage of constructions with the given morphological form that carry the information status or role in question.

	IU (%)	IE (%)			
		Fr	BGoC	Foc	PrE
des	39	20	21	14	6
mata	43	4	20	22	11
maks	63	1	1	35	0

The differences in the information status and role of -*des*, -*mata* and -*maks* constructions are primarily explained by their differences in semantic function, word order and punctuation (see Sections 7.2, 7.3, 7.5 and 7.7). If semantic function, word order and punctuation are similar, then their information status and role are also likely to be similar, regardless of the morphological form of the converb. Due to the fact that there are explanatory variables that are more strongly correlated with the information status and role than the morphological form of the converb, the variation in the information structure of -*des*, -*mata* and -*maks* constructions can be predicted by the same general variables.

However, morphological form serves to explain some relatively regular cases in which some explanatory variable deviates from the general pattern. For example, morphological form helps to explain how the position of the -*maks* converb (see Section 7.2) and the punctuation of the -*maks* construction (see Section 7.3) relate to information status.

7.2 Position of the converb

99 % of Estonian converb constructions either begin or end with the converb.[8] Both -*des* and -*mata* converbs may occur alone, as single-word constructions, and over half of -*des* and -*mata* converbs are the last constituents of the construction. -*maks* converbs, on the other hand, do not occur alone and are overwhelmingly positioned as the first constituents of the construction. Table 6 gives an overview of the distribution of converb position by the form of the converb.

Converb-initial constructions typically function as information units, while converb-final and single-word constructions usually function as elements of an information unit. Converb-initial constructions generally feature a construction-final focus, which is a characteristic of information units (48; see Section 3.2).

Table 6: Position of the converb by morphological form. The numbers indicate the percentage of converbs of the given form that occur in the position in question.

	First	Last	Alone	Other
des	290 (35 %)	**421 (51 %)**	98 (12 %)	16 (2 %)
mata	159 (33 %)	**240 (50 %)**	70 (15 %)	9 (3 %)
maks	**495 (99 %)**	5 (1 %)	0	0

Values in bold indicate the most frequent position of each morphological form.

---

8 This figure includes single-word constructions, which are both converb-initial and converb-final.

(47)   [*Tema jagab   momendil   6.-7. kohta*]_{IU}, [*jäädes*
       he   share.3SG moment:ADE 6.-7. place.PRT stay.DES
       favoriitidest   maha   [28   sek]_{Foc}]_{IU}.
       Favorite.PL.ELA behind 28   sec.PRT
       'He's currently tied for 6th place, **staying 28 seconds behind the favorites**.'
       (NEWS)

Converb-final and single-word constructions usually do not contain a focus; rather, they contextualize the focus of the main clause or function as the focus of the whole sentence (49).

(48)   [*Ihalesin   neid   **[isegi hukatuse   äärel   seistes]***_{Foc}]_{IU}.
       Desire.PST.1SG they.PL.PRT even ruin.GEN edge.ADE stand.DES
       'I desired them **even standing on the edge of ruin**.'
       (FIC)

Table 7 summarizes the relationship of the position of the converb with information status. Only construction-initial and construction-final converbs are shown in the table, because converbs positioned between modifiers are too rare (cf. Table 6) and single-word converb constructions always function as elements.

     Roughly one third of converb-initial *-maks* constructions and one fifth of converb-final *-mata* constructions deviate from the general pattern described above. *-maks* constructions tend to begin with the converb regardless of information status (49).

(49)   [*Savi on   neutraalne ja   kompetentne isik   **[juhtimaks***
       Savi be.3SG neutral   and competent   person lead.MAKS
       Riigikogu]_{Foc}]_{IU}.
       Parliament.PRT
       'Savi is a neutral and competent person **to lead the Parliament**'
       (BC).

Table 7: Information status of the converb construction by the position of the converb. The numbers indicate how often a converb-initial construction functions as an information unit and how often a converb-final construction functions as an element.

	First (IU)	Last (IE)
des	98 %	93 %
mata	94 %	80 %
maks	64 %	100 %

The strong tendency of the -*maks* form to appear at the beginning of a construction could be explained by its artificial origin, stylistic markedness, association with specific registers, syntactic restrictedness, semantic meaning of purpose and its functions as a complement and a bound modifier (Erelt 2017c: 801–802; Uuspõld 1980: 729, 736). The relevance of these factors is a matter for a separate study, but in any case the -*maks* form tends to occur construction-initially regardless of information status.

Almost one fifth of converb-final -*mata* constructions function as information units, consisting solely of a focus (26) and complementing the preceding main clause with a supplementary piece of information.

(50) [*Kuid   nad   olid   tulnud   [igaüks   oma*
     But   they be.PST.3PL come.PST.PTCP each   own.GEN
     vabast   tahtest]_{Foc}]_{IU}, **[*kellegi   käsku   või***
     free.ELA   will.ELA   who.GEN=CL   order.PRT   or
     korraldust   kuulmata*]_{IU}.
     instruction:PRT   hear:MATA
     'But they had each come of their own free will, **without following anyone's orders or instructions.'**
     (FIC)

In summary, the position of the converb indicates the presence of a focus, the construction's informational completeness and information status. A construction beginning with the converb resembles a subject-less finite clause: it begins with the core element – the converb or predicate – and ends with a focused modifier. Single-word and converb-final constructions are similar to typical core-final (noun) phrases in Estonian, which usually do not contain a distinct focus and are not informationally independent.

7.3 Punctuation

Punctuation is the clearest formal indicator of the information status of Estonian converb constructions (cf. Figure 1). The relationship between punctuation and information structure is difficult to interpret, because in written Estonian, punctuation is regulated by rules and the material of this study consists of edited texts. However, punctuation is an empirical linguistic feature whose consideration as an explanatory variable does not diminish the value of other explanatory variables (see Section 6.2). Therefore, I will describe how strongly and in what way the punctuation of the converb construction is correlated with information status. The interpretation of the extent to which the relationship between punctuation and information status is

based on the Simmul's or editor's perception of meaning, pause or rules is beyond the scope of the tree and forest method (see Section 6) as well as of this study.

The existence of punctuation rules does not in itself mean that punctuation is arbitrary in the sense of lack of information-structural function or motivation. Rules are one factor that affects the punctuation and that may interact in many ways with other factors, e.g. perception of meaning and pause. The punctuation rules concerning the Estonian converb constructions are not explicitly tied to information structure but nevertheless align well with information status and may be implicitly based on information structure.

Estonian punctuation rules have focused on word order, and mainly the position of the converb. According to the general rule, a converb-initial construction should be separated from the main clause by punctuation (example 23), but a converb-final or single-word construction should not be separated (example 24; Erelt 2006: 151–152; Saari 1993: 400; Vääri 1980: 144–145). On the level of information structure (see Section 7.2), this rule essentially implies that an information unit, i.e. a construction containing a focus, should be separated from the main clause, while an information element, i.e. a construction not containing a focus, should not be separated from the main clause. Thus there is a clear functional link between the punctuation rule and information status: punctuation marks the boundary between information units.

Most (96 %) Estonian converb constructions functioning as information units are separated from the main clause by punctuation (example 48), whereas most (95 %) of those functioning as elements are not (example 49). Sometimes, a converb construction functioning as an information unit occurs without punctuation (52). Most (72 %) of these cases are -*maks* constructions, which can be explained by the fact that punctuation in -*maks* constructions is by rule considered optional (Erelt 2006: 151–152) and the avoidance of punctuation has sometimes been preferred (see Saari 1993: 400).

(51)　　　[*Belgia　　　sadamatöölised　alustasid　　　esmaspäeval*
　　　　　Belgian.GEN　port-worker-PL　begin.PST.3PL　Monday.ADE
　　　　　[*ööpäevast　streiki*]Foc]IU **[*väljendamaks　oma***
　　　　　24-hour.PRT　strike.PRT　　express.MAKS　　own.GEN
　　　　　pahameelt　　　　plaanidele* [*avada　kaupade　　käsitlemine
　　　　　dissatisfaction.PRT　plan.PL.ALL　open.INF　goods.PL.GEN　handling
　　　　　***konkurentsile*]Foc]IU**.
　　　　　Competition.ALL
　　　　　'On Monday, Belgian port workers began a 24-hour strike **to express their dissatisfaction with the plans to open the handling of goods to competition**'
　　　　　(BC)

Table 8: Number of words in the converb construction and number of modifiers of the converb by information status.

	NoW	NoM
IU	6.4	1.4
IE	3.6	1

On the basis of English, punctuation and intonation units have been treated as oral and written manifestations of information units (Chafe 1988; Halliday 1967: 201, 1985; Moore 2016). Similarly, it can be said that punctuation indicates the information status of Estonian converb constructions quite accurately, though not perfectly. The question through which processes the correspondence between punctuation and information status emerges requires a separate analysis.

7.4 Number of words in the converb construction and number of modifiers of the converb

A converb construction functioning as an information unit contains on average nearly twice as many words and one-third more modifiers than a construction functioning as an element. The relationship of information status to the number of words and modifiers is summarized in Table 8.

An information unit typically contains more words and modifiers than an information element. However, the relationship between information status and the number of words and modifiers is far from absolute. An information element (example 42) can be longer and contain more modifiers than an information unit (example 8). Thus the number of words and modifiers by itself is not a definitive indicator of information status.

7.5 Position of the converb construction

The position of the converb construction is to some extent correlated with its information status and strongly correlated with its information role. The distribution of position of constructions functioning as elements is relatively even, but constructions functioning as information units are usually postposed. Table 9 provides an overview of these relationships.

Table 9: Information status of the converb construction by its position. The numbers indicate what percentage of constructions with a given information status occur in each position.

	IU	IE
Pre	189 (22 %)	263 (27 %)
Intra	43 (5 %)	312 (32 %)
Post	610 (72 %)	386 (40 %)
Total	842	961

Converb constructions functioning as information units are rarely interposed, because a sentence in which one information unit is coopted (see Heine et al. 2017) between the parts of another requires tracking two ideas at once, which is complicated and risks violating the maxim of manner (see Grice 1975: 46).

The position of the converb construction is the main indicator of its information role. If a converb construction functions as an information element, its position relates to the information role like the position of a typical Estonian phrase, except for the role of topic (see Lindström 2017: 543, 549–550; Tael 1988: 38–40). A postposed construction usually functions as a focus (50), an interposed construction is typically in the role of background part of the comment (40), and a preposed construction is typically a frame (37). Table 10 gives a breakdown of the frequency of information roles in different positions.

The relationship of the position of the information element to information role is somewhat weakened by the prominent element role, as prominent elements occur relatively equally in both preposed (45) and interposed (46) positions. The characteristic emphasis of a prominent element is primarily marked not by position, but by the presence of a lexicogrammmatical marker.

Table 10: Information role of the converb construction by position.

	Fr	BGoC	Foc	PrE
Pre	**191 (73 %)**	9 (3 %)	2 (1 %)	61 (23 %)
Intra	0	**258 (83 %)**	14 (4 %)	40 (13 %)
Post	0	2 (1 %)	**383 (99 %)**	1

Values in bold indicate the most frequent position of each morphological form.

7.6 Lexicogrammatical prominence markers

In Estonian, lexicogrammatical prominence markers are an important means of emphasis (Erelt 2017a: 53–54; Lindström 2017: 545–546). Such markers are found in roughly half (52 %) of converb constructions functioning as a prominent element (example 53) and in one ninth (11 %) of those functioning as a focus (54).

(52)   (*Kõiki neid unistusi Billy otsesõnu välja ei ütle*),
         [*kuid   see   on   **[ütlemata=gi]**_PrE   [arusaadav]_Foc]_IU.*
         but   it   is   say.MATA=CL   clear
         '(Billy doesn't express all those dreams explicitly), but it's clear **even when
         left unsaid**'
         (FIC).

(53)   [*Maailmas   viljeldakse   maad   **[ka   hoopis   ilma**
         world.INE   cultivate.IPS   land.PRT   also   completely   without
         ***kündmata]**_Foc]_IU.*
         plough.MATA
         'In some parts of the world land is cultivated **even without ploughing**'
         (NEWS).

Converb constructions functioning as information units are rarely (4 %) emphasized by prominence markers. The roles of frame or background part of comment are not marked by prominence markers, because according to the definition used in this study (see Section 4.2), the emphasis of the prominence marker would change the role of frame and the role of background part of comment into the role of prominent element.

   As the primary indicator of a focus in Estonian is its information unit-final position and an element in its typical word order position is usually not lexically marked (see Lindström 2017: 544–545), lexicogrammatical markers are used to highlight only those foci which are particularly notable, e.g. those that run counter to expectation (examples 43, 49, 54).

   However, lexicogrammatical markers are the primary means of highlighting prominent elements, because the prominent – final and initial – word order positions are typically occupied by the focus and the topic or frame (see Lindström 2017: 547, 549, 552). The division of prominence-marking work between the focus and the prominent element is thus economical; the focus, as the more common information role, is typically marked by word order, while the less common role of prominent element is marked via additional lexicogrammatical means.

7.7 Semantic function

The relationship between semantic function and information structure is diverse, but some concordant patterns do emerge. Table 11 summarizes the occurrence of different semantic functions in different information status and roles. Functions 1–7 are characteristic of information units and functions 8–15 are characteristic of information elements.

The converb construction has a twofold meaning. It serves to expresses an event, while simultaneously indicating that this event is a circumstance of a larger event. In general it can be said that in the case of constructions with functions 1–7 the facet of an event tends to be contextually prominent. Conversely, for constructions with functions 8–15 the facet of a circumstance tends to be contextually prominent.

On the informational level, functions 1–7 tend to be associated with relative conceptual independence and completeness of the event. Converb constructions with functions 1–7 usually express a different situation (32) or the same situation from a different perspective (48) than the main clause. On the information level, converb constructions with these functions relate to the main clause as a comparative whole to a whole.

Table 11: Information status and information role of the converb construction by semantic function. The most frequent information status or role for each semantic function is shown in bold.

	SF	IU (%)	IE (%)			
			Foc	**BGoC**	**PrE**	**Fr**
1	C_SPECIFICATION	**100**	0	0	0	0
2	C_CONTRAST	**97**	3	0	0	0
3	C_CAUSE	**83**	2	7	1	7
4	c_none	**81**	6	6	1	6
5	C_RESULT	**78**	12	6	4	0
6	C_PURPOSE	**63**	35	1	0	0
7	T_CAUSE	**53**	3	9	6	28
8	C_CONCESSION	26	**43**	15	17	0
9	C_MEANS	15	**35**	26	4	20
10	C_MANNER	10	**38**	38	14	0
11	T_CONCESSION	18	18	9	**45**	9
12	C_CONDITION	21	8	4	**38**	29
13	T_MEANS	18	5	20	5	**51**
14	T_none	12	9	30	6	**43**
15	T_CONDITION	20	4	25	10	**41**

Values in bold indicate the most frequent position of each morphological form.

Functions 8–15, on the other hand, tend to be associated with informational partiality and conceptual incompleteness of the event of the converb construction. Expressing a circumstance of the event of the main clause, converb constructions with these functions relate to the main clause as a part to the whole. Furthermore, semantic functions 8–15 tend to be associated with different information roles. The role of focus is most characteristic of functions 8–10, in which the circumstance of the construction is interpreted as the most unpredictable aspect of the event of the sentence (example 53). The role of frame is most typical of functions 13–15, wherein the circumstance of the construction contextualizes the event of the main clause (example 36). Functions 11–12 are associated with the role of prominent element; constructions carrying this function tend to express a remarkable, though somewhat peripheral circumstance (example 52).

To some extent, the observed patterns of association between semantic function and information structure may apply to other constructions as well. This possibility is suggested by the typical morphosyntactic structures commonly associated with these functions in Estonian. Semantic functions SPECIFICATION, CONTRAST, CAUSE, RESULT and PURPOSE, characteristic of the information unit status, tend to be expressed by clausal-sentential and co-ordinated morphosyntactic structures, in contrast to functions like MANNER and MEANS, which are more often expressed as phrases (see Erelt 2017b: 611–620; Erelt et al. 2020: 429–430; Hennoste 2017: 486; Veismann et al. 2017). The relationships between semantic function, information structure and morphosyntactic structure is a wide-ranging topic worthy of further study encompassing a variety of linguistic units.

8 Conclusions

This paper has described the variation in the information status and information role of Estonian *-des*, *-mata* and *-maks* converb constructions. Estonian converb constructions may function as information units, expressing a complete idea, or as information elements of a larger information unit, expressing a part of an idea. In the latter case, converb constructions take on different informational roles, namely the role of focus, background of the comment, frame and prominent element.

Neither the information status nor the information role of converb constructions is unambigously determined by any of the morphological, syntactic, semantic or orthographic features analyzed in this study. Nevertheless, some of these features

are so strongly correlated with information status or role that they effectively predict the variation in the information structure of converb constructions.

This paper has discussed seven features which predict the variation in information status of converb constructions. These features are the presence of punctuation separating the converb construction from the main clause, the position of the converb within the construction and the position of the converb construction with regards to the main clause, the semantic function of the construction, the morphological form of the converb, the number of words in the construction, and the number of modifiers of the converb.

Converb constructions functioning as information units tend to be separated from the main clause by punctuation, begin with the converb, and be positioned after the main clause. Converb constructions functioning as information elements, by contrast, tend to be one-word or converb-final constructions, are usually not separated from the main clause by punctuation and are fairly evenly distributed before, within, and after the main clause. Semantically, information units tend to carry the basic function of CONCOMITANCE and the complementary functions of SPECIFICATION, CONTRAST, CAUSE, RESULT and PURPOSE. Information elements are associated with the basic function of TIME and the complementary functions of MANNER, MEANS, CONDITION and CONCESSION. On average, information units are longer and contain more modifiers than information elements (1–2 modifiers for information units and 1 modifier for information elements). Over half of *-des* and *-mata* constructions function as information elements, while *-maks* constructions mainly function as information units.

Three features have been identified that predict the variation in information role among converb constructions functioning as information elements: the position of the construction, the presence of a lexicogrammatical prominence marker, and the semantic function. Preposed constructions usually carry the role of frame, interposed constructions carry the role of background part of the comment, and postposed constructions carry the role of focus. Preposed and interposed constructions which contain a lexicogrammatical prominence marker typically carry the role of prominent element.

Semantically, the role of focus is associated with the complementary functions of MANNER, MEANS and CONCESSION, the role of frame with the basic function of TIME and the complementary functions of CONDITION and MEANS, and the role of prominent element with the functions of CONCOMITANCE-CONDITION and TIME-CONCESSION. The role of background part of the comment frequently carries the basic function of TIME and the complementary functions of MANNER and MEANS.

Table 12 summarizes the key features characteristic of the information status and role of converb constructions.

Table 12: Characteristic features of information status and roles.

	IU	IE			
		Foc	**BGoC**	**Fr**	**PrE**
Punct	yes	no	no	no	no
PoCC	post	post	intra	pre	pre, intra
PoCv	first	first, last	last, alone	last	last, alone
SF	C_SPEC, C_CONT, C_CA, C_RES, C_none, C_PURP, T_CA	C_CONC, C_ME, C_PURP, C_MA	C_MA, C_ME, T_none	T_ME, T_none, T_COND	T_CONC, C_COND
PrM	no/yes	no/yes	no	no	yes/no
Form	*maks, des, mata*	*maks, mata, des*	*des, mata*	*des*	*mata, des*

This paper has demonstrated the usefulness of an idea-based and focus-centered category of information unit: a linguistic unit that has an information structure of its own. The category of information unit is connected to the category of focus: each information unit has one and only focus. An information unit consists of one focus or one focus and background, i.e. information elements that are organized around this focus. The information unit provides a common unit of analysis of simple sentences, complex sentences, independent and subordinate clauses. The information unit category is particularly crucial for analyzing the information structure of the subordinate clause of a complex sentence. If a clause functions as an information unit, its internal information structure is of primary importance. If a clause functions as an information element, its most relevant feature is its external information structure, i.e. information role.

The analysis herein illustrates the usefulness of the well-known information role categories of focus, topic, and comment, and adds to these the roles of frame and prominent element, which facilitate the detailed description of complex information units.

This paper has explicated the extent of information-structural variation in the Estonian converb construction, which ranges from sentence-like to phrase-like. On the one hand, a converb construction can have its own information structure, such that it resembles a sentence. On the other hand, a converb construction can be part of a larger information structure, such that it is similar to a phrase.

This analysis shows how the Estonian converb construction as a subordinate non-finite clause can have its own information structure. Thereby, the analysis serves to illustrate the independence of the syntactic level and the information level. There is no one-to-one correspondence between syntactic embeddedness and

informational embeddedness. The syntactically subordinate clause does not necessarily belong to the same information unit as the main clause. A subordinate clause can have its own focus and an information structure distinct from the main clause. An information unit can be a smaller linguistic unit than a sentence. A sentence can contain multiple foci and multiple information units.

This analysis also shows how the Estonian converb construction can function as an information element, participating in an information structure organized around a focus that connects the converb construction with other constituents of the main clause. A converb construction which functions as an information element is similar to a typical Estonian phrase both in terms of possible information roles and in terms of how these information roles are expressed. The converb construction can have the same information roles as other Estonian information elements, except for topic, which is not characteristic of converb constructions. Also, the information role of the converb construction is primarily expressed by word order and lexicogrammatical markers, which also express the information role of a typical Estonian phrase.

However, the Estonian converb construction is neither a completely typical (sentence-like) information unit nor a completely typical (noun phrase-like) information element. Both the main difference from the typical information unit and the main difference from the typical information element relate to the category of topic. The Estonian converb construction generally does not contain an explicit topic nor does it take on the information role of topic.

The extensive information-structural variability of the Estonian converb construction is enabled by the functional variability of its formal features. The word order, punctuation, length, lexicogrammatical markers and modifiers of the converb construction vary widely and functionally with respect to information structure. The central feature is the syntactically flexible and information-structurally functional word order, which is characteristic of the Estonian language in general. In expressing information structure, word order is supported by punctuation, which is closely related to word order, but is even more strongly correlated with the information status of converb construction than is word order.

In the future, it is worth investigating to what extent the dimensions of the variation in the information structure described here are general and to what extent they are specific to Estonian. On the one hand, the information structure of converb constructions may be more limited in a language where the word order is not as flexible and expressive as in Estonian. On the other hand, it may not be impossible for a converb construction to contain an explicit topic or function as a topic in a language where it is not prevented by grammatical constraints or other constructions more established in the informational role of topic.

Abbreviations

1–3	person
ABL	ablative
ADE	adessive
ALL	allative
CL	clitic
CMP	comparative
CNG	connegative
COM	comitative
COND	conditional
DES	*-des* converb
ELA	elative
GEN	genitive
INF	infinitive
INE	inessive
IPS	impersonal
MAKS	*-maks* converb
MATA	*-mata* converb
NEG	negation
PL	plural
PRT	partitive
PST	past
PTCP	patriciple
Q	question word
SG	singular
SUP	supine
TRSL	translative

Corpus references

NEWS = Newspaper texts from the years 1990–1999. korp.keeleressursid.ee.
FIC = Fiction texts from the years 1990–1999. korp.keeleressursid.ee.
BC = The Balanced Corpus of Estonian. http://www.keeleveeb.ee/.

References

Adamou, Evangelia, Katharina Haude & Matrine Vanhove. 2018. Investigating information structure in lesser-known and endangered languages: An introduction. In Evangelia Adamou, Katharina Haude & Matrine Vanhove (eds.), *Information structure in lesser-described languages: Studies in prosody and syntax* (Studies in Language Companion Series 199), 1–14. Amsterdam: John Benjamins.
Asu, Eva Liina, Pärtel Lippus, Pajusalu Karl & Pire Teras. 2016. *Eesti keele hääldus* [The pronunciation of Estonian] (Eesti keele varamu II). Tartu: Tartu Ülikooli kirjastus.

Baayen, R. Harald, Anna Hendersen, Laura A. Janda, Anastasia Makarova & Tore Nesset. 2013. Making choices in Russian: Pros and cons of statistical methods for rival forms. *Russian Linguistics* 37(3). 253–291.

Bisang, Walter. 2020. Verb serialization and converbs – differences and similarities. In Martin Haspelmath & Ekkehard König (eds.), *Converbs in cross-linguistic perspective: Structure and meaning of adverbial verb forms – adverbial participles, gerunds*, 137–188. Berlin, Boston: De Gruyter Mouton.

Breiman, Leo. 2001. Random forests. *Machine Learning* 45. 5–32.

Chafe, Wallace. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Charles N. Li (ed.), *Subject and topic*, 25–55. New York: Academic Press.

Chafe, Wallace. 1979. The flow of thought and the flow of language. In Thomas Givón (ed.), *Discourse and syntax*, 159–181. New York: Academic Press.

Chafe, Wallace. 1985. Linguistic differences produced by differences between speaking and writing. In David R. Olson, Andrea Hildyard & Nancy Torrance (eds.), *Literacy, language and learning. The nature and consequences of reading and writing*, 105–123. Cambridge: Cambridge University Press.

Chafe, Wallace. 1987. Cognitive constraints on information flow. In Tomlin Russell (ed.), *Coheremce and grounding in discourse*. Typological Studies in Language, vol. IX, 21–52. Amsterdam: John Benjamins.

Chafe, Wallace. 1988. Punctuation and the prosody of written language. *Written Communication* 5(4). 395–426.

Chafe, Wallace. 1992. Information flow in speaking and writing. In Pamela Downing, Susan D. Lima & Michael Noonan (eds.), *The linguistics of literacy*, 17–29. Amsterdam: John Benjamins.

Creissels, Denis. 2010. Specialized converbs and adverbial subordination in Axaxdərə Akhvakh. In Isabelle Bril (ed.), *Clause-hierarchy and clause-linking: Syntax and pragmatics*, 105–142. Amsterdam: John Benjamins.

Croft, William. 2012. *Verbs: Aspect and causal structure*. Oxford: Oxford University Press.

Diessel, Holger. 2005. Competing motivations for the ordering of main and adverbial clauses. *Linguistics* 43(3). 449–470.

Ebert, Cornelia. 2009. *Quantificational topics – a scopal treatment of exceptional wide scope phenomena* (Studies in Linguistics and Philosophy 86). Berlin: Springer.

Erelt, Mati, Tiiu Erelt & Kristiina Ross. 2020. *Eesti keele käsiraamat* [Handbook of Estonian]. Tallinn: Eesti Keele Instituut.

Erelt, Mati. 2006. *Lause õigekeelsus. Juhatused ja harjutused* [Sentence ortography. Directions and exercises]. Tartu: Emakeele Selts.

Erelt, Mati. 2017a. Sissejuhatus süntaksisse. In Mati Erelt & Helle Metslang (eds.), *Eesti keele süntaks* [The syntax of Estonian] (Eesti keele varamu III), 53–89. Tartu: Tartu Ülikooli kirjastus.

Erelt, Mati. 2017b. Rinnastus. In Mati Erelt & Helle Metslang (eds.), *Eesti keele süntaks* [The syntax of Estonian] (Eesti keele varamu III), 603–646. Tartu: Tartu Ülikooli kirjastus.

Erelt, Mati. 2017c. Sekundaartarindiga laused. In Mati Erelt & Helle Metslang (eds.), *Eesti keele süntaks* [The syntax of Estonian] (Eesti keele varamu III), 756–840. Tartu: Tartu Ülikooli kirjastus.

Erteschik-Shir, Nomi. 2007. *Information structure: The syntax-discourse interface*. Oxford: Oxford University Press.

Gijn, Rik van, Jeremy Hammond, Dejan Matić, Saskia van Putten & Ana Vilacy Galucio (eds.), 2014. *Information structure and reference tracking in complex sentences* (Typological Studies in Language 105). Amsterdam: John Benjamins.

Grice, H. Paul. 1975. Logic and conversation. In Peter Cole & Jerry L. Morgan (eds.), *Syntax and semantics, Vol. 3, Speech acts*, 41–58. New York: Academic Press.

Gundel, Jeanette & Thorstein Fretheim. 2004. Topic and focus. In Larry Horn & Gregory Ward (eds.), *The handbook of pragmatics*, 175–196. Malden, MA: Blackwell Publishers.

Halliday, Michael A. K. 1967. Notes on transitivity and theme in English: Part 2. *Journal of Linguistics* 3(2). 199–244.

Halliday, Michael A. K. 1985. *An introduction to functional grammar*. London: Arnold.

Haspelmath, Martin. 1995. The converb as a cross-linguistically valid category. In Martin Haspelmath & Ekkehard König (eds.), *Converbs in cross-linguistic perspective*, 1–55. Berlin: Mouton de Gruyter.

Heine, Bernd, Gunther Kaltenböck, Tania Kuteva & Haiping Long. 2017. Cooptation as a discourse strategy. *Linguistics* 55(4). 813–855.

Hennoste, Tiit. 2017. Üldlaiend, kiil, irdelemendid. In Mati Erelt & Helle Metslang (eds.), *Eesti keele süntaks* [The syntax of Estonian] (Eesti keele varamu III), 481–502. Tartu: Tartu Ülikooli kirjastus.

Ibarluzea, Patxi Laskurain. 2014. Mood selection in the complement of negation matrices in Spanish. In Rik van Gijn, Jeremy Hammond, Dejan Matić, Saskia van Putten & Ana Vilacy Galucio (eds.), *Information structure and reference tracking in complex sentences* (Typological Studies in Language 105), 193–228. Amsterdam: John Benjamins.

Jacobs, Joachim. 2001. The dimensions of topic–comment. *Linguistics* 39(4). 641–681.

Killie, Kristin & Toril Swan. 2009. The grammaticalization and subjectification of adverbial *-ing* clauses (converb clauses) in English. *English Language and Linguistics* 13(3). 337–363.

Kinberg, Naphtali. 2001. Adverbial clauses as topics in Arabic: Adverbial clauses in frontal position separated from their main clause. In Naphtali Kinberg & Kees Versteegh (eds.), *Studies in the linguistic structure of classical Arabic*, 43–102. Leiden: Brill.

Klavan, Jane, Maarja-Liisa Pilvik & Kristel Uiboaed. 2015. The use of multivariate statistical classification models: Comparing textual and behavioral evidence. *SKY Journal of Linguistics* 28. 187–224.

Klumpp, Gerson & Elena Skribnik. 2022. Information structuring. In Marianne Bakró-Nagy, Johanna Laakso & Elena Skribnik (eds.), *The Oxford guide to the Uralic languages*, 1018–1036. Oxford: Oxford University Press.

Komagata, Nobo. 2003. Information structure in subordinate and subordinate-like clauses. *Journal of Logic, Language and Information* 12. 301–318.

Komen, Erwin R. 2014. Chechen extraposition as an information ordering strategy. In Rik van Gijn, Jeremy Hammond, Dejan Matić, Saskia van Putten & Ana Vilacy Galucio (eds.), *Information structure and reference tracking in complex sentences*, 99–126. Amsterdam: John Benjamins.

Kortmann, Bernd. 1991. *Free adjuncts and absolutes in English: Problems of control and interpretation*. London/New York: Routledge.

Krifka, Manfred. 2007. Basic notions of information structure. In Catherine Féry, Gisbert Fanselow & Manfred Krifka (eds.), *The notions of information structure* (Interdisciplinary Studies on Information Structure 6 (2007)), 13–55. Potsdam: Universitätsverlag Potsdam.

Krifka, Manfred & Catherine Féry. 2008. Information structure. Notional distinctions, ways of expression. In Piet van Sterkenburg (ed.), *Unity and diveristy of languages*, 123–136. Amsterdam: John Benjamins.

König, Ekkehard. 1995. The meaning of converb constructions. In Martin Haspelmath & Ekkehard König (eds.), *Converbs in cross-linguistic perspective* (Empirical Approaches to Language Typology 13), 57–96. Berlin: Mouton de Gruyter.

Lambrecht, Knud. 1994. *Information structure and sentence form*. Cambridge: Cambridge University Press.

Lambrecht, Knud. 2000. When subjects behave like objects: An analysis of the merging of S and O in sentence-focus constructions across languages. *Studies in Language* 24(3). 611–682.

Leino, Jaakko. 2013. Information structure. In Graeme Trousdale & Thomas Hoffmann (eds.), *The Oxford handbook of construction grammar*, 329–344. Oxford: Oxford University Press.

Levshina, Natalia. 2015. *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam: John Benjamins.

Lindström, Liina. 2005. *Sõnajärg ja seda mõjutavad tegurid suulises eesti keeles* [The position of the finite verb in a clause: Word order and the factors affecting it in spoken Estonian] (Dissertationes philologiae estonicae Universitatis Tartuensis 16). Tartu: Tartu Ülikooli kirjastus.

Lindström, Liina. 2006. Infostruktuuri osast eesti keele sõnajärje muutumisel [On the role of information structure in the change of Estonian word order]. *Keel ja Kirjandus* 49(11). 875–888.

Lindström, Liina. 2017. Lause infostruktuur ja sõnajärg. In Mati Erelt & Helle Metslang (eds.), *Eesti keele süntaks* [The syntax of Estonian] (Eesti keele varamu III), 537–565. Tartu: Tartu Ülikooli kirjastus.

Matić, Dejan & Daniel Wedgwood. 2013. The meanings of focus: The significance of an interpretation-based category in cross-linguistic analysis. *Journal of Linguistics* 49(1). 127–163.

Matić, Dejan, Rika van Gijn & Robert D. van Valin, Jr. 2014. In Rik van Gijn, Jeremy Hammond, Dejan Matić, Saskia van Putten & Ana Vilacy Galucio (eds.), *Information structure and reference tracking in complex sentences* (Typological Studies in Language 105), 1–41. Amsterdam: John Benjamins.

Matić, Dejan & Irina Nikolaeva. 2018. From polarity focus to salient polarity: From things to processes. In Christine Dimroth & Stefan Sudhoff (eds.), *The grammatical realization of polarity contrast: Theoretical, empirical, and typological approaches* (Linguistik Aktuell/Linguistics Today 249), 9–54. Amsterdam: John Benjamins.

Moore, Nick. 2016. What's the point? The role of punctuation in realising infomation structure in written English. *Functional Linguistics* 3(6). 1–23.

Nedjalkov, Igor. 1998. Converbs in the languages of Eastern Siberia. *Language Sciences* 20(3). 339–351.

Nedjalkov, Vladimir P. 1995. Some typological parameters of converbs. In Martin Haspelmath & Ekkehard König (eds.), *Converbs in cross-linguistic perspective* (Empirical Approaches to Language Typology 13), 97–136. Berlin: Mouton de Gruyter.

Nikolaeva, Irina. 2001. Secondary topic as a relation in information structure. *Linguistics* 39(1). 1–49.

Plado, Helen. 2015a. *des-* ja *mata*-konverbi kasutusest eesti murretes [On the use of -*des* and -*mata* converb in dialects of Estonian]. *Emakeele Seltsi Aastaraamat* 60(2014). 195–218.

Plado, Helen. 2015b. The subject of the Estonian *des*-converb. *SKY Journal of Linguistics* 28. 313–348.

Prince, Ellen F. 1981. Toward a taxonomy of given-new information. In Peter Cole (ed.), *Radical pragmatics*, 223–255. New York: Academic Press.

Prince, Ellen F. 1992. The ZPG Letter: Subjects, definiteness, and information-status. In William C. Mann & Sandra A. Thompson (eds.), *Discourse description: Diverse linguistic analyses of a fund-raising text*, 295–325. Amsterdam: John Benjamins.

Putten, Saskia van. 2014. Left dislocation and subordination in Avatime (Kwa). In Rik van Gijn, Jeremy Hammond, Dejan Matić, Saskia van Putten & Ana Vilacy Galucio (eds.), *Information structure and reference tracking in complex sentences* (Typological Studies in Language 105), 71–98. Amsterdam: John Benjamins.

Reesink, Ger P. 2014. Topic management and clause combination in the Papuan language Usan. In Rik van Gijn, Jeremy Hammond, Dejan Matić, Saskia van Putten & Ana Vilacy Galucio (eds.), *Information structure and reference tracking in complex sentences* (Typological Studies in Language 105), 231–262. Amsterdam: John Benjamins.

Reinhart, Tanya. 1981. Pragmatics and linguistics: An analysis of sentence topics in pragmatics and philosophy I. *Philosophica anc Studia Philosophica Gadensia Gent* 27(1). 53–94.

Remmel, Nikolai. 1963. Sõnajärjestus eesti lauses [Word order in Estonian sentence]. In *Eesti keele süntaksi küsimusi* (Keele ja Kirjanduse Instituudi uurimused VIII), 216–389. Tallinn: Eesti Riiklik Kirjastus.

Saari, Henn. 1993. Lisa: Kiri. In Mati Erelt, Reet Kasik, Helle Metslang, Henno Rajandi, Kristiina Ross, Henn Saari, Kaja Tael & Silvi Vare (eds.), *Eesti keele grammatika II* [The grammar of Estonian II]. Süntaks. Lisa: Kiri, 323–425. Tallinn: EKI.

Sahkai, Heete. 1999. Eesti verbifraasi sõnajärg [Word order of Estonian verb phrase]. *Keel ja Kirjandus* 42(1). 24–32.

Sahkai, Heete & Anne Tamm. 2019. Verb placement and accentuation: Does prosody constrain the Estonian V2? *Open Linguistics* 5(1). 729–753.

Shagal, Ksenia, Pavel Rudnev & Anna Volkova. 2022. Multifunctionality and syncretism in non-finite forms: An introduction. *Folia Linguistica* 56(3). 529–557.

Simmul, Carl Eric. 2018. *des*- ja *mata*-konverbitarindi funktsioonid [The semantic functions of Estonian -*des* and -*mata* converb construction]. Keel ja. *Keel ja Kirjandus* 59(11). 847–867.

Simmul, Carl Eric. 2020. Süüvides jonni jätmata avamaks sõnajärjemustreid. *des*-, *mata*- ja *maks*-konverbitarindi sõnajärg [Word order of Estonian -*des*, -*mata* and -*maks* converb construction]. Keel ja. *Keel ja Kirjandus* 61(3). 221–242.

Simmul, Carl Eric. 2021. *des*-, *mata*- ja *maks*-konverbitarindi inforoll [The information role of Estonian -*des*, -*mata* and -*maks* converb constructions]. *Journal of Estonian and Finno-Ugric Linguistics* 12(1). 303–334.

Storto, Luciana. 2014. Constituent order and information structure in Karitiana. In Rik van Gijn, Jeremy Hammond, Dejan Matić, Saskia van Putten & Ana Vilacy Galucio (eds.), *Information structure and reference tracking in complex sentences* (Typological Studies in Language 105), 163–192. Amsterdam: John Benjamins.

Strobl, Carolin, Anne-Laure Boulestix, Thomas Kneib, Thomas Augustin & Achim Zeileis. 2008. Conditional variable importance for random forests. *BMC Bioinformatics* 9(307). https://doi.org/10.1186/1471-2105-9-307.

Strobl, Carolin, James Malley & Gerhard Tutz. 2009. An introduction to recursive partitioning: Rationale application and characteristics of classification and regression trees, bagging and random forests. *Psychological Methods* 14(4). 323–348.

Tael, Kaja. 1988. *Sõnajärjemallid eesti keeles (võrrelduna soome keelega)* [Word order patterns in Estonian (as compared to Finnish)] (Preprint KKI-56). Tallinn: ENSV TA Keele- ja Kirjanduse Instituut.

Tagliamonte, Sali A. & R. Harald Baayen. 2012. Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change* 24(2). 135–178.

Thompson, Sandra A., Robert E. Longacre & Shin Ja J. Hwang. 2007. Adverbial clauses. In Shopen Timothy (ed.), *Language typology and syntactic description*, 2nd edn., 171–234. Cambridge: Cambridge University Press.

Uuspõld, Ellen. 1966. Määrusliku *des*-, *mata*-, *nud*- (~*nuna*-) ja *tud*- (~*tuna*-) konstruktsiooni struktuur ja tähendus [The structure and meaning of adverbial -des, -mata, -nud (-nuna) and -tud (-tuna) constructions]. In *Keele modelleerimise probleeme I* (Tartu Riikliku Ülikooli toimetised 188), 1–196. Tartu: Tartu Riiklik Ülikool.

Uuspõld, Ellen. 1980. *Maks*-vorm ja teised finaaladverbiaalid [-maks form and other final adverbials]. *Keel ja Kirjandus* 23(12). 729–736.

Valijärvi, Riitta-Liisa. 2003. Estonian converbs – with special emphasis on early 18th-century literary languages. *Ural-Altaische Jahrbücher* 18. 24–67.

Valin, Robert D. van, Jr. & Randy J. LaPolla. 1997. *Syntax: Structure, meaning and function*. Cambridge: Cambridge University Press.

Veismann, Ann, Mati Erelt & Helle Metslang. 2017. Määrus [Adverbial]. In Mati Erelt & Helle Metslang (eds.), *Eesti keele süntaks* (Eesti keele varamu III), 300–375. Tartu: Tartu Ülikooli kirjastus.

Vääri, Eduard. 1980. *Eesti keele õpik keskkoolile* [The handbook of Estonian for high school]. 10. trükk. Tallinn: Valgus.

Wal, Jenneke van der. 2014. Subordinate clauses and exclusive focus in Makhuwa. In Rik van Gijn, Jeremy Hammond, Dejan Matić, Saskia van Putten & Ana Vilacy Galucio (eds.), *Information structure*

and reference tracking in complex sentences (Typological Studies in Language 105), 45–70. Amsterdam: John Benjamins.

Ylikoski, Jussi. 2003. Defining non-finites: Action nominals, converbs and infinitives. *SKY Journal of Linguistics* 16. 185–237.

Zimmermann, Malte & Caroline Féry. 2010. Introduction – information structure. In Malte Zimmermann & Caroline Féry (eds.), *Information structure: Theoretical, typological, and experimental perspectives*, 1–11. Oxford: Oxford University Press.