

# On the Role of Contrast in Pronoun Interpretation: Some Insights from Estonian

Elsi Kaiser\*

## 1 Introduction

In this paper, I use corpus data to test two competing claims concerning the referential properties of the Estonian long pronominal form *tema* 'she/he', namely that its use is sensitive to the salience/accessibility of its antecedent, or that its use is linked to contrast. I present the results of detailed corpus analyses investigating (i) the grammatical role of the antecedents of *tema*, (ii) the anaphoric forms used in Finnish versions of the same texts, and (iii) the presence/absence of contrast associated with uses of *tema*. As we will see, the corpus data indicate that salience does not appear to be relevant for use of *tema*, and that use of this form is triggered by the presence of contrast. Based on the data patterns, I present a detailed formulation of the referential properties of *tema* that—by making reference to different kinds of relations between propositions—aims to spell out what it means to say that this form is used contrastively.

Many researchers have noted that the form of referring expressions is connected to the salience of their referents. The claim is that the most reduced referring expressions refer to highly salient referents, and fuller expressions are used for less salient referents (e.g. Givón, 1983, 1984; Ariel, 1990; Gundel, Hedberg & Zacharski, 1993).<sup>1</sup> Thus, according to this kind of accessibility hierarchy type approach, if a language has both full and reduced pronouns, the full forms are said to refer to less salient referents than the reduced forms (see e.g. Givón, 1984; Schwartz, 1986; Bresnan, 2001).

Estonian is a language with both full and reduced forms of the gender-neutral third person pronoun: the reduced form *ta* 'she/he' and the full form *tema* 'she/he'. The prediction of accessibility hierarchy theories is thus that

---

\*I'd like to thank audiences at the University of Helsinki, the University of Pennsylvania and the University of Texas at Austin for useful comments and feedback. Thanks also to Christine Gunlogson, Renate Pajusalu, Ellen Prince and Maribel Romero. All errors are mine.

<sup>1</sup>These approaches resemble each other in that they all propose a ranking of referential forms depending on the salience or accessibility of the antecedent. However, there are other important differences between them, such as the degree of detail in their claims concerning the relation between particular referential expressions and accessibility statuses (see Ariel, 2001, for a detailed discussion).

*ta* refers to highly salient referents and *tema* to less salient referents. However, a different claim has been put forth by Pajusalu (1995, 1997), namely that use of *tema* has to do with comparison/opposition between referents. In this paper, I use corpus data to test these two competing views: (i) *tema* is used for lower-salience referents and (ii) *tema* is used for referents that are being compared to/opposed to other referents in the discourse. The structure of this paper is as follows. In section 2, I present an overview of the existing work on referring expressions in Estonian. In section 3, I use corpus data to test the claim that the long form *tema* is used to refer to referents that are lower in salience, and in section 4, I investigate the claim that use of *tema* is triggered by contrast/comparison, and attempt to spell out in more detail what it means for a pronoun to be contrastive. Section 5 is the conclusion.

## 2 Estonian Pronominal and Demonstrative Forms

In this section, I review the existing work on pronouns and demonstratives in Estonian. According to Pajusalu, the short pronominal form *ta* 's/he' is the most common way of referring to the entity that is currently in the focus of attention (Pajusalu, 1997:107). In contrast, the demonstrative *see* 'it, this' is described as referring to a less salient referent. For example, if there are two third human person referents in a clause, the short pronoun *ta* is used to refer to the first referent, and *see* to the second (Erelt *et al.*, 1993:209; Tauli, 1983:323). Some Southern dialects of Estonian use the distal demonstrative *too* 'that' instead of *see* 'it/this' to refer to the second-mentioned referent (Erelt *et al.*, 1993:209). Since, in this kind of 'personal pronoun' use, *see/too* seem to be simply dialectal variants of each other, in this paper I will group them together.

- (1) Tüdruk<sub>1</sub> vilksas pois<sub>2</sub> poole; ta<sub>1</sub>/see<sub>2</sub> oli kahvatu.  
 Girl-NOM glanced boy-GEN towards; ta-NOM/see-NOM was pale.  
 'The girl glanced towards the boy; she/he was pale.'

Now, if we combine the pattern exemplified in (1) with the well-known finding that subjects are more salient than objects and obliques (Brennan, Friedman & Pollard, 1987; Matthews & Chodorow, 1988; Stevenson *et al.*, 1994; McDonald & MacWhinney, 1995; *inter alia*), it looks like the use of *ta* vs. *see/too* is guided by the salience of the referent. The short pronoun *ta* is used to refer to a preceding subject, i.e. a higher-salience referent, and the demonstratives *see/too* are used to refer to the preceding object/oblique, which is lower in salience (see also Kaiser & Hiietam, 2004).

Now, let us consider the referential properties of the long pronominal form *tema*. According to Pajusalu, this form is used when the referent is being opposed or compared to another referent (Pajusalu, 1997:109).<sup>2</sup> It is worth noting that the *ta/tema* distinction doesn't map clearly onto a unstressed/stressed pronoun distinction: while *ta* cannot be stressed, *tema* can be stressed or unstressed (Pajusalu, 1995; citing Palmeos, 1981; Vsevirov, 1983). The example below (from Pajusalu, 1997), illustrates the use of *tema*. This is a context where the different dancers in a ballet are being announced. The first dancer has just been introduced, and then the speaker continues the list by introducing the role performed by the second dancer. The speaker uses *tema* when specifying what role the second dancer performs, since that is being implicitly compared/opposed to the first dancer's role. In the third clause, when the second dancer is salient and not being compared to anyone else, *ta* is used.

- (2) – ja teine klient on Maarika Aidla(.)  
         and second-NOM client-NOM is Maarika Aidla  
         tema tantsib seltsidaami(.)  
         tema-NOM dances lady-in-waiting-PART  
 – miks ta sinna prostituudi juurde satub (.)  
     why ta-NOM there prostitute-GEN at-ILL drops-in  
 – seda peab ise vaatama  
     see-PART must self-NOM watch  
     '– and the second dancer is **Maarika Aidla**.  
       *She dances the role of a lady-in-waiting*  
     – why does *she* visit the prostitute  
     – that you have to see for yourself'

## 2.1 Two Different Predictions

If we compare Pajusalu's claims about the discourse functions of *ta* and *tema* and the predictions of the accessibility hierarchy approaches, it becomes clear that these two approaches make different claims. If, as accessibility theories argue, full forms of pronouns are used for less salient referents than

---

<sup>2</sup>Pajusalu focuses on *ta/tema* in subject position because for non-nominative cases, choice of *ta* vs. *tema* is somewhat morphologically idiosyncratic and tends to be influenced by the particular case marker. She notes that when the pronoun is marked for a so-called exterior local case (e.g. *tal* s/he-ADES 'on him/her'), the tendency is to use the short form, whereas in interior local cases and when followed by a comitative postposition, the longer form is used (e.g. *temas* s/he-INES 'in him/her'; *temaga kaasa* s/he-COM with 'with him/her').

the reduced forms, we predict that *tema* is used for less salient referents and *ta* for highly salient referents. However, if as Pajusalu claims, *tema* is used for referents that are being compared or opposed to other referents present in the discourse, the salience of the referent is not the factor guiding the *ta/tema* alternation.

In this paper, by using a corpus of naturally-occurring tokens of *tema* in written Estonian, my aim is to investigate its referential properties in more detail in order to see which of these two claim better fits the data. Although my work is clearly related to Pajusalu's research, it differs from her work in that I use detailed quantitative corpus analyses to assess the validity of the salience/accessibility claims, and I also make reference to relations between propositions in order to spell out more clearly what it means to say that *tema* is used when the referent is compared/opposed to something else—i.e. what it means to say that *tema* is contrastive. In the next section, section 3, I present the results of the corpus analyses that I conducted in order (i) to find out what grammatical roles the antecedents of *tema* in my corpus have and (ii) to see what form is used in Finnish when Estonian uses *tema*—in particular, whether use of the Finnish demonstrative *tämä* 'this', which is used for low-salience referents, is correlated with use of *tema*. In section 4, we turn to the question of how to define contrast.

### 3 Salience

The corpus used for this study consists of 50 occurrences of *tema*, as well as their Finnish counterparts, from the four novels listed below. I wanted to compare Finnish and Estonian to see how *tema* shows up in the Finnish versions of the texts, given that there exists a demonstrative in Finnish, *tämä* 'this,' which is historically related to *tema* (Kulonen *et al.*, 2000). The referential properties of Finnish *tämä* have been fairly extensively studied, the general conclusion being that *tämä* is used for low-salience referents (e.g. Varteva, 1998; Halmari, 1994; Kaiser, 2000a, 2003, to appear).

- Finnish original: Utrio, Kaari. (1989). *Vendela*. Helsinki: Tammi.  
Estonian translation: Utrio, Kaari. (1996). *Vendela*. Tallinn: Sinisukk.  
(translator Maire Jürima)
- Finnish original: Utrio, Kaari. (1992). *Vaskilintu*. Helsinki: Tammi.  
Estonian translation: Utrio, Kaari. (1994). *Vasklind*. Tallinn: Sinisukk.  
(translator Andres Lepp)
- English original: Rowling, J.K. (2000). *Harry Potter and the Goblet of Fire*. New York: Scholastic.

Finnish translation: Rowling, J.K. (2001). *Harry Potter ja Liekehtivä Pikari*. Helsinki: Tammi. (translator Jaana Kapari)

Estonian translation: Rowling, J.K. (2000). *Harry Potter ja Tulepeeker*. Tallinn: Varrak. (translators Krista Kaer and Kaisa Kaer)

- English original: Rowling, J.K. (1999). *Harry Potter and the Prisoner of Azkaban*. New York: Scholastic.  
Finnish translation: Rowling, J.K. (2001). *Harry Potter ja Azkabanin Vanki*. Helsinki: Tammi. (translator Jaana Kapari)  
Estonian translation: Rowling, J.K. (2000). *Harry Potter ja Azkabani Vang*. Tallinn: Varrak. (translators Krista Kaer and Kaisa Kaer)

The data set was restricted in a number of ways, in order to ensure that I only analyzed instances of *tema* that could have surfaced as *ta*. I did not want to look at occurrences of *tema* that were required for grammatical, non-discourse-related reasons. Thus, I only included nominative occurrences of *tema* that were in subject position, because in the nominative case, both *tema* and *ta* are possible, whereas in some more oblique cases, *tema* is required or strongly preferred (e.g. Pajusalu, 1997, see also footnote 2).<sup>3</sup>

To test the validity of the claim that *tema* is used for referents that are not highly salient, I coded the 50 tokens in my corpus for (i) the grammatical role of the most recent preceding mention of the antecedent and (ii) what form was used in the Finnish version of the text.<sup>4</sup> If, as the salience approach predicts, full forms are used for referents that are not highly salient, we predict that *tema* will tend to refer back to nonsubject referents (given that subjects have been found to be more salient than objects), and we also predict that use of *tema* in Estonian will correlate with use of the low-salience anaphor *tämä* 'this' in the Finnish texts.

However, if use of *tema* has to do with contrast, then we do not expect it to show a preference for antecedents in non-subject position, since referents in any position can be interpreted contrastively, nor do we expect its use to

---

<sup>3</sup>I also did not include occurrences of *tema* in noun-noun coordinations, since *ta* is not possible in noun-level coordinations (see also Cardinaletti & Starke, 1999). Moreover, I did not include uses of *tema* modified by *ka* ('also,' 'too') or by the clitic *-ki/-gi* ('also, too, as well'), since these modifiers strongly prefer *tema* and rarely, if ever, occur with *ta* (see Tauli, 1983:334). I focused on reasons prompting the use of *tema* because *tema* is the marked form and *ta* seems to be the default form. In her corpus study, Pajusalu (1997) found 25 cases of *tema* and 126 cases of *ta*.

<sup>4</sup>It is worth pointing out that the majority of the sentences preceding the sentence with *tema* had canonical word order, and thus I do not investigate the role of word order in this paper (but see Kaiser, 2003, on Finnish). When I refer to objects or obliques, they can, for the most part, be assumed to be postverbal.

correlate with use of *tämä* in Finnish, since *tämä* has not been found to have contrast-related referential functions. (In fact, to the best of my knowledge, no anaphoric form in Finnish has been described as specifically having a contrast-related referential function.)

### 3.1 Results

Even a cursory examination of the data reveals that *tema* does not show a clear preference for low-salience referents, and can indeed be used to refer to a preceding subject, as in (3) (from *Harry Potter ja Azkabani Vang*, 305).

- (3) [context: Lupin's rights to attend the Hogwarts school are in danger, because of other parents' possible concerns about him—a werewolf—being near their children.]  
 Aga siis sai Dumbledore<sub>i</sub> direktoriks ..... Ta<sub>i</sub> ütles, et senikaua kui me võtame tarvitusele teatud abinõud, ei näe tema<sub>i</sub> mingit põhjust, miks ma ei peaks koolu tulema.  
*'But then Dumbledore<sub>i</sub> became Headmaster ..... He<sub>i</sub> said that as long as we took certain precautions, he<sub>i</sub> saw no reason why I shouldn't come to school.'*<sup>5</sup>

The observation that *tema* does not seem to act like a strictly object/oblique-preferring form is corroborated by corpus data. Figure 1 (on the next page) shows that although the grammatical roles of the antecedents of *tema* do not show an overwhelming preference for any one grammatical role, there does seem to be a slight preference for subjects (38%). For comparison purposes, note that preliminary corpus studies indicate that the short form *ta* has a strong subject bias (76% refer to matrix subject) and the demonstrative *see* a strong object/oblique preference (93% of occurrences refer to object or oblique) (Kaiser, 2003; see also Kaiser & Hiietam, 2004). The antecedents of *tema* do not show the same kind of clear preference for a particular grammatical role that we see with the antecedents of *ta* and *see*.

Now, we will turn to the Finnish counterparts of *tema*. Figure 2 (next page) shows what referential form was used in the Finnish version of the texts when *tema* was used in Estonian. Clearly, there is one dominant form used in Finnish where Estonian uses *tema*, namely the Finnish third person pronoun *hän* 'she/he' (80%). This preference, and the low occurrence of the Finnish demonstrative *tämä* 'this' (2%), shows that the discourse properties

<sup>5</sup>In the original English text, the structure is slightly different: "...He said that as long as we took certain precautions, there was no reason I shouldn't come to school."

of *tema* do not match the discourse properties of the *tämä*, even though the two words are historically related (Kulonen *et al.*, 2000). As mentioned above, Finnish *tämä* is used for low-salience referents.

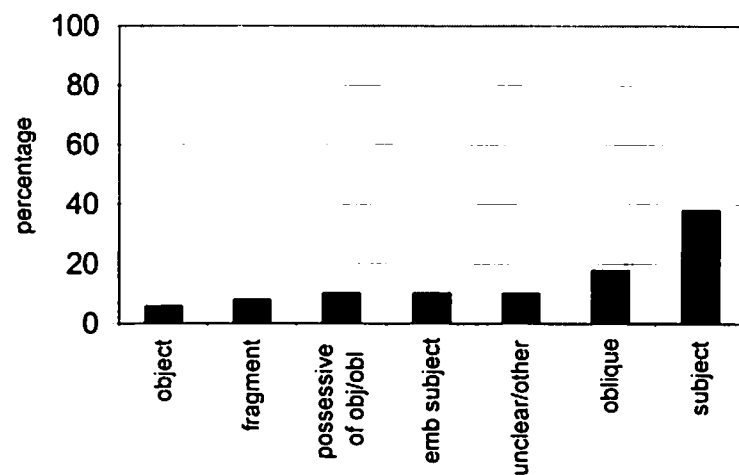


Figure 1. Grammatical role of the antecedent of *tema*<sup>6</sup>

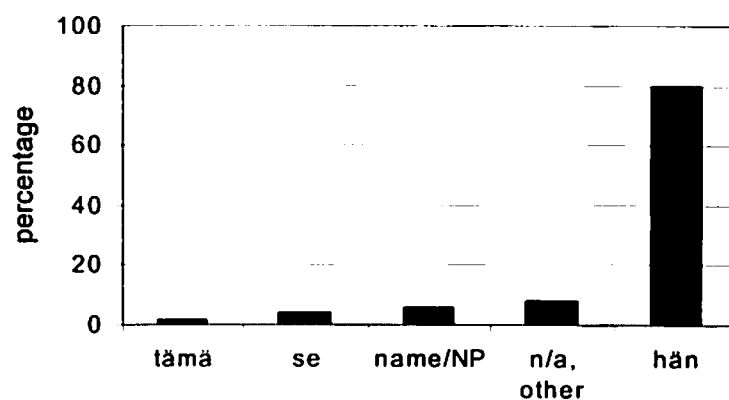


Figure 2. Counterpart of *tema* in Finnish versions of texts  
(Translations: *tämä* 'this', *se* 'it', *hän* 'she/he')

<sup>6</sup>'Fragment' means that the antecedent occurred in an exclamation or some other kind of 'sentence fragment' that did not constitute a complete sentence.

In sum, based on the corpus counts of the grammatical role of the antecedent of *tema* and the Finnish counterparts of *tema*, we can conclude that (i) the long form *tema* does not have a clear preference for nonsubjects, and (ii) it does not have similar referential properties as the Finnish demonstrative *tämä*. These findings do not support the claim that the full form *tema* is used for lower-salience referents, and this raises the question: If the salience of the antecedent is not what distinguishes *ta* and *tema*, is contrast the relevant factor? This is the question we tackle in section 4.

#### 4 Contrast

In the preceding section we saw that the antecedents of the long form *tema* are not associated with low-salience grammatical roles, nor does *tema* pattern like the Finnish low-salience anaphor *tämä* 'this'. The question thus remains, what is the discourse function of *tema*? Why would a speaker opt to use *tema* instead of the short form *ta*? In this section, we explore in more detail Pajusalu's claim that *tema* is used to refer to entities that are being compared to or opposed to some other referent(s) in the discourse. Before we turn to the Estonian data, let us think a bit more about the notions of contrast and set membership. I take as my starting point Prince's (1998, 1999) work on the discourse functions of topicalization in English, because, as we will see later, it has some interesting similarities with the Estonian data. Let's go through the basics of Prince's analysis by looking at the example below (from Prince, 1999):

- (4) She had an idea for a project. She's going to use three groups of mice. One, she'll feed mouse chow, just the regular stuff they make for mice. Another she'll feed them veggies. **And the third she'll feed junk food.**

Prince (1998, 1999; see also Hirschberg, 1985; Ward, 1985; Ward & Prince, 1991) shows that in English, a so-called topicalization can be structured into two main parts with different discourse-semantic properties:

- the topicalized element (e.g. 'the third') is a member of a contextually-salient set (in this example, the set of three mouse groups);
- the remaining sentence becomes a salient open proposition (OP) once the new information it contains is replaced with a variable.
  - OP = She'll feed the third X
  - X = junk food



It is worth noting that in English, for topicalization to be felicitous, the open proposition needs to be currently salient in the discourse. In contrast, in Yiddish, it is sufficient for the open proposition to be “known or at least plausible to the hearer at that point in the discourse” (Prince, 1999; see also Kaiser, 2000b, on Finnish OSV). Thus, there is crosslinguistic variation on what the ‘minimal salience’ of the open proposition needs to be.

Now, keeping this analysis in mind, let’s return to the Estonian data. Out of the 50 examples in my corpus, it turns out that all 50 involve some kind of contrast. More specifically, in light of the data patterns, I claim that the referential properties of *tema* can be summarized as follows:

- (5) Use of *tema* triggers an inference that the proposition P expressed by the sentence with *tema* is such that there exists some other relevant referent (y) for which a related form Q of that proposition holds.

This of course raises the question of what is the nature of relation between proposition P and proposition Q. The corpus data suggest that there are three (related) options:

- (i)  $Q = \neg P$
- (ii)  $Q = P$
- (iii) Q is related to P in the following way: P contains some new information, and when this new information is replaced with a variable, the result is an open proposition that is previously mentioned, known, or inferable/plausible.

Relation (iii) makes use of the notion of ‘open proposition’ as employed by Prince (1998, 1999). However, as we will see, in Estonian the open proposition does not need to be highly salient at that point in the discourse: It is sufficient for it to be known or inferable/plausible, similar to the requirements for OPs in Yiddish and Finnish topicalization (see Prince, 1999; Kaiser, 2000b). Thus, even though the third relation between P and Q seems ‘looser’ than the first two, it is nevertheless a specific kind of relation that is also instantiated in other aspects of natural language pragmatics.

#### 4.1 Contrastive use of *tema*

Let us now turn to some actual examples from the corpus, in order to see some of the concrete examples that motivate my formulation of *tema*’s referential properties presented in the preceding section. First, let us consider a case where  $Q = \neg P$ , i.e. a case where the sentence with *tema* is such that a

negated form of a related proposition holds for some other referent. In example (6a), the sentence with *tema* states that Sir Hartman cannot read, but the preceding context states that Vendela can read. Thus, this is a case of *tema* being used in a context where there exists some other relevant referent (Vendela) for which the negated form of the proposition “x can’t read” holds.

Now, let us consider a case of  $Q = P$ , i.e. where the sentence with *tema* is such that the same proposition has been claimed to hold of another referent in the preceding discourse. This type of use is exemplified by example (6b), where the sentence with *tema* states that Vendela will take care of the knight, and this is in a context where Father Henrik has earlier stated that he will take care of the knight. Thus, in this context, use of *tema* indicates a kind of corrective function.

- (6a) [context: Vendela has just told Sir Hartman that she can read and that she even owns a book, which was quite a rare possession in Finland in the year 1371]

Rüütel Hartman mõtiskles selle üle, lebades mõnusa laas voodis.

Tema ei osanud lugeda, selleks polnud mingit vajadust – lugemine oli pastorite osa. (Vendela, 107)

*‘Sir Hartman thought about this, resting comfortably in the wide bed.*

*He couldn’t read, there was no need for it – reading was for pastors.’*

- (6b) [context: Father Henrik wants to come along to take care of Sir Hartman, who is seriously ill. The head of Sir Hartman’s men explains to him:]

See [...] sõltub täielikult sellest, kas Domina Vendela lubab sul kaasa tulla või mitte. Domina Vendela on ravitseja. Tema ravib rüütli...

(Vendela, 94)

*‘It [...] depends entirely on whether Domina Vendela allows you to come along or not. Domina Vendela is a healer. She will take care of the knight...’*

An example of the third kind of relation between P and Q, i.e. a case where the proposition P expressed by the sentence with *tema* can be transformed into an open proposition that is previously mentioned or known/inferable, is exemplified in (6c).

- (6c) [context: A new woman is arriving at the village. Everyone has gathered to look at her, including another woman called Hailvi:]  
Naine oli ümmarguse näo ja väikese suuga. Tema vammus oli valmistatud kirkalt läikivast kangast; Rodarve Hailvi ütles, et see on

siid, ja *tema* juba teadis, sest Rodair Rodmundsson oli toonud talle  
siidriet Gaardarikest. (Vasklind, 98)

*'The woman had a round face and a small mouth. Her coat was made  
from brightly sparkling material; Hailvi of Rodarve said that is silk,  
and she indeed knew, because Rodair Rodmunsson had brought her  
silk from Gardarike.'*

In this example, the proposition expressed by the sentence with *tema* ('x knows what the material is') has not been explicitly applied to any referents in the preceding discourse, nor has its negated form been explicitly claimed to hold of anyone. Thus, the first two conditions do not hold ( $Q = \neg P$ ,  $Q = P$ ). However, it can be easily inferred from the context that the other onlookers do not know that the 'brightly sparking material' is silk. Thus, the proposition P ('x knows what the material is') is easily turned into an open proposition that is inferable: I follow Prince (1998) and treat the polarity of a proposition as new information that can be replaced with a variable to create an open proposition. Thus, 'x knows what the material is' becomes the open proposition 'x knows/doesn't know what the material is', which is inferable from context.

In fact, if we treat the polarity of an utterance as a variable that can be used to create an open proposition, then we could probably subsume the first two relations between P and Q (listed in (5)) under the third relation, since the negation relation  $Q = \neg P$  and the identity relation  $Q = P$  can be viewed as subcases of relation (iii). However, for reasons of expository clarity, and also because certain examples that fall under (iii) are much more complex and harder to code than those that fall under (i) and (ii) (to be discussed below), I keep the three relation types separate at this stage of research.

Another example that illustrates the third kind of relation between P and Q is ex. (2) from Pajusalu (1997). Here, the proposition P expressed by the sentence with *tema* ('x dances the part of a lady-in-waiting') becomes a salient open proposition ('x dances the part of y'), once the new information is replaced with a variable.

- (2) – ja teine klient on **Maarika Aidla**(.)  
and second-NOM client-NOM is Maarika Aidla  
**tema** tantsib seltsidaami(.)  
tema-NOM dances lady-in-waiting-PART  
– miks **ta** sinna prostituudi juurde satub (.)  
why ta-NOM there prostitute-GEN at-ILL drops-in  
– seda peab ise vaatama  
see-PART must self-NOM watch

- *and the second dancer is Maarika Aidla. She dances the role of a lady-in-waiting*
- *why does she visit the prostitute*
- *that you have to see for yourself*

In sum, in this section I have provided some naturally-occurring corpus examples to support my claim that use of *tema* triggers an inference that the proposition P expressed by the sentence with *tema* is such that there exists some other relevant referent for which a related form Q of that proposition holds. Clearly, a number of questions remain open for future work. In particular, a larger corpus of *tema* should be collected in order to see if the patterns I observed in my corpus also occur on a larger scale and in different types of texts. Along with the collection and coding of a larger corpus comes another challenge, namely the need to establish reliable coding guidelines, in particular concerning the relations between P and Q. The first two relations ( $Q = \neg P$ ,  $Q = P$ ) are straightforward to code, but deciding whether or not we can derive from P an open proposition that is previously mentioned or inferable/plausible, is more difficult. Detailed coding guidelines, based on insights derived from natural corpus examples and existing coding manuals, will be necessary. This is an important direction for future work.

## 5 Conclusions

In this paper, by using a corpus of naturally-occurring examples of the Estonian third person pronoun *tema*, I investigated its referential properties in detail in order to find out whether the long form *tema* (as opposed to the short form *ta*) is used to refer to referents that are not highly salient, or whether it used for contrastive referents. The corpus data concerning the grammatical role of the antecedents of *tema* indicate that this form does not show a preference for low-salience referents, and this conclusion is corroborated by the Finnish data which shows that Estonian *tema* and Finnish *tämä* ‘this’ clearly do not have the same referential properties.

On the other hand, the idea that *tema* is used in cases of contrast receives support from the corpus data. More specifically, the corpus data show that use of *tema* triggers the following inference: the proposition P expressed by the sentence with *tema* is such that there exists some other relevant referent y for which a related form Q of that proposition holds. The relation between the propositions P and Q can take one of three related forms: It can be the identity relation ( $Q = P$ ), the negation relation ( $Q = \neg P$ ) or an ‘open proposition’-based relation which states that when the new information in P is replaced with a variable, the result is an open proposition

that is previously mentioned, known or inferable/plausible (see also Prince, 1998, 1999 on topicalization).

On a more general level, the finding that the referential properties of the long form *tema* are not driven by salience fits in with evidence from Dutch and Finnish (Kaiser, 2003, to appear) indicating that different anaphoric forms within one language can be sensitive to different factors, and that their referential properties cannot necessarily be captured by a unified notion of salience. Findings concerning the interpretation of stressed pronouns (Venditti, Stone, Nanda & Tepper, 2002; de Hoop, 2003, *inter alia*) provide further evidence in support of this view.

In other words, crosslinguistic data indicate that the 'retrieval instructions' carried by different referential forms are not restricted to statements about the level of salience of the antecedent, and can in fact inform the hearer how to locate the antecedent by using other kinds of information. In the case of *tema*, it is worth noting that use of this form in fact provides new information about its referent, as *tema* can be used to signal that its referent (with respect to the predicate in which *tema* occurs) contrasts with some other referent(s) present in the discourse. Thus, use of *tema* can add to our knowledge of its referent. In contrast, 'regular pronouns' like *he*, *she* etc. do not add any new information to our knowledge of the referent. These forms simply encode retrieval instructions which make reference to the properties that the antecedent already possesses.

## References

- Ariel, Mira. 2001. Accessibility Theory: An overview. In T. Sanders, J. Schilperoord and W. Spooren, eds., *Text Representation, Linguistic and Psycholinguistic Aspects*. Amsterdam/Philadelphia: John Benjamins, 29–87.
- Ariel, Mira. 1990. *Accessing NP Antecedents*. London: Routledge, Croom Helm.
- Arnold, Jennifer. 1998. Reference Form and Discourse Patterns. Doctoral dissertation, Stanford University, Stanford, CA.
- Brennan, Susan E., Marilyn A. Friedman, and Carl J. Pollard. 1987. A Centering approach to pronouns. In *Proceedings of the 25th ACL*, 155–162.
- Bresnan, Joan. 2001. The Emergence of the Unmarked Pronoun. In *Optimality-Theoretic Syntax*, ed. G. Legendre, J. Grimshaw & S. Vikner, 113–142. Cambridge, MA: The MIT Press.
- Cardinaletti, Anna & Starke, Michel. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In Henk van Riemsdijk, ed., *Clitics in the Languages of Europe*. Berlin/New York: Mouton de Gruyter, 145–233.
- Givón, Talmy. 1983. *Topic Continuity in Discourse: A Quantitative Cross-Language Study*. Amsterdam: John Benjamins.

- Givón, Talmy. 1984. *Syntax. A Functional-Typological Introduction*, Vol. 1. Amsterdam/Philadelphia: John Benjamins.
- Gundel, Jeanette K., Nancy Hedberg, and Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69:274–307.
- Halmari, Helena. 1994. On accessibility and coreference. *Nordic Journal of Linguistics* 17:35–59.
- Hirschberg, Julia. 1985. A Theory of Scalar Implicature. Doctoral dissertation, University of Pennsylvania.
- De Hoop, Helen. 2003. On the interpretation of stressed pronouns. In M. Weisgerber, ed., *Proceedings of the 7<sup>th</sup> Sinn und Bedeutung Conference, Arbeitspapier Nr. 114, FB Sprachwissenschaft*. Universität Konstanz, Germany, 159–172.
- Kaiser, Elsi. To appear. When salience isn't enough: Pronouns, demonstratives and the quest for an antecedent. To appear in Ritva Laury, ed., *Minimal Reference in Finnic: The Use and Interpretation of Pronouns and Zero in Finnish and Estonian Discourse*. Helsinki, Finland: Suomalaisen Kirjallisuuden Seura.
- Kaiser, Elsi. 2003. The Quest for a Referent: A Crosslinguistic Look at Referent Resolution. Doctoral dissertation, University of Pennsylvania.
- Kaiser, Elsi. 2000a. Pronouns and demonstratives in Finnish: Indicators of referent salience. In P. Baker, A. Hardie, T. McEnery and A. Siewierska., eds., *Proceedings of the Discourse Anaphora and Anaphor Resolution Conference*. Lancaster, UK: University Center for Computer Corpus Research on Language, Technical Papers vol.12, 20–27.
- Kaiser, Elsi. 2000b. The discourse functions and syntax of OSV word order in Finnish. In A. Okrent, and J. Boyle, eds., *The Proceedings from the Main Session of the Chicago Linguistic Society's Thirty-sixth Meeting*, vol. 36(1). Chicago, Ill.: Chicago Linguistic Society, 179–194.
- Kaiser, Elsi and Katrin Hietam. 2004. A comparison of the referential properties of third person pronouns in Finnish and Estonian. *Nordlyd* 31(4):654–667
- Kulonen, Ulla-Maija et al. 2000. *Suomen sanojen alkuperä*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Matthews, A. and M. Chodorow. 1988. Pronoun resolution in two-clause sentences: Effects of ambiguity, antecedent location, and depth of embedding. *Journal of Memory and Language* 27:245–260.
- McDonald, Janet L. and Brian MacWhinney. 1995. The time course of anaphor resolution: Effects of implicit verb causality and gender. *Journal of Memory and Language* 34:543–566.
- Pajusalu, Renate. 1995. Pronominit *see*, *tema* ja *ta* viron puhekielessä. *Sananjalka* 37:81–93.
- Pajusalu, Renate. 1996. Pronoun systems of Common Estonian and Estonian dialects in a contrastive perspective. In *Estonian: Typological Studies*. Publications of the Department of Estonian of the University of Tartu, Volume 4, 19–164.
- Pajusalu, Renate. 1997. Eesti pronomeneid I. Ühiskeele *see*, *too* ja *tema/ta*. *Keel ja Kirjandus*, 24–30 and 106–115.
- Palmes, Paula. 1981. *Eesti keele grammatika II Morfoloogia. Arv- ja asesõna*. Tartu.

- Prince, Ellen F. 1999. How not to mark topics: 'Topicalization' in English and Yiddish. *Texas Linguistics Forum*, Chapter 8.
- Prince, Ellen F. 1998. On the limits of syntax, with reference to Topicalization and Left-Dislocation. In P. Culicover and L. McNally, eds., *Syntax and Semantics. The Limits of Syntax* 29. New York: Academic Press, 281–302.
- Schwartz, Linda. 1986. The function of free pronouns. In U. Wiesemann, ed., *Pronominal systems*. Tübingen: Gunter Narr, 405–36.
- Stevenson, Rosemary J., Rosalind J. Crawley, and David Kleinman. 1994. Thematic roles, focus and the representation of events. *Language and Cognitive Processes* 9:519–548.
- Tauli, Valter. 1983. *Standard Estonian Grammar. Part II Syntax*. Acta Universitatis Upsaliensis, Studia Uralica et Altaica Upsaliensia. Uppsala.
- Vallduví, Enric and Maria Vilks. 1998. On rheme and kontrast. In P. Culicover and L. McNally, eds., *Syntax and Semantics. The Limits of Syntax* 29. New York: Academic Press, 79–108.
- Varteva, Annukka. 1998. Pronominit *hän* ja *tämä* tekstissä. *Virittäjä* 2/1998:202–223.
- Venditti, Jennifer, Matt Stone, Preetham Nanda, and Paul Tepper. 2002. Discourse constraints on the interpretation of nuclear-accented pronouns. *Proceedings of the 2002 International Conference on Speech Prosody*. Aix-en-Provence, France.
- Vesio, Tiiu. 1983. Lausetevaheline pronominalisatsioon eesti keeles. Thesis, University of Tartu.
- Ward, Gregory and Ellen F. Prince. 1991. On the topicalization of indefinite NPs. *Journal of Pragmatics* 16:167–177.
- Ward, Gregory. 1985. The Semantics and Pragmatics of Preposing. Doctoral dissertation, University of Pennsylvania.

Center for Language Sciences, Meliora Hall  
University of Rochester  
Rochester, NY 14627  
[ekaiser@ling.rochester.edu](mailto:ekaiser@ling.rochester.edu)