# The Estonian pronoun *tema* and the role of contrast in pronoun interpretation

The form of referring expressions is said to be connected to the accessibility / topicality of their referents: The most reduced referring expressions refer to highly accessible referents; more marked expressions are used for less accessible referents (e.g. Ariel 1990, Gundel, Hedberg & Zacharski 1993). In languages with full and reduced pronouns, full forms are said to refer to less accessible referents than reduced forms (e.g. Bresnan 2001).

In this paper, I use corpus data to test the predictions these accessibility-based claims make for the Estonian long pronoun *tema* (he/she) in relation to the reduced form *ta* (he/she). In particular, I explore how these predictions relate to Pajusalu's claim (1995, 1997) that *tema* is used for referents that are being compared or opposed to other referents. In light of my corpus data, I claim that an accessibility-based approach makes incorrect predictions about *tema*, but also that Pajusalu's claim is not adequate to capture the referential properties of *tema*. I present a new analysis based on Büring's notion of 'contrastive topic,' and illustrate how this analysis can be used to unify the different uses of *tema* in a theoretically meaningful way.

Let us now take a look at the corpus data, the empirical foundation for my analysis. If we combine the accessibility-based claim that full forms refer to less accessible referents with the finding that subjects are more accessible than objects/obliques (e.g. Brennan, Friedman & Pollard 1987), we predict that *tema* will not be used to refer back to subjects. However, in my corpus (50 tokens of *tema* in subject position, randomly selected from four novels), I found that if we look at the grammatical roles of the antecedents of *tema*, we find a preference for subjects (48%) over objects or obliques (both combined = 24%). (A detailed discussion of grammatical roles, as well as frequency information about reference to antecedents in sentence fragments and other contexts not mentioned here will be provided in the talk.) In other words, *tema* is more likely to have a subject antecedent than an object or oblique antecedent. (Not surprisingly, a comparison corpus of 50 occurrences of the short pronominal form *ta* reveals this same pattern, but more strongly.) This subject-preference pattern does not match the prediction that *tema* is used for low-accessibility, non-subject referents.

A thorough analysis of the data reveals that almost all uses of *tema* in the corpus involve explicit or implicit contrast. In some cases, the contrast between the antecedent of *tema* and other referent(s) is explicit, and many of these 'explicit contrast' cases fit Pajusalu's criteria of comparison or opposition. However, in many cases the contrast is implicit and not inferrable without the help of context. At first glance (and out-of-context), these two usage types – explicit and implicit contrast – look rather different. However, I show that they can be unified by the notion of 'contrastive topic' in the sense of Büring (2002). Büring uses 'contrastive topic' to mean a member of a salient set that contrasts with other members, as specified by the predicate. I extend this notion to Estonian, and discuss in detail what it means to claim that *tema* is used for contrastive referents. In fact, my analysis aims to (i) capture the referential properties of *tema* and to show that it is not a 'low-salience' version of the short for *ta*, and also (ii) to provide a theoretically sound and testable way of describing the notion of 'contrast,' which is a term that often crops up in the pronoun literature but whose precise meaning, with respect to pronoun interpretation, is rarely discussed in any detail.

The imporance of making a clear distinction between the referential properties of contrastive form *tema* and the 'regular pronoun' *ta* is highlighted by the fact that the antecedent of *tema* is not interpreted contrastively in the discourse preceding the sentence with *tema*. In other words, use of *tema* provides new information about its referent (i.e., that it is contrastive), whereas a regular pronoun like *ta* simply acts as a retrieval instruction for the intended referent and does not convey new information by virtue of its referential form.

In sum, in this paper I use corpus data to show that the referential properties of the long pronominal form *tema* do not fit accessibility-based predictions, since use of *tema* does not depend on salience. Moreover, I claim that we can extend Büring's notion of 'contrastive topic' to provide a precise definition of what kind of contrast characterizes use of *tema*.

**References:**

Ariel, M. 1990. *Accessing NP antecedents*. London: Routledge, Croom Helm.

Brennan, S. E., Friedman, M. A. & Pollard, C. J. 1987. A Centering approach to pronouns. In *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics*, 155-162. Stanford, CA.

Bresnan, J. 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.

Büring, D. 2002. *On D-Trees, Beans, and B-Accents*. UCLA Manuscript (to appear in Linguistics & Philosophy).

Gundel, J.K., Hedberg, N. & Zacharski, R. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69:274-307.

Pajusalu, R. 1995. Pronominit *see*, *tema* ja *ta* viron puhekielessä. *Sananjalka* 37:81-93.

Pajusalu, R. 1997. Eesti pronoomeneid I. Ühiskeele *see, too* ja *tema/ta*. *Keel ja Kirjandus*, 24-30 and 106-115.