# THIRD PERSON PRONOUN FORMS IN ESTONIAN IN THE LIGHT OF CENTERING THEORY

**Helen Hint**
*University of Tartu*

**Abstract.** This paper explains the distinctions between the Estonian 3rd person overt pronoun and the zero person marker in spoken narratives. As both forms express the most salient entities in discourse, the saliency criterion cannot distinguish them. The Centering Theory is used to explore if the overt pronoun and zero have different effects on discourse coherence, i.e. whether there is a difference between transition types relating to zero and those signaling the overt pronoun. Additionally, factors such as grammatical role, case and clause type affecting the choice of pronominal forms are studied to supplement results from the Centering analysis. It is hypothesized that the use of the zero form connects to the CONTINUE transition, while the overt pronoun combines with other Centering-based transition types as well. Furthermore, results show that the zero form is more restricted in its usage contexts and signals mainly nominative subjects in main clauses, while the overt form can appear more widely in different linguistic environments.

**Keywords:** zero person marker, third person pronoun, Centering Theory, reference resolution, spoken narrative, Pear Stories

## 1. Introduction

There are various forms available to a speaker who wants to refer to a particular entity. Expressions such as "a lady", "my mother", "she" and "Ülle" might all refer to the same person in different contexts. It is assumed that the choice of referring expression depends on the attentional/cognitive state of the referent: more reduced forms point to more salient entities, whereas more elaborate phrases are needed for less salient (new) entities (see e.g. Ariel 1990, Gundel et al. 1993). One theory concerned with the relationship between the attentional state and the form of referring expressions is Centering Theory (Grosz et al. 1995, Walker et al. 1998). As a theory of focus of attention and anaphora, it is applicable to analyze reference resolution.

This article employs the Centering framework to explore the distinctions between the use of the Estonian 3rd person overt pronoun and

the zero person marker in spoken narrative discourse. The Pear Film (Chafe 1980) is used to collect the narratives. In Estonian some contexts allow the use of zero forms as anaphoric devices, although this is not considered to be a default choice.[1] Rather, the 3rd person short pronoun *ta* 's/he' (PL *nad* 'they') is the most expected choice for referring to highly salient entities. Usually, zero can be replaced with an overt pronoun without any changes in the grammaticality of the clause. The opposite possibility is quite rare. It has been noted that in narrative contexts it is possible to leave out the subject if it remains unchanged in a longer sequence of sentences (Lindström 2005: 175). As yet there are no systematic analyses regarding the distinction of zero and the overt 3rd person pronoun *ta* in Estonian.

This paper aims to fill the gap by exploring the relation between Centering transitions and different forms of referential pronouns in Estonian spoken narratives. The main reason for applying the Centering Theory in the present work is that it is the only carefully implemented framework that allows us to test the differences between referential forms by following a rigid, fixed procedure. As far as I know, it is the first attempt to apply the Centering framework to Estonian. However, note that Centering is used here merely as an explanatory linguistic tool and not as a computational method with practical application. To give a more extensive picture of factors triggering the use of overt 3rd person pronouns or the zero person marker, the Centering analysis is supplemented by qualitative analysis considering various grammatical features of corresponding utterances.

The questions asked are as follows.

1) Is there a difference between the transition types signaling zero and those signaling the overt pronoun? It is hypothesized that the use of zero forms strongly connects to the CONTINUE transition. The relationship appears to be a cross-linguistic phenomenon (see Di Eugenio 1998, Iida 1998). At the same time, the pronoun *ta* should combine with other Centering-based transition types as well.

2) How do grammatical factors such as grammatical role, case and clause type (main *vs.* subordinate) influence the choice of pronominal forms, and how do they supplement the results from the Centering analysis?

The organization of the article is as follows. Section 2 gives a brief overview of Estonian 3rd person anaphoric expressions and the research

---

1    A similar phenomenon is observed for example in English as a diary-style null subject, see e.g. Scott (2013).

done in this area. The basic notions and ideas of Centering Theory are explained in section 3. In section 4, the material and method of the present study are introduced. Section 5 presents the results of the analysis and answers the research questions. Finally, general discussion and conclusions can be found in section 6.


## 2. The Estonian 3rd person anaphoric paradigm

Various anaphoric devices are available for referring back to 3rd person entities in Estonian. These include
1) short and long forms of personal pronouns: SG *ta*/*tema* 's/he' (PL *nad*/ *nemad* 'they');
2) demonstrative pronouns: SG *see* 'this' (PL *need* 'these') and SG *too* 'that' (PL *nood* 'those')
3) zero person marker (mostly zero subjects).

In this article, the term zero person marker stands for the absence of a phonological form in grammatical person interpretation, as Siewierska (2004: 22) has defined it. Note that Kibrik (2011: 121) includes only overt devices under the notion of pronoun and opposes zero reference to overt forms, thereby excluding zero reference from pronouns. However, as the present work considers zero forms in contexts similar to overt pronoun usage, the term zero pronominal is deemed more precise. Works discussing Estonian pronouns do not usually list the zero form as a feature of the language's pronominal system (e.g. Pajusalu 2005, 2009, Kaiser and Hiietam 2003). Estonian traditional grammar describes the phenomenon as contextual ellipsis, stating that it serves as an anaphoric substitute for a previously mentioned initial form (EKG II: 223). The term 'subjectless clause' has also been used (Lindström 2001).
It is true that the usage of referential zero person markers is somewhat restricted. According to the traditional view, only 1st and 2nd person pronouns can be realized by zero in sentences where the predicate verb form also carries a personal ending (EKG II: 223). However, the 3rd person zero form (ellipsis) is considered acceptable in subsequent sentences sharing similar context and a common subject (EKG II: 225). It has also been noted that the 3rd person zero is mainly characteristic of narrative contexts and spoken language (see Lindström 2005: 175, Kivik 2010). Lindström (2001) has studied subject ellipsis in narratives and concludes that ellipsis (zero subject) is possible in contexts where the agent remains the same for a longer period, and the utter-

ances containing zero build up a whole episode in the narrative. Also, entities expressed with zero are usually subjects (*Ibid*.). More recent studies have also shown that the use of ellipsis is highly usual in Estonian conversational data (see Vihman 2015). Therefore, Estonian may be described as a language including both full and reduced pronominal forms[2] (Siewierska 2003, 2004).

Traditionally, the choice of referential expressions is explained in terms of salience or the accessibility of the referent. According to different accessibility hierarchies (e.g. Gundel et al. 1993, Ariel 1990, Givón 1983, see also Kaiser and Trueswell 2008), more reduced forms, i.e. those with the least phonetic content, refer to most salient entities, whereas fully specified longer forms refer to less salient entities. Following this idea, two hypotheses may be drawn: firstly, Estonian demonstrative pronouns refer to less salient entities compared to short 3rd person pronouns[3], and secondly, the referent of the zero pronominal should be more accessible than *ta*.

Also taking into consideration the animacy hierarchy as described by Siewierska (2004: 149), stating that human entities are more salient than other animate entities and animate entities outperform abstract ones, the salience/accessibility view is compatible with the differences in Estonian personal pronouns and demonstratives. The personal pronoun *ta*/*tema* is mostly restricted to human or animate referents recently mentioned in the discourse. The demonstratives *see* and *too* typically express inanimate referents, although in some contexts they may indicate animate entities as well. When there are two similar animate referents in the discourse, *ta* is used to refer to the first and a demonstrative to the second entity mentioned. (EKG II: 208–209, see also Pajusalu 2006) More recent studies have shown that *ta* refers to the most activated referents in discourse and demonstratives tend to refer to entities not yet in the focus of attention of the addressee (Pajusalu 2006, 2009).

---

2   Siewierska (2003, 2004) has used the term *reduced pronominal* to denote pronominal forms with no overt form, such as zero pronouns, as opposed to full pronominal forms, such as personal pronouns. This term should not be confused with the notion of *reduced referential device*, which includes all possible pronominal forms as opposed to lexically full referential devices (e.g. Kibrik 2011).

3   This conclusion takes into account only the short form of the 3rd person pronoun. The long form *tema* is actually longer than demonstrative pronouns. Note that the usage of long forms is largely guided by pragmatic factors, such as contrast (Pajusalu 2009, Kaiser and Hiietam 2003). Therefore, it is not considered the common neutral choice when referring to highly salient entities. Although *tema* refers to just as salient entities as the short form *ta*, its special function must be taken into consideration and the position of Estonian long pronoun forms on the salience scale needs further investigation.

Additionally, the personal pronoun *ta* is sensitive mostly to the grammatical role of the referent, preferring subjects; the use of the demonstrative *see* is more complex, expressing preferences in grammatical role (object) as well as word order (postverbal arguments) (Kaiser and Vihman 2010). The choice between a short and a long pronoun form is considered to be pragmatic and factors such as contrast come into play when choosing them (Pajusalu 2009, Kaiser and Hiietam 2003).

While the distinction between Estonian personal and demonstrative pronouns is relatively well studied, the difference between the referential properties of the zero person marker and the personal pronoun *ta/tema* is not so clear. The widely approved Givenness Hierarchy (Gundel et al. 1993) states that (unstressed) personal pronouns and zero forms may only refer to highly salient entities and both express the highest status *in focus*. This leaves salience quite incapable of differentiating overt and zero pronominal forms. The equal level of salience of these forms is supported by the circumstance that zero and *ta* are interchangeable in many contexts. Two excerpts from narratives (examples 1 and 2) suggest that there is no difference in the cognitive salience of *ta* and zero: both of them refer to the most salient entity in the ongoing discourse.

(1)   1   ja= ja    siss em (.) see: se= talumees ee korja-s      pirn-e.
          and and  then      this this farmer     pick-3SG.PST pear-PL.PART

      2   mm: ja   ø=        roni-s        sinna 'redeli    otsa
              and 3SG.ZERO climb-3SG.PST there  ladder.GEN up there

      3   jaa= ja   siis  ø        korja-s       ästi  hoolega= neid
          and  and then 3SG.ZERO pick-3SG.PST very  carefully them.PL.PART

      'and and then *umm* this this farmer *ee* was picking pears; *mm* and he climbed up there onto a ladder; and then was picking them very carefully'

(2)   1   ja= sis   natukese   aja       pärast ää lähe-b      mööda
          and then a little.GEN time.GEN after      go-3SG.PRS by

          mingi: väike põnn (.) ästi  suure     jalgratta-ga.
          some   little kid      very  big.GEN  bike-COM

2   jaa: m (.)  sis=  ta          alguse-s      mõtle-b
    and          then  3SG.SHORT  beginning-IN  think-3SG.PRS

3   et=  ta          sõida-b        nisama  mööda
    that  3SG.SHORT  ride-3SG.PRS  just      by

'and then in a little time *ää* some little kid with a very big bike goes by; and *m* at the beginning he thinks; that he will just ride by.'

As the zero person marker and the pronoun *ta* express an equal level of salience, some other factors must be crucial in differentiating them. Recent cross-linguistic studies (see e.g. Kaiser and Trueswell 2008, Kaiser 2009, Brown-Schmidt et al. 2005, Byron et al. 2008) have also shown that although salience has a clear effect on the choice of referring expressions, various form-specific factors should be taken into consideration in explaining the use of distinct referential forms. A multiple-constraint approach may outperform the uniform dimension (e.g. salience) perspective. Adapting this idea and the Centering Theory framework in this article, it will be examined if zero and *ta*, although equal in salience, have distinct effects on discourse coherence.

## 3. Centering theory: an overview

A brief introduction concerning the key terms of Centering Theory is given in this section. The overview is based on two central detailed works on Centering: Grosz, Joshi, and Weinstein (1995) and Walker, Joshi, and Prince (1998).

The Centering framework explains discourse coherence within a discourse segment by bringing together attentional state factors and the choice of referring form (Grosz et al. 1995: 204, Walker et al. 1998). The theory is concerned with the distribution and interpretation of referring expressions in natural language (Gundel 1998: 190). The notion *inference load* is used to describe changes in the hearer's attentional focus during the task of identifying and tracking entities through the text. In a coherent text, this task of tracking entities in a given discourse should be made as easy as possible to reduce the inference load placed upon a hearer. (Grosz et al. 1995, Walker et al. 1998) The choice of linguistic expression directly influences the inference load and hence one of the basic claims of Centering states that the perceived discourse coherence

depends on the distinct inferential demands positioned on a hearer/ reader by the usage of different referring expressions and syntactic forms (Grosz et al. 1995: 206–207). Thus the Centering model imitates psychological processes involved in tracking entities through the discourse (Ballantyne 2004: 50). The theory is mainly interested in the interpretation of anaphoric expressions (e.g. pronouns, zero anaphora, definite descriptions) and is focused on noun phrases (NPs). (Grosz et al. 1995, Walker et al. 1998)

Centering as a model of analyzing discourse defines it as a coherent sequence of utterances. Units of discourse are utterances, i.e. "the uttering of a sequence of words at a certain point in a discourse", not sentences (Grosz et al. 1995: 208). In each discourse segment, local and global coherence are exhibited. Local coherence is coherence among utterances in a particular segment, while global coherence pertains to coherence with other segments in ongoing discourse (Grosz et al. 1995: 204). Although Centering was originally concerned with the local coherence of discourse segments (Grosz et al. 1995), there have been attempts to apply the centering algorithm to global coherence as well (Hedberg and Dueck 1999, Cristea et al. 1998, Di Eugenio 1998). Walker (1998) even suggests that the within-segment approach of Centering should be replaced with a version of Centering that accounts for global discourse structure. The present work also deviates from the original notion of Centering in that segment boundaries are not explicitly taken into account and no global/local discourse distinction is made.

### 3.1. Centers

The basic term of the theory is center. **Centers** are semantic entities in an utterance, part of a discourse model concerning the utterance in a segment, that link the utterance to other utterances in a particular segment (Grosz et al. 1995: 208, Walker et al. 1998: 3). As Grosz and her colleagues (1995: 208) point out, centers are discourse constructs and semantic objects, not words, phrases or syntactic forms. Centers are considered to regulate the local information flow in discourse (Hu and Pan 2001: 139).

Three kinds of centers are distinguished in the centering framework: forward-looking, backward-looking and preferred centers (Grosz et al. 1995, Walker et al. 1998, see also Ballantyne 2004, Taboada and Hadic Zabala 2008).

1) All discourse entities mentioned in an utterance $U_n$ in a discourse segment make up a set of **forward-looking centers** (Cf), which are ranked by means of salience.
2) The first, i.e. the most salient member of the Cf set in utterance $U_n$ is called the **preferred center** (Cp). The Cp predicts what the backward-looking center Cb of the following utterance will presumably be.
3) A special member of the Cf set, the **backward-looking center** (Cb), is the entity which is in the center of attention in the utterance $U_n$. Cb is the highest-ranking entity from the set of Cf of the previous utterance $U_{n-1}$, which is realized also in the current utterance. Therefore, Cb serves as the link between the present and previous utterance in a particular discourse segment.

When applying Centering to a new language, the Cf ordering principles must be specified. Researchers have applied different Cf ranking criteria to different languages. The most traditional criterion has been the grammatical function (subject > object(s) > other), where subjects are classed higher than objects and objects higher than other functions (Walker et al. 1998: 7). Several other criteria have been used in different works related to Centering as well, such as topicality (Walker et al. 1994), empathy (Di Eugenio 1998), the Givenness Hierarchy (Ballantyne 2004) and animacy (see Taboada 2008). The description of ordering principles followed in this article can be found in section 4, Material and Method.

### 3.2. Transitions

By comparing the centers of utterances $U_n$ and $U_{n-1}$, the transition types are computed within the Centering framework. Two factors are important in the typology of transitions: the comparison of Cb of $U_n$ and $U_{n-1}$, and the comparison of the Cb of $U_n$ with the Cp of $U_n$. The transition types are shown in Table 1, where $Cb(U_{n-1}) = [?]$ stands for cases where $Cb(U_{n-1})$ does not exist. (Walker et al. 1998: 5–6)

**Table 1.** Transition types in CT (Walker et al. 1998: 6)

| | $Cb(U_n) = Cb(U_{n-1})$ or $Cb(U_{n-1}) = [?]$ | $Cb(U_n) \neq Cb(U_{n-1})$ |
|---|---|---|
| $Cb(U_n) = Cp(U_n)$ | CONTINUE | SMOOTH SHIFT |
| $Cb(U_n) \neq Cp(U_n)$ | RETAIN | ROUGH SHIFT |

$U_n$ and $U_{n-1}$ are linked by the CONTINUE transition when the Cb and the Cp of the current utterance are the same and the Cb at hand simultaneously coincides with the Cb of the previous utterance. This means that the speaker is talking about a certain entity and aims to keep it in the center of attention in the next utterance(s) as well (Taboada and Hadic Zabala 2008: 70–71, Walker et al. 1998: 5) A RETAIN transition relates two utterances in cases where the Cb of utterances $U_n$ and $U_{n-1}$ are the same, but the Cb and the Cp of the present utterance do not coincide. If the Cb of the current utterance is not concurrent with the Cb of the previous utterance, then there is a SHIFT relation between two utterances: a SMOOTH SHIFT in cases where the Cb is the same as the Cp of $U_n$, and a ROUGH SHIFT in cases where the Cb is not the same as the Cp in $U_n$. Transitions help to explain how coherence is achieved: a text is more coherent if it maintains the same centers (Taboada and Hadic Zabala 2008: 71).

Example (3), an excerpt from a coherent discourse segment, helps to understand these basic centering notions. In the first sentence, there are two centers: *boy* (realized with a zero form) and *basket*. They are listed in a set of Cf according to the grammatical function. The first member of the Cf list is *boy*, so it is the preferred center Cp. As this is not the first utterance of the segment, it has a backward-looking center Cb, which is *boy*. The second utterance has three centers: *boy*, *bike* and *pear*. The highest of them is *boy*, which is also the Cb of this utterance and the Cp for the following utterance. Therefore, the transition between the first and second utterance is CONTINUE. In the last utterance, three Cf-s are realized: *girl*, *boy* and *road*. However, the first member of the Cf list is no longer *boy*, instead it is *girl*, a grammatical subject of this utterance. *girl* is also the Cp in this utterance. Still, the Cb is *boy*, thus the transition is RETAIN.

(3)  1  ja=  sis  ø  võttis  terve  selle  korvi
and  then  3SG.ZERO  take.3SG.PST  whole  this.GEN  basket.GEN

'and then he took the whole basket'
Cf: boy, basket
Cb: boy – Cp: boy
transition: CONTINUE

2 ja:   ø         pani        punuma        koos= oma  ratta-ga
    and  3SG.ZERO  put.3SG.PST  pleach.mINF  with  his  bike-COM

    selle      pirni-ga.
    this.GEN   pear-COM

    'and (he) turned tail with his bike with the pear'
    Cf: boy, bike, pear
    Cb: boy – Cp: boy
    transition: CONTINUE

3  ää  ja   siis  ee  tee   peal tuli           ta-lle
       and  then       road  on   come.3SG.PST  3SG.SHORT-ALL

    vastu     mingi:  tüdruk
    against   some    girl

    '*ää* and then *ee* on the road some girl came towards him'
    Cf: girl, boy, road
    Cb: boy – Cp: girl
    transition: RETAIN

### 3.3. Rules and constraints

Besides transitions, the Centering framework involves rules and constraints.

1) Rule 1, also known as a Pronoun Rule, states that if any element of the set of Cf of the previous utterance $U_{n-1}$ is realized as a pronoun in the current utterance $U_n$, then the Cb of the current utterance $U_n$ must also be a pronoun. This indicates that when the speaker continues to talk about the same entity, it is a salient one, which is (usually) referred to by pronouns. Therefore, according to rule 1, the least marked referential form in the utterance refers to the most salient entity. (Walker et al. 1998: 4–5, Grosz et al. 1995)
2) Rule 2 ranks the transition types within Centering, stating that CONTINUE is preferred to the RETAIN transition, which is preferred to the SMOOTH-SHIFT transition, which is preferred to the ROUGH-SHIFT transition (Walker et al. 1998: 4). The ordering of the transition types is based on the processing time they require from the addressee and thus makes it possible to measure the coherence of a

discourse segment, as the most coherent segments require less time for processing (Walker et al. 1998: 6). It implies that in a coherent discourse, CONTINUE transitions should predominate, while ROUGH-SHIFT should occur rarely.

Constraints, as they are formulated by Walker and colleagues (1998: 3), state that:
1) each utterance in a discourse segment has precisely one Cb;
2) every element listed in the Cf set must be realized in that utterance;
3) the Cb of the current utterance is the highest-ranked element in the set of Cf of the previous sentence that is realized in the current utterance.

The term *realized element* includes pronouns, zero pronominals, explicitly realized discourse entities and implicitly realized entities inferable from the discourse situation (see Walker et al. 1998: 4). Constraint 1 is also regarded as the central claim of the theory (Poesio et al. 2004). A classic example explaining Rule 1, the importance of Cf ranking and Cb realization by a pronoun comes from Gordon et al. (1993: 313). Passage (4) is given two possible final utterances, 3 and 3'.

(4)   1  Susan gave Betsy a pet hamster.
          Cf: Susan, Betsy, hamster

      2  She reminded her such hamsters were quite shy.
          Cf: Susan, Betsy, hamsters
          Cb: Susan

      3  She asked Betsy whether she liked the gift.
      3' Susan asked her whether she liked the gift.
          Cf: Susan, Betsy, gift = hamster
          Cb: Susan;

Here, utterance 1 has no Cb as it is the first utterance of the segment. Its Cf list is ranked according to grammatical role, yielding Susan as the first element. In utterance 2, Susan is the Cb and again the highest-ranking Cf. In utterance 3, the Cb (*Susan*) is preserved and realized by a pronoun, and it is a natural-sounding coherent continuation of the previous segment. However, in 3' the Cb (*Susan*) is realized by a proper name and a pronoun is used to refer to the second entity in the previous

Cf list (*Betsy*). This results in a very unnatural result and the utterance 3' reads poorly.


## 4. Material and method

In this study, a Centering analysis is applied to experimentally elicited spoken narratives based on the Pear Film[4] (Chafe 1980). The design of the research follows the one worked out by Wallace Chafe and his coworkers. 16 female high-school or university students, aged 18–26, all native speakers of Estonian voluntarily participated in the study. To hedge against possible gender-based differences in pronoun use in this study, only female subjects are included. The subjects are shown the Pear Film: a short motion picture without explicit language about a man who is picking pears and a boy who steals some pears from him. After seeing the footage, the subject is asked to retell the content of the film to a friend who has not seen it. All hearers are approximately the same age as the speaker. Note that in the original experiment, the hearer was not acquainted with the subject. However, I find it best to have the story told to friends since it makes one feel more comfortable and therefore provides more natural narratives. Having just seen the movie, all referents must be known for the speaker. Yet for the hearer every single one of them is new at the beginning of the narrative. Therefore, the experiment provides us with information about how the speaker estimates the cognitive states of the hearer.

The coding procedure largely followed the manual by Hadic Zabala and Taboada (2006, see also Taboada 2008), as it takes into consideration the characteristics of spoken language. In this study, an utterance is defined as an intonation unit (Taboada 2008: 181). An utterance is generally a tensed clause (Hadic Zabala and Taboada 2006). If the unit does not fill the requirements of a clause, it might be considered an utterance if it contains entities (*Ibid.*). Only utterances containing the 3rd person overt pronoun or the zero person marker are coded and analyzed for this study. In total, 517 cases of pronoun use (overt or zero form) were included in the study.

Besides the proper centering analysis, grammatical factors pertaining to the use of pronoun forms under study, such as (1) case, (2) grammatical role (subject, other), and (3) clause type (main or subordi-

---

4   The film can be viewed at http://pearstories.org/.

nate clause) are coded in order to gain a fuller picture of the differences between these forms in Estonian spoken language.

In the analysis procedure, every narrative was first segmented into utterances and for each utterance, the list of forward-looking centers Cf, the backward-looking center Cb and the preferred center Cp were coded. As deciding on the Cb and the Cp of an utterance directly depends on the ordering of the Cf list, it is crucial in adapting the Centering Theory to a new language to determine what ordering best suits that language. The most appropriate ordering for a language would result in the least violation of the Pronoun Rule. The most often used ranking principle follows grammatical relations, according to which subjects are ranked higher than objects and objects are ranked higher than adverbials (Walker et al. 1998). However, as the ordering of entities realized in a sentence is considered to be language-dependent, different Cf templates occur (see Taboada 2008). In this article the original grammatical function ordering principle has been modified to fit Estonian data.

In Estonian, the grammatical subject does not always indicate the most agentive argument in the clause (EKG II: 14). Besides prototypical subjects, many subject-like arguments exhibit some subject properties (see Metslang 2013). According to Erelt and Metslang (2006), clauses can be divided into unmarked and marked basic clauses, based on the realization of the main topic of the clause. Marked basic clauses include existential clauses, possessive clauses, source-marking resultative clauses and experiential clauses. While in unmarked basic clauses, the main topic is realized as a grammatical subject, in unmarked ones it appears as an adverbial, an oblique or a direct object. (Erelt and Metslang 2006) However, in all marked basic clauses, the topic has some subject-like features (e.g. the clause-initial position of the experiencer or possessor) and it is often an animate entity (*Ibid.*). Examples of marked basic clauses can be found in (5) (a possessive clause) and (6) (an experiential clause).

(5)  Mari-l     on       raamat.
     Mary-AD  be.3SG   book
     'Mary has a book.'


(6)  Mari-le    meeldi-b   muusika.
     Mary-ALL  like-3SG   music
     'Mary likes music.'

As the ordering criteria are based on the salience of the referent, it is clear that the topic must rank higher in the hierarchy than grammatical subject. This is also supported by the fact that if utterance (6) were followed by another utterance in a segment, it is most probably Mari who would be mentioned by a pronoun, not music.

In this article, the notion of oblique subject is applied to account for the non-nominative animate subject-like topic of marked basic clauses. It is clear that sticking strictly to grammatical function in ordering the Cf list in Estonian would result in more frequent violation of the Pronoun Rule. The Cf ordering criteria followed in this article are as follows (7):

(7)  oblique subject > grammatical subject > grammatical object2 > grammatical object1 > other[5]

In ranking the Cf list of an utterance, it must also be decided what to do with NPs containing more than one entity (e.g. possessives, such as *minu vend* 'my brother', and conjoined NPs, such as *Jaan ja Mari* 'John and Mary'). According to Di Eugenio (1998), if the possessee is inanimate, the possessor is ranked before the possessee; and if the possessee is animate, the possessor is ranked after the possessee. For conjoined NPs, the individual entities are less salient than the whole group denoted by the conjoined NP. Therefore, the group as a whole must rank first and individual entities are subsequent, following the linear order.

The decision as to whether there is a zero in an utterance is based on the valency of the verb. For example, in the case of intransitive verbs which take one argument, the utterance must contain at least one verb and a subject. If there is no overt NP expressing the subject, then it must be a zero form. In case of transitive verbs with two arguments, the utterance must have at least a subject and an object along with the verb; when one of them is not expressed with an explicit NP, it is considered to be a zero. Only cases where the obligatory argument is missing are considered to be instances of a zero form.

For the analysis, all zero forms appearing in the data are considered to be informative regarding their cognitive status, therefore every single zero form is included. The question of coordination and its influence on the usage of zero forms is considered to be a syntactic issue and is therefore neglected here.

5  When several NPs apper in function other than oblique subject, grammatical subject or object, linear order is followed in ordering.

Cases where the interpretation of the zero person marker is ambiguous are excluded from the analysis. Example (8) illustrates such a situation. Here, the zero form in the second utterance may either refer to the boy with a bike or to one of the boys who took the hat to the boy with a bike. Also omitted are speakers' comments not related to the narrative discourse.

(8)  1  ja   siis ee (.)üks  poiste-st      vii-s          selle (.)  ta-lle?
        and  then     one  boy.PL.GEN-EL bring-3SG.PST this.GEN  3SG.SHORT-ALL

        'and then *ee* one of the boys took this to him'

     2  ja   siis õ (.) õ  ø            võttis=      sea-lt      kolm  pirni.
        and  then     3SG.ZERO take.3SG.PST there-ABL three  pear.PL.PART

        'and then *õ õ* (he) took three pears from there'

## 5. Results

The quantitative data for pronominal forms under examination is presented in this section. Note that due to the infrequent occurrence of plural forms, chi-square tests reported in this study are only applied to compare the 3rd singular overt form *ta* and the 3rd singular zero form[6]. There were 417 instances of SG pronominal forms altogether. Even so, for the sake of the complete overview of the data, the contingency tables present details for the plural forms also (100 instances altogether).

The initial assumption that the zero person marker signals CONTINUE and the overt pronoun is used more in cases of other transitions is not fully supported. It is true that there appears to be a tendency of zero only being used with CONTINUE, while overt forms may express transitions other than CONTINUE as well. However, for both the overt form as well as zero, CONTINUE is the most exploited transition. Therefore other differentiating aspects besides Centering transitions are also examined.

The overall distribution of pronominal forms with respect to the Centering transition type is shown in Table 2. What the results clearly indicate is that the ordering rule (Rule 2) applies to analyzed data and therefore the narratives used for this study are examples of coherent

---

6   As the observations from the corpus used here are not totally independent, the chi-square test results are used as a suggestive evidence, not as a strong statistic proof.

text. The most frequent transition is CONTINUE (424 utterances), as the theory predicts. At the other end are SHIFTS (35), which are rather rare in the data. In between are RETAINS (38). In the data there were also 20 utterances without Cb and thus no transition type was applicable to them. A small number of shifts is of course anticipated since the theory itself predicts that full NPs are more likely to signal them. The cases where no Cb can be determined might be considered as the beginnings of new discourse segments, although this assumption needs further investigation. Note that in Table 2, the SMOOTH- and ROUGH SHIFT are not distinguished. This is due to the fact that it is very uncommon for utterances containing pronouns to express ROUGH SHIFTS (see also Di Eugenio 1998). Even Grosz and colleagues (1995) in their original paper on Centering see center shifting as a uniform category.

It is important to note that the zero pronominal is used even more frequently than the overt form. In singular forms, it can be observed that the frequency of the 3rd person pronoun *ta* and the zero form is roughly equal in the data (204 and 213, respectively). The occurrence of plural pronouns is modest (*nad* is used in 34 cases and plural zero in 66 cases), but a significant fact is that the plural zero is used more often than the plural overt form. Note also that while for the CONTINUE transition all observed pronoun options are possible, other transitions are more constrained regarding the alternative forms. Form-wise, it appears to be significant that the overt pronoun *ta* exhibits more variation in transition types compared to zero. The zero form, on the other hand, is mainly restricted to the CONTINUE transition ($X^2(3, N = 417) = 44,686$, $p < .01$).

**Table 2.** Distribution of transitions by pronoun form

| FORM | CONTINUE | RETAIN | SHIFT | No Cb | TOTAL |
|---|---|---|---|---|---|
| *ta* | 140 | 30 | 17 | 17 | 204 |
| zero (SG) | 197 | 3 | 11 | 2 | 213 |
| SG forms total | 337 | 33 | 28 | 19 | 417 |
| *nad* | 25 | 5 | 3 | 1 | 34 |
| zero (PL) | 62 | – | 4 | – | 66 |
| PL forms total | 87 | 5 | 7 | 1 | 100 |
| TOTAL | 424 | 38 | 35 | 20 | 517 |

In (9) a typical example of using the zero person marker in a narrative is given. All the transitions are CONTINUES, there is only one animate entity referred to in these utterances and the zero form is always in the nominative case and the grammatical subject of the clause.

(9)  1  emm  seal= oli         üks mees kes  korja-s        pirne,
        umm  there be.3SG.PST one man who pick-3SG.PST pear.PL.PART

        pirni-puu  otsast,      redeli-ga.
        pear-tree  from upon ladder-COM

        'umm there was a man who was picking pears from the pear tree, with a ladder'
        Cf: man, pears, tree, ladder
        No Cb – Cp: man
        transition: NONE

     2  jaa:  ø        korja-s        pirne          rahulikult.
        and  3SG.ZERO pick-3SG.PST pear.PL.PART calmly

        'and (he) was picking pears calmly'
        Cf : man, pears
        Cb: man – Cp: man
        transition: CONTINUE

     3  jaa  ø        pani=      nad        siis= tõi
        and  3SG.ZERO put.3SG.PST 3PL.SHORT then bring.3SG.PST

        pirni-d  ilusti  alla    oma korvi,
        pear-PL neatly down his  basket.ILL

        'and (he) put them then brought the pears down neatly to his basket'
        Cf: man, pears, basket
        Cb: man – Cp: man
        transition: CONTINUE

     4  ø        jättis         nad        sinna (.)
        3SG.ZERO leave.3SG.PST 3PL.SHORT there

        '(he) left them there'
        Cf: man, pears
        Cb: man – Cp: man
        transition: CONTINUE

A similar excerpt with CONTINUE transitions is given in (10), but here the overt pronoun *ta* appears in contexts where zero was used in the previous example.

(10) 1   sis   ää   ø       jäi           seisma     sinna: (.) puu  alla.
         then  ää   3SG.ZERO stay.3SG.PST stand.mINF there       tree under

         'then *umm* (he) stopped under that tree'
         Cf: boy, tree
         Cb: boy – Cp: boy
         transition: CONTINUE

     2   s=   **ta**         vist       taht-is       alguse-s      võtta
         then 3SG.SHORT  probably   want-3SG.PST  beginning-IN  take.dINF

         ühte          pirni=       aga
         one.PART    pear.PART    but

         'then he probably at first wanted to take one pear but'
         Cf: boy, pear, beginning
         Cb: boy – Cp: boy
         transition: CONTINUE

     3   s=   **ta**        ütle-s=      et     no    kui  juba   sis    juba.
         then 3SG.SHORT  say-3SG.PST  that   well  if   already then  already

         'then he said that if already so, so be it'
         Cf: boy
         Cb: boy – Cp: boy
         transition: CONTINUE

     4   ja   sis   ta          võttis       kogu   korvi.
         and  then  3SG.SHORT  take.3SG.PST  whole  basket.GEN

         'and then he took the whole basket'
         Cf: boy, basket
         Cb: boy – Cp: boy
         transition: CONTINUE

Yet another display of a sequence where the overt pronominal form is used is presented in (11). Here other transitions besides CONTINUE are also present, hence the occurrence of zero would be unlikely in contexts similar to this one.

(11)  1   sis=  **ø**           sõitis        seal  mingi  kruusatee    peal.
          then   3SG.ZERO   ride.3SG.PST  there  some   gravel road   on

      'then rode there on some gravel road'
      Cf : boy, gravel road
      Cb: boy – Cp: boy
      transition: CONTINUE

      2   sis=  **ta**-lle          tuli          vastu  üks  tüdruk, ka  ratta-ga.
          then   3SG.SHORT-ALL  come.3SG.PST  against one  girl     also bike-COM

      'then a girl came towards him, also with a bike'
      Cf: girl, boy, bike
      Cb: boy – Cp: girl
      transition: RETAIN

      3   ja    siis  ää   se=  poiss  jäi            'vaatama  kuidagi
          and   then      this  boy    stay.3SG.PST  look.mINF  somehow

      'and then *ää* the boy kept looking somehow'
      Cf: boy
      Cb: boy – Cp: boy
      transition: CONTINUE

      4   ja=  sis  ee  se=  tüdruk  ee  lükka-s=   **ta**-lt            mütsi
          and  then     this  girl        push-3SG.PST  3SG.SHORT-ABL  hat.GEN

      pea-st    ära. (.)
      head-EL   off

      'and then *ee* the girl *ee* pushed his hat off his head'
      Cf: girl, boy, hat
      Cb: boy – Cp: girl
      transition: RETAIN

   Although some difference in preferred transition types of overt forms and zeros can be observed, as in examples (9) and (11), there is still a coinciding tendency for both to code mainly CONTINUES, as examples (9) and (10) illustrate. This indicates that discourse coherence factors alone cannot completely explain why one form is chosen instead of other. Hence other possible defining factors must be considered as

well. The important aspects that need closer examination, as the examples also indicate, are grammatical role and case, as well as clause type.

A prototypical grammatical subject is expressed with the unmarked nominative case in Estonian (Metslang 2013: 224), while other grammatical roles tend to exhibit other cases. Studying the behaviour of subjects and subject-like arguments in Estonian written texts, Helena Metslang (2013: 243) has found that zero-anaphora tends to express more prototypical subjects and therefore helps to distinguish subjects from objects. As it is often difficult to see the case marking in zero forms, it is tempting to assume that referents in grammatical roles other than subject cannot be expressed with zero. However, Lindström and her colleagues (2014) have observed that it is very common in Estonian to omit the experiencer marked with an oblique case (adessive or allative) in certain experiential constructions. The distribution of pronominal forms by grammatical role presented in Table 3 also shows that although the non-nominative arguments occur rarely in the data, they are still possible in Estonian.

**Table 3.** Pronoun forms by grammatical role

| FORM | SUBJECT | OTHER | TOTAL |
|---|---|---|---|
| *ta* | 109 | 95 | 204 |
| zero (SG) | 207 | 6 | 213 |
| SG forms total | 316 | 101 | 417 |
| *nad* | 30 | 4 | 34 |
| zero (PL) | 66 | – | 66 |
| PL forms total | 96 | 4 | 100 |
| TOTAL | 412 | 105 | 517 |

As expected, zero pronominals almost exclusively (207 out of 213 in SG; 66 out of 66 in PL) express grammatical subjects. By contrast, overt forms are not restricted to one grammatical relation: the overt pronoun *ta* is used 95 times in a grammatical role other than subject. The difference between *ta* and SG zero appears to be significant ($X^2(1, N = 417) = 108.675$, $p < .01$).

This data is strongly related to the results shown in Table 4, which presents the distribution of Estonian cases used with the pronoun forms

under examination[7]. The nominative, genitive and partitive are grammatical cases, while all others are semantic cases expressing oblique arguments. Subjects most often take the nominative case, but partitive is also possible in some constructions. Singular objects are expressed with the partitive or genitive, plural objects with the partitive or nominative. It can be seen that pronominal forms are mainly used in the nominative case, which is not surprising considering the fact that the nominative is the typical case for marking the grammatical subject. Note that some instances of exterior locative cases (allative and adessive) belong to marked basic clauses, where the experiencer of the clause is not the grammatical subject of the clause. In these clauses, the oblique subject expressed with a locative case is nevertheless considered to bear the most important and active role in a sentence. Note also that the allative and adessive cases can be expressed with the zero form. This result is especially interesting as it poses a contradiction to the cross-linguistic observation that there are probably no languages with zero obliques (Siewierska 2003). Other studies have also shown that zero obliques are quite usual in Estonian experiential clauses (Lindström et al. 2014).

**Table 4.** Pronoun form by case

| FORM | NOM | GEN | PART | EL | ALL | AD | ABL | COM | TOTAL |
|---|---|---|---|---|---|---|---|---|---|
| *ta* | 114 | 5 | 11 | 1 | 29 | 40 | 3 | 1 | 204 |
| zero (SG) | 207 | – | 1 | 1 | 2 | 2 | – | – | 213 |
| SG forms total | 321 | 5 | 12 | 2 | 31 | 42 | 3 | 1 | 417 |
| *nad* | 31 | – | – | – | 3 | – | – | – | 34 |
| zero (PL) | 66 | – | – | – | – | – | – | – | 66 |
| PL forms total | 97 | – | – | – | 3 | – | – | – | 100 |
| TOTAL | 418 | 5 | 12 | 2 | 34 | 42 | 3 | 1 | 517 |

Example (12) is a typical instance of a zero person marker, as it expresses the grammatical subject in the nominative case. Similar conditions apply for the pronoun *ta* in example (13), but examples (14) and (15) show *ta* in contexts where zero is most unlikely to appear.

---

7   Note that in the data, not all possible Estonian cases (14) are present with pronominal forms.

(12)  ja=  sis   **ø**        kuidagi  läks       selle      ratta-ga  minema
      and   then  3SG.ZERO  somehow  go.3SG.PST  this.GEN  bike-COM  away

      'and then (he) somehow went away with this bike'

(13)  ja=  sis=  **ta**         tühjenda-s    selle:    nagu  need
      and   then  3SG.SHORT  empty-3SG.PST  this.GEN  like  this.PL

      põlle-tasku     ära
      apron-pocket  away

      'and then he emptied it, like these in an apron pocket'

(14)  ee ja=  sis   sõida-b    mingisugune     ee: õ  'kooli-tüdruk
      ee and  then  ride-3SG  some-kind-of          school-girl

      **ta-lle**           vastu.
      3SG.SHORT-ALL   against

      '*ee* and then some kind of *ee õ* schoolgirl rides towards him'

(15)  jaa (...)  mm  **ta-l**=          oli              juba      umbes  kaks
      and              3SG.SHORT-AD  have.3SG.PST  already  about  two

      korvi         korja-tud
      basket.PART  pick-PASS.PPTCP

      'and (...) umm he had picked about two baskets already'

The preference to use the zero form for mostly nominative subjects is also supported by the fact that a wide majority of zero forms expresses the backward-looking center Cb in an utterance. As can be seen in Table 5, besides marking the Cb in an utterance (146 utterances), the overt form *ta* is related to instances where the pronoun is not the Cb of the current utterance (40 utterances) or there is no Cb in the utterance (18 utterances). However, except for some marginal cases, the zero form tends to express only Cb-s (208 utterances) ($X^2$(2, N = 417) = 55,327, p < .01). According to the Centering Theory, the Cb is the most central entity in an utterance and each utterance can contain only one Cb. However, in actual language use, often more than one pronoun may appear in one utterance. In these cases, the zero form is reserved for the Cb and the overt form refers to the less central entity, a non-Cb.

**Table 5.** Pronoun form by Cb

| Form | pro = Cb | pro ≠ Cb | NA | Total |
|---|---|---|---|---|
| *ta* | 146 | 40 | 18 | 204 |
| zero (SG) | 208 | 3 | 2 | 213 |
| SG forms total | 354 | 43 | 20 | 417 |
| *nad* | 31 | 2 | 1 | 34 |
| zero (PL) | 64 | 2 | – | 66 |
| PL forms total | 95 | 4 | 1 | 100 |
| Total | 449 | 47 | 21 | 517 |

Example (16) illustrates cases like these. Here, the zero form is reserved for the Cb of the utterance (*three boys*), whereas the less central entity, i.e. an entity other than the subject (*the boy with a bike*), is referred to by using an overt pronoun.

(16) ja    siis    veel          **ø**          and-si-d        **ta-lle**
     and   then    furthermore   3PL.ZERO    give-PST-3PL    3SG.SHORT-ALL

     mütsi     tagasi,
     hat.GEN    back

     'and then (they) also gave him the hat back'

Also, in utterances where there is no Cb, a pronoun can sometimes be used. As example (17) suggests, in such cases the general discourse topic is realized as a pronoun. Here, utterance 4 has no Cb, as in utterance 3 neither *man* nor *pears* are mentioned. However, as utterance 3 does not realize any animate entities, it is clear that *ta* in utterance 4 refers back to the man, who is the general discourse topic. The same applies to utterance 6. Although here the previous utterance 5 introduces a new animate entity into discourse, it is clear from the content of the film that the pronoun *ta* in utterance 6 cannot refer to the man with a goat. It has been assumed that the absence of a backward-looking center may indicate a segment boundary (Passonneau 1998, see also Walker 1998), and the use of zero in these cases is unlikely. However, Walker (1998) has claimed that pronouns actually can be used in the initial utterance of the next segment. In Estonian, an overt pronominal is used in such cases.

(17) 1   õ   seal= oli (.)        mees: farmer= või keegi.        seline (.)
         there be.3SG.PST man     farmer   or   somebody  such

     mehiko   tüüpi.
     mexican  type

     'õ there was a man, a farmer or somebody, like Mexican type'
     Cf: man, film
     Cb: film – Cp: man

2   jaa:  **ta**          korja:-s (.)    puu-l (.)   ta=
    and   3SG.SHORT   pick-3SG.PST tree-AD     3SG.SHORT

    korja-s         pirn-e?
    pick-3SG.PST   pear-PL.PART

    'and he was picking up on a tree, picking pears'
    Cf: man, pears, tree
    Cb: man – Cp: man
    transition: SMOOTH SHIFT

3   ää  jaa: (.)  ja   se= oli=        igav   töö.
        and      and  it   be.3sg.pst  boring work

    'ää and and it was boring work'

4   ja=  siis  **ta** (.)    ää  korja-s          pirn-e
    and  then  3SG.SHORT     pick-3SG.PST   pear-PL.PART

    'and then he ää was picking pears'
    Cf: man, pears
    Cb: ø – Cp: man

5   ja=  si  vahepeal= sea-lt      jaluta-si-d     mingisugune   tüüp
    and  then meanwhile there-ABL walk-PST-3PL  some-kind-of   guy

    jaluta-s         vel      kitse-ga  mööda?
    walk-PST.3SG   as well  goat-COM  by

    'and then meanwhile walked some kind of guy, walked by with a
    goat as well
    Cf: man, goat
    Cb: ø – Cp: man

6   aga: (.)  ja=   sis    no=   **ta**=        oli        juba    'kolm
    but        and   then  well   3SG.SHORT  be.3SG.PST  already  three

    selist        suur-t      ä  korvi=          oli          pirne.
    such.PART  big-PART       basket.PART  be.3SG.PST  pear.PL.PART

    'but and then well he had already three such big *ä* baskets of pears
    he had'

    Cf: man, baskets, pears
    Cb: ø – Cp: man

Differences in pronoun use between clause types (main and subordinate) were also examined. Table 6 presents the distribution of pronoun forms by clause type. 169 instances of the overt form *ta* appeared in main clauses, 35 in subordinate clauses. The zero form was used in the main clause 209 times, but only 4 times in subordinate clauses. There appears to be a significant relationship between the pronominal form and the clause type ($X^2(1, N = 417) = 28{,}693$, $p < .01$): the zero person marker is more likely to occur only in the main clause than is the overt pronoun *ta*. In the subordinate clause, the overt form is preferred to zero.

**Table 6.** Pronoun form by clause type

| FORM | Main clause | Subordinate clause | TOTAL |
|---|---|---|---|
| *ta* | 169 | 35 | 204 |
| zero (SG) | 209 | 4 | 213 |
| SG forms total | 378 | 39 | 417 |
| *nad* | 31 | 3 | 34 |
| zero (PL) | 65 | 1 | 66 |
| PL forms total | 96 | 4 | 100 |
| TOTAL | 474 | 43 | 517 |

However, it must be kept in mind that the preference for overt forms in subordinate clauses might be connected to the circumstance that in subordinates, a non-nominative case may be the reason triggering the use of the overt form, as in example (18).

(18) ja=   sis   ø           avasta-s            et   ta-l=           on
    and  then  3SG.ZERO  discover-3SG.PST  that  3SG.SHORT-AD  be.3SG

    üks   korv   kadunud=  ja:
    one   basket  missing     and

    'and then he discovered that he has one basket gone missing'

## 6. Discussion and conclusion

This article has examined how the usage of the Estonian 3rd person overt pronoun and the zero person marker in spoken narrative discourse differ from each other. The starting point was that the overt pronoun and zero express equal levels of salience, therefore other factors must explain the distinction between them. The Centering Theory was applied to see if and how Centering transitions may help to interpret the difference between those two forms. Although a certain disparity in the transition types was found, there was still an overlapping area in the usage of the two forms, indicating that the effect on discourse salience is not completely different among the overt and zero forms. Therefore additional factors such as grammatical role, grammatical case, the possibility to realize the backward-looking center, and clause type were studied.

The Estonian data is compatible with previous findings that the use of zero for CONTINUE is a robust cross-linguistic phenomenon (see e.g. Di Eugenio 1998) and fuller forms are used for shifting Cb (see e.g. Hedberg and Dueck 1999). The results reported in the previous section indicate that the main difference between the Estonian overt form and zero lies in the fact that the use of the latter is restricted to only particular contexts, whereas the overt form can be used more freely in all possible semantically appropriate contexts. The zero form is applied to signal CONTINUE transitions, taking the role of a grammatical subject in the nominative case. It also must realize the backward-looking center Cb of an utterance, and is restricted to main clauses only. While all these traits are also very common to the overt form, other grammatical surroundings are likewise fully acceptable with the overt form *ta*: it can be found in all possible cases in Estonian, it can take any grammatical role in a sentence and it can occur in subordinate clauses.

Note that as the majority of data represents singular *ta* and zero, conclusions are valid primarily for these forms. Table 2 indicates that

plural forms behave slightly differently from singular forms, as in plural, zero forms are used noticeably more than overt forms. However, due to the scarce occurrence of plural forms, this cannot be claimed with full certainty and more language data is needed to analyze plural forms more reliably.

As the zero form can only be employed in the aforementioned conditions, it is probably more appropriate to apply the term zero subject instead. In previous research (Lindström 2005, Kivik 2010) this definition is indeed preferred. Moreover, the likelihood of reduced pronominal forms to realize higher arguments, i.e. subject rather than object1 and object1 rather than object2, appears to be a cross-linguistic tendency (Siewierska 2003). However, the scarcity of zero objects in the given study alone is not a sufficient argument for neglecting the possibility of zero object in Estonian. Rather, more supplementary and diverse data is needed to research the phenomenon.

The present study suggests that the initial hypothesis about the equal level of salience of zero and overt forms is valid, as the CONTINUE transition, which is the most frequent transition for both forms, is generally used when referring to highly salient entities. Nevertheless, considering the fact that the zero person marker is so restricted in its use, it is possible to see it as representing only the highest point of salience, as it has been argued that entities mentioned in main clauses are more salient than entities in subordinate clauses, and subjects are more salient than objects (see e.g. Kaiser 2000, Yoshida 2011). Still, the overt form is also completely acceptable in similar situations, so it is not possible to fully distinguish the two forms.

One possible way to account for why the speaker might use zero instead of the overt pronoun is the need to avoid the "repeated-name penalty" (Gordon et al. 1993). Originally, the repeated name penalty was reported to account for cases where a definite description was used instead of a pronoun to refer back to Cb (*Ibid*.). However, in Estonian, where there are two possible equally salient devices for referring back to an already-mentioned entity, it might be probable that the overt form is avoided in order to reduce processing time. It is also likely that the use of a repeated pronoun instead of zero where it would be acceptable indicates a new discourse segment boundary. (*Ibid*.)

This study has concentrated on spoken monologues in narrative contexts. High occurrences of zero person markers in the data clearly indicate that this form cannot be omitted from the description of the Estonian pronominal system. However, as the possible contexts for the

zero pronominal are undoubtedly restricted, further investigation is needed to fully determine its place in the system.

## Acknowledgements

**Address:**
Helen Hint
Eesti ja üldkeeleteaduse instituut
Tartu Ülikool
Jakobi 2, 51014 Tartu
E-mail: helen.hint@ut.ee

## References

Ariel, Mira (1990) *Accessing noun-phrase antecedents*. London and New York: Routledge.

Ballantyne, Keira Gebbie (2004) "Givenness as a ranking criterion in centering theory: evidence from Yapese". *Oceanic Linguistics* 43, 1, 49–72.

Brown-Schmidt, Sarah, Donna K. Byron, and Michael K. Tanenhaus (2005) "Beyond salience: interpretation of personal and demonstrative pronouns". *Journal of Memory and Language* 53, 292–313.

Byron, Donna K, Sarah Brown-Smith and Michael K. Tanenhaus (2008) "The overlapping distribution of personal and demonstrative pronouns". In Jeanette K. Gundel and Nancy Hedberg, eds. *Reference: interdisciplinary studies*, 143–175. Oxford: Oxford University Press.

Chafe, Wallace, ed. (1980) *The pear stories: cognitive, cultural, and linguistic aspects of narrative production*. Norwood, N.J.: Ablex.

Cristea, Dan, Nancy Ide and Laurent Romary (1998) "Veins theory: a model of global discourse cohesion and coherence". *Proceedings of ACL/COLING98*, 281–285.

Di Eugenio, Barbara (1998) "Centering in Italian". In Marilyn A. Walker, Aravind K. Joshi, and Ellen F. Prince, eds. *Centering theory in discourse*, 115–137. Oxford: Oxford University Press.

EKG II = Erelt, Mati, Reet Kasik, Helle Metslang, Henno Rajandi, Kristiina Ross, Henn Saari, Kaja Tael, Silvi Vare (1993) *Eesti keele grammatika II. Süntaks. Lisa: kiri*. Tallinn: Eesti Teaduste Akadeemia Keele ja Kirjanduse Instituut.

Erelt, Mati and Helle Metslang (2006) "Estonian clause patterns – from Finno-Ugric to Standard Average European". *Linguistica Uralica* 42, 4, 254–266.

Givón, Talmy (1983) *Topic continuity in discourse: a quantitative cross-language study*. Philadelphia and Amsterdam: John Benjamins.

Gordon, Peter C., Barbara J. Grosz, and Laura A. Gilliom (1993) "Pronouns, names, and the centering of attention in discourse". *Cognitive Science* 17, 311–347.

Grosz, Barbara J., Aravind K. Joshi, and Scott Weinstein (1995) "Centering: a framework of modelling the local coherence of discourse". *Computational Linguistics* 21, 2, 203–225.

Gundel, Jeanette (1998) "Centering theory and the givenness hierarchy: towards a synthesis". In. Marilyn A. Walker, Aravind K. Joshi, and Ellen F. Prince, eds. *Centering theory in discourse*, 183–198. Oxford: Oxford University Press.

Gundel, Jeanette, Nancy Hedberg, and Ron Zacharski (1993) "Cognitive status and the form of referring expressions in discourse". *Language* 69, 2, 274–307.

Hadic Zabala, Loreley and Maite Taboada. (2006) "Centering theory in Spanish: coding manual". Manuscript. Simon Fraser University.

Hedberg, Nancy and Sandra Dueck (1999) "Cakchiquel reference and centering theory". *Proceedings of the workshop on structure and constituency in the languages of the Americas. University of British Columbia working papers in linguistics*, 59–74.

Hu, Jianhua and Haihua Pan (2001) "Processing local coherence of discourse in Centering Theory". *Proceedings of the 15th Pacific Asia Conference on Language, Information and Computation*, 139–150. Hong Kong: City University of Hong Kong.

Iida, Masayo (1998) "Discourse coherence and shifting centers in Japanese texts". In Marilyn A. Walker, Aravind K. Joshi, and Ellen F. Prince, eds. *Centering theory in discourse*, 161–180. Oxford: Clarendon Press.

Kaiser, Elsi (2000) "Pronouns and demonstratives in Finnish: indicators of referent salience". In Paul Baker, Andrew Hardie, Tony McEnery, and Anna Siewierska, eds. *Proceedings of the 3rd discourse anaphora and reference resolution colloquium*, 20–27. (Technical Papers, 12.) Lancaster: University Center for Computer Corpus Research on Language.

Kaiser, Elsi (2009) "Investigating effects of structural and information-structural factors on pronoun resolution". In Malte Zimmermann and Caroline Féry, eds. *Information structure from different perspectives*, 332–353. Oxford: Oxford University Press.

Kaiser, Elsi and John C. Trueswell (2008) "Interpreting pronouns and demonstratives in Finnish: evidence for a form-specific approach to reference resolution". *Language and Cognitive Processes* 23, 5, 709–748.

Kaiser, Elsi and Katrin Hiietam (2003) "A typological comparison of third person pronouns in Finnish and Estonian". *Nordlyd* (Tromsø) 31, 4, 654–667.

Kaiser, Elsi and Virve-Anneli Vihman (2010) "On the referential properties of Estonian pronouns and demonstratives". In Hans Götzsche, ed. *In Memory, Mind and Language*, 193–205. Newcastle: Cambridge Scholars Publishing.

Kibrik, Andrej A. (2011) *Reference in discourse*. New York: Oxford University Press.

Kivik, Piibi-Kai (2010) "Personal pronoun variation in language contact: Estonian in the United States". In M. Norde, C. Hasselblatt, and B. de Jonge, eds. *Language contact – new perspectives*, 63–86. Philadelphia and Amsterdam: John Benjamins.

Lindström, Liina (2001) "Verb-initial clauses in narrative". In Mati Erelt, ed. *Estonian: Typological Studies* V, 138–168. Tartu.

Lindström, Liina (2005) *Finiitverbi asend lauses. Sõnajärg ja seda mõjutavad tegurid suulises eesti keeles*. (Dissertationes philologiae Estonicae Universitatis Tartuensis.) Tartu: Tartu University Press.

Lindström, Liina, Kristel Uiboaed, and Virve-Anneli Vihman (2014) "Varieerumine *tarvis-/vaja-* konstruktsioonides keelekontktide valguses". *Keel ja Kirjandus* 8–9, 609–630.

Metslang, Helena (2013) "Coding and behaviour of Estonian subjects". *Journal of Estonian and Finno-Ugric Linguistics* 4, 2, 217–293.

Pajusalu, Renate (2005) "Anaphoric pronouns in spoken Estonian: crossing the paradigms". In Ritva Laury, ed. *Minimal reference: the use of pronouns in Finnish and Estonian discourse*, 107–134. Helsinki: Suomalaisen Kirjallisuuden Seura.

Pajusalu, Renate (2006) "Death of a demonstrative: person and time: the case of Estonian *too*". *Linguistica Uralica* 42, 4, 241–253.

Pajusalu, Renate (2009) "Pronouns and reference in Estonian". *Sprachtypologie und Universalienforschung: Estonian in typological perspective* 62, 1, 122–139.

Passonneau, Rebecca J. (1998) "Interaction of discourse structure with explicitness of discourse anaphoric noun phrases". In Marilyn A. Walker, Aravind K. Joshi, and Ellen F. Prince, eds. *Centering Theory in Discourse*, 327–358. Oxford: Clarendon Press.

Poesio, Massimo, Rosemary Stevenson, Barbara di Eugenio and Janet Hitzeman (2004) "Centering: a parametric theory and its instantiations". *Computational Linguistics* 30, 3, 309–363.

Scott, Kate (2013) "Pragmatically motivated null subjects in English: a relevance theory perspective". *Journal of Pragmatics* 53, 68–83.

Siewierska, Anna (2003) "Reduced pronominals and argument prominence". In Miriam Butt and Tracy Holloway, eds. *Nominals: inside and out*. Stanford: CSIL Publications, 119–150.

Siewierska, Anna (2004) *Person*. Cambridge: Cambridge University Press.

Taboada, Maite (2008) "Reference, centers, and transitions in spoken Spanish". In Jeanette K. Gundel and Nancy Hedberg, eds. *Reference: Interdisciplinary studies*, 176–215. Oxford: Oxford University Press.

Taboada, Maite and Loreley Hadic Zabala (2008) "Deciding on units of analysis within Centering Theory". *Corpus Linguistics and Linguistic Theory* 4, 1, 63–108.

Vihman Virve-Anneli (2015) "Pick it up: a look at referential devices in Estonian child-directed speech". *Journal of Estonian and Finno-Ugric Linguistics* 6, 2, 63–83 (this volume).

Walker, Marilyn, Aravind Joshi and Ellen Prince (1998) "Centering in naturally occurring discourse: an overview". In Marilyn A. Walker, Aravind K. Joshi, and Ellen F. Prince, eds. *Centering theory in discourse*, 1–28. Oxford: Clarendon Press.

Walker, Marilyn, Masayo Iida and Sharon Cote (1994) "Japanese discourse and the process of Centering". *Computational Linguistics* 20, 2, 193–232.

Walker, Marylin A. (1998) "Centering, anaphora resolution, and discourse structure". In Marilyn A. Walker, Aravind K. Joshi, and Ellen F. Prince, eds. *Centering theory in discourse*, 401–435. Oxford: Clarendon Press.

Yoshida, Etsuko (2011) *Referring expressions in English and Japanese: patterns of use in dialogue processing*. Philadelphia and Amsterdam: John Benjamins.

## Abbreviations

| | | | | |
|---|---|---|---|---|
| 3 | – third person | | mINF | – *ma*-infinitive |
| ABL | – ablative | | NP | – noun phrase |
| AD | – adessive | | PART | – partitive |
| ALL | – allative | | PASS | – passive |
| Cb | – backward-looking center | | PL | – plural |
| Cf | – forward-looking center | | PPTCP | – past participle |
| COM | – comitative | | PRO | – pronoun |
| Cp | – preferred center | | PRS | – present |
| dINF | – *da*-infinitive | | PST | – past tense |
| EL | – elative | | SG | – singular |
| GEN | – genitive | | SHORT | – short pronoun |
| IN | – inessive | | ZERO | – zero person marker |
| | | | U | – utterance |

**Kokkuvõte**. **Helen Hint: Kolmanda isiku pronoomenivormid eesti keeles tsenderdamisteooria valguses.** Artikkel käsitleb eesti keele 3. isiku eksplitsiitse pronoomeni ja pronoomeni nullvormi erinevusi suulistes narratiivides. Mõlemad vormid väljendavad diskursuse kõige esilduvamaid entiteete, mistõttu ei saa nende erinevust selgitada pelgalt esilduvuse mõistele toetudes. Seega on artiklis kasutatud tsenderdamisteooriat (*the Centering Theory*), et näha, kas nullvorm ja eksplitsiitne 3. isiku pronoomen on seotud erinevustega diskursuse sidususes. Täpsemalt vaadeldakse, kas need kaks vormi erinevad tsenderamise üleminekute poolest. Põhjalikuma ülevaate saamiseks nulli ja *ta* erinevustest on võrreldud ka neile vormidele iseloomulikke grammatilisi faktoreid: grammatiline roll, kääne, lausetüüp. Analüüsist selgub, et nullvormi kasutuskontekstid on kitsamad, piirdudes JÄTKAMISE üleminekuga, nominatiivse subjektiga pealausetes, samas kui eksplitsiitne vorm esineb mitmesugustes kontekstides ja ümbrustes.

**Märksõnad:** pronoomeni nullvorm, 3. isiku pronoomen, tsenderdamisteooria, referents, suuline narratiiv, pirnilood