

TARTU ÜLIKOOLI EESTI KEELE ÕPPETOOLI TOIMETISED 14

---

PUBLICATIONS OF THE DEPARTMENT OF ESTONIAN OF THE  
UNIVERSITY OF TARTU 14

ESTONIAN:  
TYPOLOGICAL STUDIES IV

edited by  
Mati Ereht

---

TARTU 2000

## ESTONIAN: TYPOLOGICAL STUDIES IV

TARTU ÜLIKOOLI EESTI KEELE ÕPPETOOLI TOIMETISED 14

---

PUBLICATIONS OF THE DEPARTMENT OF ESTONIAN OF THE  
UNIVERSITY OF TARTU 14

# ESTONIAN TYPOLOGICAL STUDIES IV

edited by  
Mati Ereht

ESTONIAN: TYPOLOGICAL STUDIES IV

This work was supported

by the Estonian Science Foundation, Grant No.: 3259.

# CONTENTS

## **Mati Erelt**

On the Pro-Forms in Estonian ..... 7

## **Külli Habicht**

Grammaticalization of Adpositions in Old Literary Estonian ... 19

## **Helle Metslang**

Reflections on the Development of a Particle in Estonian ..... 59

## **Renate Pajusalu**

Indefinite Determiners *mingi* and *üks* in Estonian ..... 87

## **Urmas Sutrop**

Basic Terms and Basic Vocabulary ..... 118

## **Ilona Tragel**

Usage of the Verb *andma* 'to give'  
in the Domain of Verbal Communication ..... 146

## **Enn Veldi**

Estonian and English:  
a Lexicographer's Point of View. I ..... 173

## **Haldur Õim**

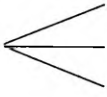
*Otse, sirge* and *õige*:  
a Domain of Metaphoric Extension in Estonian ..... 198

# ON THE PRO-FORMS IN ESTONIAN

**Mati Erelt**

*University of Tartu*

Estonian pro-forms include pronouns and pro-adverbs, but pro-verbs are absent:

pronouns  prosubstantives: *mina* 'I', *ise* 'oneself' etc.  
proadjectives: *niisugune* 'this kind of' etc.  
pronomerals: *mitu* 'several' etc.  
proadverbs: *siin* 'here', *nii* 'so' etc.

There are six kinds of **pronouns** in Estonian:

- 1) **personal pronouns**: *mina~ma* 'I' *sina~sa* 'you sg.' *tema~ta* 'he/she' *meie~me* 'we' *teie~te* 'you pl.' *nemad~nad* 'they';
- 2) **reflexive pronouns**: *ise: enese~enda* 'oneself', *oma* 'one's', *iseenda~iseenese* 'one's own' *omaenda~omaenese* 'one's own';
- 3) **reciprocal pronouns**: *teineteise* 'each other', *üksteise* 'one another';
- 4) **demonstrative pronouns**: *see (need)* 'this' ('these'), *too (nood)* 'that' ('those'), *sama* 'same' *samane* 'same' *see-sama (needsamad)* 'the same over here' *toosama (nood-samad)* 'the same over there', *samasugune* 'the same kind of' *niisamasugune* 'the same kind of' *seesamagune* 'the same kind of' *niisugune* 'this kind of' *seesugune* 'this kind of', *selline* 'such' *säärane* 'such', *säherdune* 'such' *see-sinane* 'the very', *sinane* 'the very' *sihuke(ne)* 'such', *nihuke(ne)* 'such';
- 5) **interrogative-relative pronouns**: *kes* 'who' *mis* 'what, that', *kumb* 'which' *missugune* 'which, what kind of', *milline* 'which, that' *mäherdune* 'what kind of', *misuke* 'what kind of' *määrane* 'what kind of';
- 6) **indefinite pronouns**: *keegi* 'someone', *miski* 'something' *mingisugune* 'some, a certain', *miskisugune* 'some' *ükski* 'none of the' *mitu* 'several', *mitmes* 'in several' *mitu~setu* 'several, quite a few' *mitmes~setmes, palju* 'many, a large

number of; much, a large amount of' *üks* 'a' *mõningane*  
 'some' *iga* 'every, each' *igamees* 'everyone' *kõik* 'all'  
*mõlemad* 'both' *kumbki* 'either, neither, *emb-kumb* 'either',  
*kogu* 'all, entire' *terve* 'all, entire'

The Estonian pronoun system is remarkable in that the personal pronouns have full forms and short forms (see Pool 1999 about their usage), the relative and interrogative pronouns have the same form (*Mida sa ütlesid?* 'What did you say?' See, *mida sa ütlesid, on vale* 'What you said is wrong'), and that the possessive pronouns are absent.

As far as the latter statement is concerned, one can find some disagreement among the Estonian linguists. Nobody denies the absence of the possessive pronouns of the Indo-European type, i.e. the adjectival equivalents of personal pronouns, e.g. Russ *мой, твой, наш, ваш*, Ger *mein, dein, sein, ihr*, Eng *my, your, our*, etc. However, most Estonian grammars have termed *oma* and *enese~enda* (as well as the compound pronouns formed with them) as possessive pronouns where they occur as the attribute to a noun. These grammars, including the academic grammar of Estonian, provide a classification of the pronouns, which includes two identical sets (EKG I: 28): **reflexive pronouns**: *enese~enda, oma, iseenese~iseenda, omaenese~omaenda*; and **possessive pronouns**: *enese~enda, oma, iseenese~iseenda, omaenese~omaenda*.

The pronoun *enese~enda* is claimed to have a defective paradigm (without the nominative). This pronoun, as well as the other enumerated pronouns, is considered to be reflexive when it occurs as an extension to the noun – the attribute (2).

- (1) *Mees tagus endale/omale rusikaga vastu rinda*  
 'The man was thumping his chest with his fist'
- (2) *Mees läks oma/enda naisega tülli*  
 'The man started a quarrel with his wife'

However, semantically both uses are reflexive because they refer to the (logical) subject. Thus, the classification of pronouns has been based only on the syntactic relations of the pronoun and not on the referential relationships.

On the other hand, in the grammars of Indo-European languages the adjectival equivalents of personal pronouns have been regarded as possessive pronouns, e.g. Russ *мой, твой, наш, ваш*, Ger *mein, dein, sein, ihr*, Eng *my, your, our* do not include a reference not to the subject but to the person. The possessive pronouns can be used also in those sentences where the person is not the subject as in the Russian sentence *Олег брал твою книгу* 'Oleg took your book'. Thus, possessiveness is associated with the speech act and not with the proposition. The action situation is associated with reflexivity. Some Russian grammarians have, in fact, regarded the Russian pronoun *свой* as a possessive pronoun, calling it a reflexive possessive pronoun (GRJ 1952: 393). However, we cannot follow this treatment because it would violate the consistency of the semantic classification.

One could make an attempt to justify the treatment of possessive pronouns as a separate category by stating that reflexive pronouns are pro-substantives, but possessive pronouns are pro-adjectives. On the other hand, it is only *oma* that can be adjectival. The pronoun *enese~enda* remains a pro-substantive in any position. However, adjectiveness is not a sufficiently good reason to distinguish possessiveness-reflexiveness in the case of *oma* either. The reason for this lies in the fact that there are also other pronouns with a double nature as far as the parts of speech are concerned. Nevertheless, actually they represent a single kind of pronouns, e.g. the demonstrative pronoun *see: Ma näen seda. See mees on tark.* 'I can see it. This man is wise'

If we treat the difference of the referential relationship as the basic difference between reflexiveness and possessiveness, then the possessive pronouns are absent in Estonian. In place of the possessive pronouns we use the genitive forms of the personal pronouns: *minu* 'my' *sinu* 'your' *tema* 'his, her' *meie* 'our', *teie* 'your pl.' *nende* 'their'. Both the pronoun *enese~enda* and the pronoun *oma* could be regarded as reflexive pronouns. They could be characterized as follows: the former is primarily a verb dependent and the latter primarily a substantive modifier – the attribute.

Among the Estonian grammarians, only Valter Tauli (1972: 66–67) dared to state that the possessive pronouns are absent in



Estonian, and *enese~enda* and *oma* can be regarded as reflexive pronouns also when they have an attributive nature. Unfortunately, Tauli did not provide any comments about his classification of the pronouns.

Actually, many Estonian grammars include only a single kind of pronouns, which is not singled out as a separate category in this article (similarly to Tauli's grammar). They are the so-called **determinative pronouns** (cf. EKG I: 30): *ise*, *oma*, *iga*, *igamees*, *kõik*, *mõlemad*, *kumbki*, *emb-kumb*, *kogu*, *terve*.

Estonian grammars characterize determinative pronouns as pronouns that modifies the substantive and can bring into prominence its referent or emphasize its integrity. The highlighting function is fulfilled by the pronoun *ise*, the paradigm of which coincides starting with the genitive with the paradigm of *enese~enda*, and the pronoun *oma*. For example,

- (3) *Ta ise ei teadnud sellest midagi*  
 'He himself knew nothing about it'  
 (4) *See on tema enda asi*  
 'That's his own business'  
 (5) *See on mu oma asi*  
 'That's my own business'

Syntactically, these pronouns function as the apposition. The function of 'emphasizing the integrity' which is essentially a quantifying function, is fulfilled by the other determinative pronouns: *kõik*, *iga*, etc. For example,

- (6) *Kõik kuulajad olid esinejast vaimustatud*  
 All the listeners were enthusiastic about the presenter'

Thus, the determinative pronouns bring together two groups of pronouns that fulfill rather different essential functions.

It would be worthwhile to treat the appositive *ise* and *oma* as reflexive pronouns as well. On the other hand, there is no explicit referential relationship with the subject here as one is dealing with a noun-phrase-internal relationship. However, the appositive *ise* assumes that what is designated by its head is already known; the referent has been discussed earlier. If we extend the concept of reflexiveness from the (logical) subject to what is already known

(from the semantic point of departure to the communicative point of departure), then *ise* (and analogously *oma*) could be included among the reflexive pronouns. Also, in English grammar the appositive use of the pronouns *myself*, *yourself*, etc., as in the sentence *I myself did it*, is treated as the emphatic use of the reflexive pronouns.

Thus, I would like to claim that there is no pronoun *enese~enda* with a defective paradigm in Estonian, which in one case is a reflexive and in another case a possessive pronoun, contrasting with the determinative pronoun *ise* with a full paradigm. From the purely synchronic point of view, Estonian has the reflexive pronoun *ise*, which can occur in various syntactic positions – as a verb dependent, attribute, and the apposition. Its case depends on its current syntactic position. As a verb dependent it cannot be used in the nominative, which does not mean, however, that it has no nominative at all. As an attribute, it cannot occur in other cases than the genitive. Only as an apposition can the reflexive pronoun *ise* be used in all the cases. The different syntactic positions have different constraints on usage, but reflexivity as a uniting feature is present in all three cases.

In the same way there is only one reflexive pronoun *oma*. If *ise* functions primarily as a verb dependent and secondarily as an attribute or an apposition, then *oma* functions primarily as an attribute but secondarily as a verb dependent or an apposition.

At first it remains somewhat unclear what exactly unites all these pronouns that EKG treats as the indefinite pronouns and those determinative pronouns that are left over once *ise* and *oma* are excluded. Is it indefiniteness that unites this class, as English grammarians think about their analogous pronouns (and Valter Tauli may have followed their example when discussing the Estonian pronouns); is it the quantifying function, as mostly Russian linguists claim; or could it be both. The feature that was selected for the above-mentioned list on the first page was indefiniteness.

The majority of Estonian grammars do not treat **pro-adverbs** as a separate part of speech. They have been regarded as common adverbs, not considering their substitutionary function. EKG I is

the first grammar that treats such words as a separate part of speech (1995: 32). However, even EKG passes rather smoothly over the pro-adverbs, mentioning only some more important adverbial functions of the pro-adverbs and not providing a classification according to the substitutionary function.

Pro-adverbs function, on the one hand, as adverbs by fulfilling a semantic function (semantic role); on the other hand, they function as pronouns by fulfilling the substitutionary function. Pro-adverbs are peculiar in that both functions are realized lexically. Pro-adverbs are lexical means for fulfilling certain functions, which should otherwise be expressed by the declinable form of the pronoun (*selleks*), a combination of the pronoun and the adposition (*selle jaoks*), or a combination of the pronoun and a function substantive (*sel viisil*).

The pro-adverbs have fewer substitutionary functions than the pronouns and fewer semantic roles than the true adverbs. From the point of view of substitution Estonian has five kinds of pro-adverbs: 1) reflexive, 2) reciprocal, 3) demonstrative, 4) interrogative-relative, and 5) indefinite pro-adverbs. From the adverbial point of view Estonian pro-adverbs can be classified into at least six types: 1) local, 2) temporal, 3) modal (of manner and state), 4) quantitative (quantity and degree), 5) causal-final, and 6) pro-adverbs of concession. There are  $5 \times 6 = 30$  possible combinations, the actual number of combinations amounting to 21. The following Table describes the possible combinations (each kind is represented by one or more more typical pro-adverbs).

Table 1. Pro-adverbs in Estonian

	Refl	Rec	Dem	Int-rel	Indef
<b>Loc<sup>1</sup></b>			sinna 'to there' seal 'over there' sealt 'from there' siia 'to here' siin 'here' siit 'from here'	kuhu 'to where' kus 'where' kust 'from where'	kuskile 'to somewhere' kuskil 'somewhere' kuskilt 'from somewhere' kõikjale 'to everywhere' kõikjal 'everywhere' kõikjalt 'from everywhere'
<b>Temp</b>			siis 'then' tollal 'at that time'	millal kunas 'when' kuniks 'how long'	kunagi 'once' millalgi 'at a certain time' iial(gi) 'never' alati 'always'
<b>Mod</b>	ise 'one-self' isekeskis 'between oneselfes'	omavahel 'between oneselfes' vastastikku 'mutually' sedasi 'this way'	nii 'so' nõnda 'so'	kuidas 'how'	kuidagi 'somehow' teisiti 'otherwise' igati 'in every way'
<b>Quant</b>			niivõrd 'so much'	kuivõrd 'as'	kuigi 'although' kuigivõrd 'to a certain degree'
<b>Caus-Fin</b>			sellepärast 'therefore' seetõttu 'therefore' niigi 'anyway'	miks 'why' mispärast 'why' milleks 'for what'	millegipärast 'for some reason'
<b>Conc</b>			nagunii 'anyway' sellegipoolest 'nevertheless'		

Some examples:

(7) Loc/Dem

*Lähen sinna, olen seal, tulen sealt*

'I'll go there, I'll be there, I'll come from there'

<sup>1</sup> Local adverbs and pro-adverbs form three-member series in Estonian – lative, locative, and separative.

- (8) Loc/Int-Rel  
*Kuhu lähed? Kus sa oled? Kust sa tuled?*  
 'Where are you going? Where are you? Where do you come from?'  
*Koht, kuhu me jõudsime, kus me oleme, kust me tuleme*  
 'The place where we arrived, where we are, where we come from'
- (9) Loc/Indef  
*Kuskil haukus koer*  
 A dog was barking somewhere'  
*Tal oli kõikjal sõpru*  
 'He had friends everywhere'
- (10) Temp/Dem  
*Ta tuli siis, kui kõik oli juba läbi*  
 'He came when everything was already over'
- (11) Temp/Int-Rel  
*Millal me tõusma peame?*  
 'When must we get up?'  
*Täpne aeg, kunas rong saabub, pole teada*  
 'The exact time when the train arrives is unknown'
- (12) Temp/Inf  
*Ta ei tee seda kunagi*  
 'He will never do it'
- (13) Mod/Refl  
*Uks läks ise lahti*  
 'The door opened itself'
- (14) Mod/Rec  
*Nad vahetavad vastastikku komplimente*  
 'They exchange compliments with each other'
- (15) Mod/Dem  
*Pane käed nii!*  
 'Put your hands like this!'
- (16) Mod/Int-Rel  
*Kuidas te elate?*  
 'How are you?'
- (17) Mod/Inf  
*Tahaksin teda kuidagi aidata*  
 'I would like to help him somehow'
- (18) Quant/Dem

*Ei taha sellest kuuldagi, niivõrd on ära tüüdanud*

'I don't want to hear anything about it, it has become so boring'

(19) Quant/Int-Rel

*Maksin niipalju, kuipalju küsiti*

'I paid as much as I was asked'

(20) Quant/Inf

*Vesi polnud kuigi soe*

'The water was not very warm'

(21) Caus-Fin/Dem

*Ta hilines sellepärast, et auto oli rikkis*

'He was late because the car had broken down'

(22) Caus-Fin/Int-Rel

*Miks sa kiirustad?*

'Why are you in a hurry?'

(23) Caus-Fin/Inf

*Millegipärast ta ei tulnudki*

'For some reason he did not come'

(24) Conc/Dem

*See mahl on niigi magus, milleks veel suhkrut*

'The juice is sweet anyway, why do you need the sugar in addition'

## REFERENCES

**EKG I** = Mati Erelt, Reet Kasik, Helle Metslang, Henno Rajandi, Kristiina Ross, Henn Saari, Kaja Tael, Silvi Vare, Eesti keele grammatika I. Eesti Teaduste Akadeemia Keele ja Kirjanduse Instituut. Tallinn 1993. 26–31.

**Pool, Raili 1999**: About the use of different forms of the first and second person singular pronouns in Estonian cases. – Mati Erelt (ed.), Estonian: Typological Studies III. (Publications of the Department of Estonian of the University of Tartu 11.) Tartu 1999. 158–184.

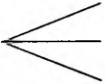
**Tauli, Valter 1972**: Eesti grammatika I. Hääliku-, vormi- ja sõnaõpetus. Uppsala. 66–67

**GRJ** = Грамматика русского языка. Москва 1952. 393.

## ASENDUSVORMIDEST EESTI KEELES

### Mati Erelt

Eesti keeles on provormidest olemas pronoomenid ja proadverbid, kuid puuduvad proverbid:

pronoomenid 

- prosubstantiivid: *mina, ise, kes* jne
- proadjektiivid: *niisugune, selline, milline* jne
- pronumeraalid: *mitu, mitmes* jne

proadverbid: *siin, nii* jne

**Pronoomenid** on eesti keele akadeemilises grammatikas (EKG I: 26–31) liigitatud kaheksasse liiki. Käesolevas kirjutises väidetakse, et tegelikult on liike vaid kuus. Eesti keeles puuduvad possessiiv- ja determinatiivpronoomenid. Pronoomenid *enese~enda* (gen) ja *oma*, samuti liitsõnalised *iseenese~iseenda* ja *omaenese~omaenda*, mida seni on käsitletud nii refleksiivpronoomenitena (*Mees tagus endale/omale vastu rinda*) kui ka possessiivpronoomenitena (*Mees läks oma/enda naisega tülli*), on tegelikult vaid refleksiivpronoomenid. Atribuutiivselt kasutatuna viitavad nad tegevussubjektile, mitte kõnesituatsiooni osalistele. Pronoomenit *enese~enda* on seni käsitletud vaegmuutelise, ilma nominatiivita pronoomenina, determinatiivpronoomenit *ise* aga täisparadigmaga pronoomenina, mille vormid alates genitiivist langevad kokku pronoomeni *enese~enda* vormidega (*Ta ise ei teadnud sellest midagi. See on tema enda asi*). Tegelikult võib ka apositiivselt kasutatud pronoomenit *ise* käsitleda refleksiivpronoomenina. Seega puhtsünkroonilisest aspektist vaadelduna on eesti keeles olemas refleksiivpronoomen *ise*, mis võib esineda erinevates süntaktilistes positsioonides – verbi laiendina, substantiivi atribuudina ning apositsioonina. See, millises käändes ta esineda saab, sõltub sellest, millises süntaktilises positsioonis ta parajasti on. Verbi laiendina ei saa tal olla nominatiivi, mis aga ei tähenda, et tal nominatiiv üldse puudub. Atribuudina ei saa ta üldse olla muudes käänetes kui genitiiv. Ainult apositsioonina on refleksiivpronoomen *ise* kasutatav kõigis käänetes. Eri süntaktilistel positsioonidel on erinevaid kasutuspiiranguid, kuid enesekohasus on ühendava tunnusena kõigil kolmel juhul olemas.

Täpselt samuti on ka vaid üks refleksiivpronoomen *oma*. Kui *ise* toimib primaarselt verbi laiendina ja sekundaarselt atribuudina või apositsioonina, siis *oma* toimib primaarselt atribuudina, sekundaarselt aga verbi laiendina ja apositsioonina.

Ülejäänud determinatiivpronoomenid, nagu *kõik*, *iga* jt on käsitletavat indefiniitpronoomenitena. Kuueliikmeline pronoomenite liigitus ei ole käesoleva kirjutise autori väljamõeldis. Sellise liigituse esitas Valter Tauli juba 1972. aastal. Paraku ta ei põhjendanud seda, mistõttu teised grammatikud seda liigitust ei aktsepteerinud.

**Proadverbe** enamikus eesti keele grammatikates omaette sõnaliigina käsitletud pole. Esimene grammatika, kus seda on tehtud, on EKG I (1995: 32). Mujal on neid peetud tavalisteks adverbideks, arvestamata nende asendusfunktsiooni. Kuid ka EKG libiseb proadverbidest üsna kergelt üle, nimetades vaid proadverbide mõningaid olulisemaid adverbiaalseid funktsioone, jättes aga andmata liigituse asendusfunktsiooni järgi.

Proadverbid toimivad ühelt poolt adverbidena, täites semantilist (aktantset) funktsiooni, teiselt poolt aga asesõnadena, täites asendusfunktsiooni. Proadverbide eripäraks on just see, et mõlemad funktsioonid realiseeruvad proadverbides leksikaalselt. Proadverbid on leksikaalsed vahendid teatud rollide täitmiseks, mida muidu peaks väljendama pronoomeni käändevormiga (*selleks*), pronoomeni ja adpositsiooni ühendiga (*selle jaoks*) või pronoomeni ja funktsioonisubstantiivi ühendiga (*sel viisil*).

Proadverbidel on vähem asendusfunktsioone kui pronoomenitel ja vähem semantilisi rolle kui päris adverbidel. Asendusaspektist vaadatuduna on eesti keeles viit liiki proadverbe: 1) refleksiivsed, 2) retsiiprooksed, 3) demonstratiivsed, 4) interrogatiiv-relatiivsed, 5) indefiniitsed. Adverbiaalsest aspektist liigituvad eesti proadverbid kuueks tüübiks: 1) lokaalsed, 2) temporaalsed, 3) modaalsed (viisi ja seisundi), 4) kvantitatiivsed (hulga ja määra), 5) kausaalsed-finaalsed ja 6) kontsessiivsed proadverbid. Potentsiaalseid kombinatsioone on  $5 \times 6 = 30$ , tegelikke 21. Neid kombinatsioone kirjeldab järgmine tabel (igast liigist on näidetena esitatud üks või mõni tüüpilisem proadverb).

Tabel 1. Proadverbid eesti keeles

	Refl	Rec	Dem	Int-Rel	Indef
Loc			sinna seal sealt siia siin siit	kuhu kus kust	kuskile kuskil kuskilt kõikjale kõikjal kõikjalt
Temp			siis tollal	millal kunas kuniks	kunagi millalgi iaal(gi) alati



	Refl	Rec	Dem	Int-Rel	Indef
<b>Mod</b>	ise	omavahel vastastikku isekeskis	nii nõnda sedasi	kuidas	kuidagi teisiti igati kõigiti
<b>Quant</b>			niivõrd	kuivõrd	kuigi kuigivõrd
<b>Caus- Fin</b>			sellepärast seetõttu	miks mispärast milleks	millegipärast
<b>Conc</b>			niigi nagunii sellegipoolest		

## Näiteid:

- (7) Loc/Dem Lähem sinna, olen seal, tulen sealt
- (8) Loc/Int-Rel Kuhu lähed? Kus sa oled? Kust sa tuled?  
Koht, kuhu me jõudsimme, kus me oleme, kust me tuleme
- (9) Loc/Indef Kuskil haukus koer  
Tal oli kõikjal sõpru
- (10) Temp/Dem Ta tuli siis, kui kõik oli juba läbi
- (11) Temp/Int-Rel Millal me tõusma peame?  
Täpne aeg, kunas rong saabub, pole teada
- (12) Temp/Inf Ta ei tee seda kunagi
- (13) Mod/Refl Uks läks ise lahti
- (14) Mod/Rec Nad vahetavad vastastikku komplimente
- (15) Mod/Dem Pane käed nii!
- (16) Mod/Int-Rel Kuidas te elate?
- (17) Mod/Inf Tahaksin teda kuidagi aidata
- (18) Quant/Dem Ei taha sellest kuuldagi, niivõrd on ära tüüdanud
- (19) Quant/Int-Rel Maksin niipalju, kuipalju küsiti
- (20) Quant/Inf Vesi polnud kuigi soe
- (21) Caus-Fin/Dem Ta hilines sellepärast, et auto oli rikkis
- (22) Caus-Fin/Int-Rel Miks sa kiirustad?
- (23) Caus-Fin/Inf Millegipärast ta ei tulnudki
- (24) Conc/Dem See mahl on niigi magus, milleks veel suhkrut

# GRAMMATICALIZATION OF ADPOSITIONS IN OLD LITERARY ESTONIAN

Küllli Habicht

*University of Tartu*

## Grammaticalization and the genesis of adpositions

Grammaticalization is the constant process of generating grammatical categories or, in a narrower sense, the process of changing lexical units into grammatical ones. It is part of the natural system of language – a lexical morpheme acquires the status of a grammatical morpheme. There occurs a process of generalized semantic opposition and its realization. Grammar shapes standard means for the expression and association of generalized meanings (Metslang 1994: 13), thus, forms that express new grammatical meanings come into existence. Grammaticalization is regarded as a universal process in the world's languages, which can be associated with various grammatical functions (Heine, Claudi, Hünemeyer 1991: 2).

The formation of adpositions and adposition-based case endings has been regarded as a universal example of grammaticalization. New adpositions develop mainly in two ways – from substantives and verbs (e.g. Heine, Claudi, Hünemeyer 1991). The Estonian adpositions are mostly fossilized locative forms. According to Lehmann, the genesis of adpositions goes through the following process: relational substantive → secondary adposition → primary adposition → agglutinative affix → fusional affix (Lehmann 1985: 304). Estonian adpositions originated from substantives with varying degrees of abstraction, more rarely from other parts of speech, e.g. verbs (*hoolimata* 'despite' and *vaatamata* 'despite').

Grammaticalization has often been regarded as an opposite to lexicalization. In the first case a word is subject to semantic change and moves from an open class to a closed class; in the

second case the direction of the development is opposite – some word form acquires a new meaning, thus leaving the inflectional paradigm of a full word. It is also possible to regard these phenomena as interrelated to each other and to claim that each semantic change towards more abstract content need not be directly related to grammaticalization. It is especially noticeable in those chains where an adverb serves as an intermediate step between the genesis of a postposition and its spread. At first the inflectional form of the noun gives rise to a lexical unit with a new meaning, which acquires a new grammatical function only after an extension into a certain syntactic environment.

Tiina Onikki, too, draws attention to this development in her article about locative expressions of state. She treats lexicalization and grammaticalization as different parts of the same development chain (spiral). On the one hand, a productive inflectional form is lexicalized and creates set phrases with specific meanings, on the other hand, the morphosyntactic and semantic pattern of a new expression type is born, which can be interpreted as an independent grammatical category. The new construction is generalized on the basis of analogy according to the pattern created by individual innovations (Onikki 1997).

It is likely that the purpose of the process may be important. In case new grammatical morphemes are born on the basis of productive rules, it is grammaticalization. By contrast, lexicalization implies the genesis of independent lexical units. Both processes trigger semantic and phonological simplification, and both may originate from the same source. Thus the crucial difference between these processes does not lie in their basic forms or in the occurring developments but in the final destination, which is either grammar or lexicon (Diewald 1997: 73).

Grammaticalization is a diachronic process – a logical chain with links that follow each other in time. However, at the same moment in a language there may exist usages of previous and following stages. Therefore, the grammaticalization theory enables us to explain both the diachronic development of the meaning and function of a grammatical device as well as synchronic variation (Metslang 1997: 228).

It is generally agreed that the grammaticalization process is launched by semantic change. The meaning of a linguistic element acquires a new interpretation in a certain context (reanalysis), which means positioning itself into new structures (restructuring). The subsequent generalization and spread brings about a change in the linguistic system. It is the context, which in a narrower sense could be defined as a certain linguistic construction, which causes the reanalysis of the linguistic elements. The change may originate from a certain combination. It has also been thought that metaphoric and metonymic (i.e. figurative) uses are the central forces that trigger grammaticalization processes (Laitinen, Lehtinen 1997: 12–13). First and foremost reanalysis has been regarded as the mechanism that launches the change. As a result, the semantic, syntactic, or morphological usage norm of the word that is involved in the process undergoes a change. Analogy changes only the external usage, whereby the norm spreads either in the linguistic system or in the speech community (Hopper, Traugott 1993: 32).

In the course of the grammaticalization of substantive forms, extension of meaning of a substantive in a locative case usually adds a grammatical function, which is similar to the function of typical locative cases, although it takes an analytic form. Such a construction fulfils grammatical functions in morphosyntactic structures. The typological study of various languages has revealed a number of general tendencies in grammaticalization. Thus, the meanings that are important from the cognitive or social point of view tend to undergo grammaticalization. Location or space is a domain that is regarded as the first step towards grammaticalization; it is followed by temporal and manner relations, which have a more abstract character. A functional gap may be absent in the location where a new grammatical means is born. A word will be more bound than before both paradigmatically and syntagmatically, it will place itself in a more clear-cut morphosyntactic paradigm with fewer members while the word's syntactic flexibility and independence will decrease. The meaning will become more specific – the form that removed itself from the paradigm of a full declinable word will acquire an

individual meaning that has usually a more general meaning than the base word.

The Finnish linguist Esa Itkonen in his article on the concept of grammaticalization supports the analogy-based theory of grammaticalization and language change. In his words nothing comes into existence from nowhere, and it is impossible to imagine that language could change fully independently of its structure. Thus, the rudiments of structural change are present in the structure itself (Itkonen 1998: 34).

Any existing adposition may provide the pattern of a new adposition, which means that once the grammaticalization process has started, analogy will guide the development of new (metaphoric) meanings and the new grammatical function. The users will perceive reanalysis (restructuring) only after the change is generalized, and the new construction type will increasingly begin to manifest itself (Ojutkangas 1998: 69).

### **On typology**

From the point of view of morphological typology, Estonian, Livonian, and Lapp form an exceptional group among Finno-Ugric languages with a higher degree of inflection, which is accompanied by higher analyticity, which is a sign of an isolating language as the next stage (Metslang 1994: 10). Estonian is rich in postpositions – the use of postpositions has been regarded as a Finno-Ugric feature in Estonian, which contrasts with Indo-European languages that use mostly prepositions. In the diachronic perspective adpositions form an open class, where new members are added by analogy with the existing adpositions. Adpositions overlap to some extent with substantives; most Estonian adpositions originated from the case-marked forms of substantives. They also overlap with adverbs – the same word may occur as an adverb or an adposition depending on whether it occurs alone or with a noun, cf. *ta seisab kõrval* ‘he is standing next’ and *ta seisab minu kõrval* ‘he is standing next to me’

Estonian grammar defines the adposition as an uninflected word, which belongs to the substantive in the sentence and adds roughly similar meanings as case markers. However, the meaning of adpositions is more specific and clearer than that of case forms

(e.g. EKG 1995: 33–34) because the semantics of adpositions has retained part of the meaning of the originally full noun.

A prototypical adposition is not independent syntactically and semantically, but it forms an adpositional phrase together with the noun. After the omission of a genitival or partitive noun the expression is syntactically and semantically unacceptable. The more prototypical the adposition, the more fossilized is its form and the more distinguishable it is from the base word (Jaakola 1997: 123).

During its initial period Written Estonian was for a long time in the sphere of strong German influence, which is manifested also in several adpositional constructions under discussion. It is understandable that in the earlier translation tradition, where morpheme-to-morpheme translation was standard practice, it was easier to translate the German prepositional constructions into the Estonian postpositional constructions than to replace them by synthetic forms.

As German is typologically an inflecting and analytic language, but Estonian is mainly an agglutinating-inflecting language, then ‘adpositionalization’ that occurs in the old literary language could be partly explained by the impact of the grammar of a typologically different language. It is confirmed by a number of cases, where the German and Estonian nouns seem to have undergone a similar process of adpositionalization. Many German prepositions have also been derived from nouns. Example include *anfangs* ‘in the beginning of’, *mittels* ‘by means of’, *kraft* ‘by force of’, *bezüglich* ‘in regard to’ *dank* ‘thanks to’ *laut* ‘according to’ *zeit* ‘in the course of’ *anstelle* ‘in place of’ and *aufgrund* ‘on the basis of’ (Duden 1997: 226; Diewald 1997).

Diewald, who has studied the grammaticalization of some German prepositions, has emphasized that adpositions include elements with highly different degrees of grammaticalization (Diewald 1997: 65). This claim is also valid for Estonian at any moment.

From the point of view of language history the relations between the postposition and its head has in many cases developed from the relation between the nominative head and its

nominal attribute in the genitive (Laanest 1975: 192), e.g. *lapse kõrv* 'the child's ear' – *lapse kõrval* 'next to the child'

### **The purpose and sources of the study<sup>1</sup>**

The adposition as a part of speech includes material with a rather varying degree of grammaticalization. The present study proceeded from some adpositions, still in the initial phase of development, which caught my attention while studying Old Literary Estonian, and from the comitative case, the genesis of which can be observed on the basis of the survived texts of Written Estonian. Therefore, the following selection is semantically rather varied and uneven from the standpoint of the genetic process. However, it should nevertheless shed light on the development of adpositions and whether these processes have taken place in a typical manner or whether there is something unusual in them. At the same time the theory of grammaticalization enables us to clarify the distribution of adposition – to distinguish them on the basis of their degree of grammaticalization.

Ascribing parts-of-speech labels to Estonian words is often problematic because one has to take into account both morphological and syntactic criteria. It is especially difficult to distinguish adpositions from nouns because the former are often fossilized forms of the locative cases of the latter (it is a characteristic feature of Balto-Finnic languages), being at the same used as full words. Therefore, the present article focuses on some typical pairs of full words and functional words in Old Literary Estonian. An attempt will be made to characterize on the basis of individual examples the stages of adpositionalization from a full substantive into an adposition and the possible reasons, which the adpositions in Modern Estonian perhaps do not manifest anymore.

The study tries to characterize the earliest stage in the development of Written Estonian on the basis of examples from Old Literary Estonian. Its purpose is to establish regularities in the

---

<sup>1</sup> I would like to express my gratitude to Prof. Mati Ereht for valuable advice

adpositionalization of lexical words and to show on the basis of the development of some adpositions the typological impact of German, which shifted Estonian towards analyticity during the beginning period of Written Estonian. The study will focus on the analysis of adpositions that originated from individual nouns. Below we will deal only with problematic cases. We will not discuss the numerous adpositions had already reached the stage of primary adpositions by the period of Old Literary Estonian (e.g. *alla* 'to under' *all* 'under', *alt* 'from under'; *sisse* 'into' *sees* 'inside', *seest* 'from inside', or *peale* 'onto' *peal* 'on', *pealt* 'from on' It goes without saying that the grammatical competence of the modern language affects the choice and analysis.

We will proceed from the understanding that adpositionalization is a step-by-step and constant process that covers long time periods, whereas different words may be at a certain moment in different stages of this development. Figuratively speaking, they may be located in various places on a scale that begins with typical full words and ends with relational words that have acquired a generalized meaning and specific usage. The study proceeds from the widespread view that new grammatical devices arise at the cost of lexical words if the latter begin to be used in a certain environment. It is important that the use of the construction become automatic and formulaic, which creates the conditions for the extension of the use of the word into other analogic associations. An attempt will be made to show that the characteristic features of Estonian were largely present already in the texts of Old Literary Estonian that were written almost four centuries ago.

The material was collected from the oldest Estonian written texts. I worked through all the existing Estonian texts from the 16<sup>th</sup> century (Ehasalu *et al* 1997), the original manuscript of 39 sermons by Georg Müller (1600–1606), and two ecclesiastical books by Heinrich Stahl. The latter included *Hand- vnd Haußbuches Für die Pfarherren/ vnd Haußväter Ehstnischen Fürstenthumbs* .. (henceforth, *HH*), parts I–IV 1632–1638 and *Leyen Spiegel* (henceforth, *LS*), parts I–II 1641 and 1649. The examples are followed by the name of the source or its abbreviation and page number in brackets. In the case of the



sermons by Müller the first number refers to the number of the sermon and the second number to the page number; in the case of *Hand- vnd Haußbuch* the part number is provided, too.

As the material has mostly a translational character, then one might assume that the text that was written by Germans contains German-like constructions. The sermons by Müller include German-language passages and comments. The works by Stahl include parallel texts in Estonian and German, which is brought out in the examples. The direction of translation is not always unambiguously clear because there is some reason to think that the authors that wrote down Old Estonian used handwritten texts. Judging by Stahl's sermons, it seems that in most cases, but not always, the translation proceeded from German. Thus the following overview reflects the earliest development stage of Literary Estonian over the course of about 150 years.

## 1. Noun → adposition

### 1.1. *ase* → *asemele*, *asemel*

In the case of this word group texts of Old Literary Estonian reveal a development from a full noun into an adposition. It seems that the corresponding German grammatical constructions may have affected the course of development. In most cases *ase* 'place' occurs as a noun (the German equivalents are *Statt*, *Stette*, *Ort*, and *Raum*), being used in a more general meaning of location, e.g.

*Kuy tema woyra rachwa secka, nick wöira aszemelle .. tulleb*  
(Müller 23/8)

'When he comes among a strange people and instead of a stranger'

*Iummal istup sel keicke körgkembal assemell*

*Gott sitzet an der höchsten Stät!* (Stahl, LS, 147)

'The Lord is sitting in the highest place'

The frequent adessive form of this word with a relatively abstract meaning of location has become generalized and expresses a more abstract substitutional relation. In the following examples it is on its way towards adpositionalization, e.g.

*Jesus Christus Iummala poick/ meije **assemel** on tulnut/  
 Jesus Christus Gottes Sohn/ an vnser **statt** ist kommen/* (Stahl,  
 HH II, 36)

‘Jesus Christ, Son of God, has come in place of us’

The previous example sentence could be interpreted also in the literal sense “above us, to our place” If one follows in making a distinction between a full noun and structural adposition the principle that an adposition cannot have any extensions, then the previous examples are closer to nouns than to adpositions (the extensions can be suggested). They would be good examples to illustrate the beginning of transfer from one variety to the other. It is accompanied by the generalization of meaning – *ase* as a concrete place has become an abstract generalization. Judging by meaning, one could say that it is an intermediate step between a full word and an auxiliary word. The position of the word in the sentence is important, too – its position is next to the pro-substantive, which is a typical pattern of adpositionalization: the literal sense *meie ase* ‘our location’ → a frequent adessive form → extension of meaning → the construction *genitival noun + asemel* (adpositional construction).

However, the following examples reflect a further development of adpositionalization:

*Sysz tullewat küll 10 willetzust se vche **aszemel*** (Müller 25/5)  
 ‘Then ten miseries will come instead of the one’

*Eth nemmat mitte kahs tullewat sesinnase wallo **assemel**  
 Auff das sie nicht auch kommen **an diesen ort** der qual* (Stahl,  
 HH III, 93)

‘That they not also come instead of this pain’

*eth se leickatusse **assemel** se pöha wedde ristminne on  
 sissesehtut/*

*weil **an stat** der Beschneidung die Heilige Wasser Tauffe  
 eingesetzt ist/* (Stahl, LS, 114)

‘that instead of the surgery the baptism by holy water has  
 been applied’

These examples reveal that the word *ase* does not refer to a place anymore; it is used to mark an abstract substitutional relation, whereas one can say the same about the German equivalent

sentences. It seems that in all the cases the adposition is in the stage of a secondary adposition. Secondary adpositions are mostly polysyllabic, they are semantically more complex, paradigmatically varied, and belong rather loosely to semantic fields. They are mostly used in association with a certain semantic domain (Diewald 1997).

## 1.2. *hea* → *heaks*

Already the oldest records of Literary Estonian reveal the fossilized translative form of the noun *hea*, which has passed into formulaic use (it can function both as an adjective and a substantive; the latter suits better for the present purpose). The previous noun, though, in the adessive in the text of a court oath that was written down around 1600, e.g.

*Se selgke n. Puchta töβidusse kellekil hehx eck kahjux, welja reckima tahhan* (court oath, about 1600)

‘I wish to speak the clear pure truth for the benefit or to the detriment of someone’

The previous example reflects the intermediate stage on the way to adpositionalization. The word form *heaks* ‘for the benefit of’ acts here already as the marker of the beneficiary role, but the case of the pronoun reserves the possibility to interpret *hea* and *kahju* still as independent nouns although it is a formulaic combination.

In the sermons by Müller *heaks* already occurs more extensively as an adposition (45 instances), mostly in the combinations *meie heaks* ‘for our benefit’ *minu heaks* ‘for my benefit’, and *sinu heaks* ‘for your benefit’ e.g.

*Ninck mea tæma teeb, se teeb tæma keick meddi heex* (Müller 36/14)

And what he does, he does everything for our benefit’

The works by Stahl also reveal repeatedly *heaks* as an adposition, above all in connection with the personal pronouns *minu* ‘my’ *meie* ‘our’ *sinu* ‘your’, *teie* ‘your pl.’ *nende* ‘their’, and the relative pronoun *kelle* ‘whose’ (“Leyen Spiegel” and “Hand- vnd Haußbuch” include 59 such occurrences.) The combination of *heaks* and a noun is rarer, e.g.

*Ananias mühs omma mah se Risti Kogkodusse hehx/  
Ananias verkauffte seine Güter/ der Christlichen Gemeine zu  
gut/* (Stahl, LS, 693)

‘Ananias sold his land for the benefit of the Christian congregation’

or

*eth nende wagkada hehx keick peap sündima  
Das den frommen muß alles zum besten dienen* (Stahl, LS, 301)

‘that everything must happen for the benefit of the pious’

The German equivalents suggest that the Estonian construction may be a translation loan. The adpositionalization of the substantival *hea* clearly shows how the use of the construction can become automatic (with regard to pronouns) and extend to other associations as well (the association with nouns must have been a later development). The frequent use of certain fixed combinations is also conditioned by the specific nature of the content of the text, which proceeds from the figurative system of religious works.

### 1.3. *kesk* → *kesk*, *keskel*

The noun *kesk*, which was used in Older Estonian, has provided adpositions first and foremost on the basis of its locative forms. Of them *keskele* ‘into the middle of’ and *keskel* ‘in the middle of’ are widely used now as well, but the noun *kesk* is not known anymore. In the texts of the old literary language *kesk* is used as a noun in the compounds *kesköö* ‘midnight’ *keskpaik* ‘midsection’ and *kesklõuna* ‘midday’. However, its use as an adposition is much more widespread, e.g.

*On tæma sen Taiwa siddes, kesck nende Englide* (Müller 6/9)  
‘He is in Heaven, among the angels’

It is noteworthy that in the sermons by Georg Müller the preposition *kesk* occurs in the nominative and not in the adessive, as sometime later in the works by Stahl. Judging by the existing written sources, the difference in grammatical usage reveals that the use of the fossilized adessive form as an adposition is a somewhat later development.

*echk küll se Ilm hucka lehx/ ninck needt Mehjet kesckil Merre  
sisse wajuxit  
wenn gleich die Welt vntergenge/ vnd die Berge mitten ins  
Meer süncken (Stahl, LS, 512)  
'if the world perished and the mountains sank into the sea'*

It is of interest to note that *keskel* occurs in the old literary language mostly as a preposition unlike the present-day postposition, which could characterize the earlier use of this adposition (cf. Finnish). However, the German language may have influenced the spread of this usage tradition in the older literary language. The use of *keskel* as a postposition, which is at present common, is rarer, e.g.

*Tullep Iesus/ kus needt uxet luckus ollit/ ninck astup nende  
kesckel/  
Kompt Iesus/ da die Thür verschlossen waren/ vnd tritt  
mitten ein/ (Stahl, LS, 430)  
'Jesus comes, where the doors were locked, and steps among  
them'*

Here, too, one might suspect that the use of the adposition may have become fixed due to the German influence in the translation (cf. the German equivalent *mitten*), whereas the corresponding adverb seems have served as the intermediate stage in Estonian.

#### 1.4. *komme* → *kombel*

It seems that the construction *adessive* + *kombel*, as well as the *genitive* + *kombel*, are the most widespread uses of the word *komme* 'custom, habit' both in the old as well as the contemporary literary language. The word *komme* occurs most commonly in the *adessive*, thus having been adpositionalized and expressing manner. Here, too, adpositionalization must have taken place because of frequent use and extension to the meaning of manner. The word *komme* is, in fact, an abstract noun that does not allow any substantial transfer of meaning. Therefore, in many cases one cannot claim for sure that it is a fully grammaticalized word. Most contexts allow interpreting this form also in the nominal sense that denotes manner. Only some examples indicate that during the

period of the oldest literary language the word was undergoing the first stage of adpositionalization.

The noun *kommel* is absent in the oldest Estonian-language sources. However, two instances of the word as adpositions in formulaic combinations were attested – *se kommel* ‘in this manner’ (Awerbach 1589) and *kahjolicul kommel* ‘in a harmful manner’ (cort oath, 1600). In the sermons by Müller, too, the adpositionalizing adessive *kommel* occurs more frequently (164 instances) than any other form. In most instances *kommel* still agrees with the preceding word and forms such collocational pairs as *sel kommel* ‘in this way’, *mõnesarnasel kommel* ‘in some way’ *imetabasel kommel* ‘miraculously’ and *mingisarnasel kommel* ‘somehow’. According to the modern grammatical tradition, one cannot regard them as adpositions as yet, but they could have provided the basis for the automatic use of this form.

In the texts by Heinrich Stahl, too, the most frequent construction is *adessive + kommel*. It forms constructions that resemble fixed phrases, but there is no adpositionalization as yet, e.g.

*Eth nühd/ armas welli/ sel sammal kommel se Kurrat sünno  
wasto tousnut/ ninck perrenkiusamisse töstnut on/Wann dann  
nun/ lieber Bruder/ gleiches falls der Teuffel wider dich sich  
gesetzt vnd Verfolgung erreget hat/ (Stahl, HH IV 212)*

‘that now, dear brother, the Devil may have risen against you  
in the same way and started to persecute you’

The following examples show more clearly that *kommel* was undergoing adpositionalization by the first half of the 17<sup>th</sup> century evidently by analogy with the above-mentioned fixed phrases:

*Nente kolme eszimesze Versi siddes tuñistame meye kaib-  
misze kommel, meddy suhre hedda (Müller 20/1)*

‘In the three first verses we witness our great misery in the  
manner of a complaint’

*Se pahharet omma kawwalusse kahs/ Mördri kommel pimme-  
dusses hulckup/*

*Der Teuffel gern mit seiner List/ als Mörder pflegen/ im  
finstern ist (Stahl, HH III, 245)*

'This Devil with its cunning lurks in darkness like a murderer'

In the previous example *kombel* has been rendered in the German text by means of the comparative conjunction *als*. This suggests that the person who wrote down the text were familiar with the corresponding grammatical construction in Estonian. Otherwise they would have translated it literally.

Generally speaking, the German equivalents to the constructions including the word *kombel*, such as *eben also*, *gleich also*, *desselbigen gleichen*, *dessgleichen*, indicate that an attempt was made to render a relationship that is expressed by non-inflectional words.

### 1.5. *kõrv* → *kõrval*

The noun *kõrv* 'ear' which denotes a body part, has contributed the auxiliary words *kõrvale* 'to next to' *kõrval* 'next to' and *kõrvalt* 'from next to' to literary Estonian. The words are absent in texts dating from the 16<sup>th</sup> century. However, the first half of the 17<sup>th</sup> century already reveals some uses of *kõrval*. The sermons by Georg Müller reveal two instances where it occurs as an adposition, e.g.

*waidt syszkit keub se Issa oma Lapse kõrwal* (Müller 29/5).  
'but Father nevertheless walks next to his child'

This adposition is rather rare also in the books by Heinrich Stahl, e.g.

*erranis oppewat temma kõrwal muhdt töiset wöhrat Teedt*  
(Stahl, LS, 588)  
'especially others learn the way beside him'

The preposition in the German equivalent sentence indicates that *kõrval* must have taken root as an auxiliary in Estonian, and it has not been made up in the course of translation. Its rare use may have been caused by the fact that the German authors who wrote down the Estonian language may have not been familiar with some words known in popular language. Therefore, they may have substituted such words by constructions that were less known in Estonian.

### 1.6. käsi → kätte, käes, käest

It is likely that the internal locative case forms of the noun *käsi* ‘hand, arm’ moved away from the paradigm due to the reinforcing influence of figurative expressions (*Kristuse käest midagi saama* ‘to receive something from Christ’ *oma käest verd andma* ‘to donate one’s own blood’). These forms express ownership in more general terms transfer into the ownership of someone (*kätte* ‘lit. into the hand’), being in the ownership of someone (*käes* ‘lit. in the hand’), and transfer from the ownership of someone (*käest* ‘lit. from the hand’). Judging by the texts of the older literary language *kätte* had not fully adpositionalized as yet. According to Lehmann, these would be secondary adpositions, which are semantically transparent, only to some extent more abstract than the base word and appropriate only in a semantic field that is associated with possessiveness. The older literary language reveals some contexts, where the word has moved away from the original meaning, but the link with the meaning of the noun is still existent, e.g.

*ninck næmat ninda nente Korgke pappide kette ülleandnuth*  
(Müller 20/73)

‘and they had thus handed it over to the high priests’

*Genitive+kätte* occurs here in a typical position of the adposition as a dependent extension of the verb *üle andma* ‘to hand over’, and the context enables us to conclude a more general transfer to ownership.

The sermons by Müller reveal seven instances where *käes* has undergone adpositionalization, e.g.

*Sesama Aick (A.R.) on io nüith parrahellis meddy käes*  
(Müller 9/3)

‘The same time is now in our hands’

The figurative expression *aeg on käes* ‘time is at hand’ clearly indicates moving away from the original meaning of the body part and the beginning of grammaticalization. By the way it points to the fact that the transfer has been extensive, involving time as



a very abstract category. The adpositional use is confirmed by the insertion of the genitival personal pronoun *meie* 'our'

The works by Heinrich Stahl also reveal that the figurative mode of expression that is characteristic of church language may have contributed to the generalization of the case forms of *käsi*, e.g.

*Surm/ Patt/ Kurrat/ Pörgkohaud/ Ello/ ninck Arm/ keick  
Iesusse Keddēs on*

*Tod/ Sünd/ Teuffel/ Hell/ leben/ vnd Gnad/ alles Iesus in  
Händen hat* (Stahl, LS, 411)

'Death, sin, devil, Hell, life and grace, everything is in the hands of Jesus'

The figurative generalization has led the body part in the state of an adposition that expresses ownership. In this case both German and Estonian have followed the same logic.

There are five instances of the adpositional *käest* in the sermons by Müller, where it expresses a relation of generalized ownership, e.g.

*tahab se Kuñinga Rickus sinu Kæddest errawotta* (Müller 15/48)

'he wants to take the kingdom away from you'

In the previous example *genitive + käest* is a dependent extension of the verb *ära võtma* 'to take away'

It seems that the semantic generalization of the forms of *käsi* may have been reinforced by the corresponding figurative constructions in German. It is especially noticeable in the case of the adposition *käest*, where most contexts in the old literary language are ambiguous, i.e. *käsi* 'hand, arm' can be interpreted as a noun as well, e.g.

*Ollemme meije se heh sahnut/ sest Issanda kehjest/*

*Haben wir das gute empfangen von der Hand deß Herren*  
(Stahl, HH II, 165)

'We have received the good from the Lord'

Here we can see a constructional translation equivalent in Estonian, where *käest* is nevertheless manifested in the

generalized meaning. The impact of the German-language construction could be seen in the following example as well:

*eth keddakit meid sest lummala Armust erralahutama/ echk  
Christusse kehjest errakisckma sahp  
dz vns nichtes von der Liebe Gottes scheiden/ noch jemand  
auß Christi Händen reissen werde* (Stahl, LS, 553)  
'that someone will separate us from God's Grace or tear us  
away from the hand of Christ'

When we examine the preceding set of words, we can see that *kätte* offers the largest number of clear instances of adpositionalization. There are fewer examples of the adpositions *käes* and *käest*. At the same time, in contemporary Estonian the entire group of adpositions has not reached beyond the stage of secondary adpositions. One could possibly explain it by the fact that the noun *käsi* 'hand, arm' is frequently used as a full word. However, in order to reach the stage of a primary adposition, the relationship must fade, and the meaning must become opaque with regard to the base word.

### 1.7. *naga* → *najal*

Of this chain, the substantive *naga* was not used in the sense 'support' anymore in the texts of the older literary language. However, there was a single occurrence of the adposition *najal* 'against, on' in the catechism by Wanradt-Koell:

*Sedda piddat teye tegkeme minu najal*  
'You are supposed to do it with my help'

Unfortunately, the two last letters of *najal* in this example sentence have perished together with a page in the book and have been reconstructed later. Therefore, we cannot fully confirm its actual use.

### 1.8. *nimi* → *nimel*

The noun *nimi* 'name' is rather common in texts of the older literary language while there are also a number of fossilized multi-word constructions, such as *nime juures* 'by (one's) name' *nime*

sees 'in (your) name' *nime pärast* 'because of (your) name' and *nime sisse* 'into the name' This frequently used noun witnessed at the beginning of the literary language the grammaticalization of its adessive form *nimel* in the construction *genitive + nimel* 'in the name of, on behalf of' which occurs mostly in the collocation *Jumala nimel* 'in the name of the Lord' e.g.

*Sesama taha meye nuith kz Jumala Nymel teha* (Müller 13/22)

'Now we want to do the same in the name of the Lord'

The genesis of this adposition could be reinforced by the German-language biblical expression because in the older literary language *nimel* occurs mostly in the combinations *Jumala nimel* 'in the name of God' or *isa, poja ning püha vaimu nimel* 'in the name of Father, Son, and the Holy Spirit' e.g.

*ninck Ristket nemmat/ se Issa/ se Poja/ ninck se pöha Waimo nimmel*

*Vnd Täuffet sie im Namen des Vaters/ vnd des Sohns/ vnd deß heiligen Geistes/* (Stahl, HH I, 40)

'and christen them in the name of Father, Son, and the Holy Spirit'

Later *nimel* may have extended by analogy from similar contexts to other constructions. The German equivalent *im Namen* is a sign of a possible translation loan.

### 1.9. *pool* → *poole*, *pool*, *poolt*; *poolest*

The word *pool* 'half; side' occurs in the oldest literary language as a noun in the fossilized combinations *kõiges pooles* 'everywhere' *kõigest poolest* 'from everywhere', and *omalt poolelt* 'on my behalf' The first occurrence of *poole* can be found in the catechism by Wannradt-Koell of 1535, e.g.

*eth meye se Sacrament wöyme önsast medy pole wotta*  
'that we can take a half of the Sacrament'

Such an adpositional use of *poole* is quite expected. It must have developed from the nominal construction *meie pool* 'our side' Such combination-like uses are common also in the sermons by

Müller, where *poole* occurs most frequently in the phrase *end kellegi poole käänama* or *pöörama* ‘to turn oneself in someone’s direction’ e.g.

*Iss: Iumall mina ollen üx Pattune Inimene, ninck tahan hend hælmelell sinu Armu pohle kænda* (Müller 16/6)

‘Lord, I’m a sinful person, and I want to turn myself willingly in the direction of your Grace’

Müller’s texts include *pool* also as a noun in the compounds *meespool* ‘male sex’ and *naispool* ‘female sex’. However, his sermons do not include the adposition *poolt*.

The works by Heinrich Stahl also reveal that *poole* in the meaning of direction is the most frequent adposition, e.g.

*ninck sahtas se igkawesse ello ninck Taiwalicko selgkuse pohle:*

*vnd brechte zu dem Ewigen Leben vnd Himlischer Klarheit:*  
(Stahl, LS, 226)

‘and sent into the eternal life and towards heavenly clarity’

The German equivalent suggests that this adpositional use had already taken root in Estonian because Stahl replaces the German preposition by the Estonian postposition and changes the word order.

Some interesting and unlikely uses can be found in the ecclesiastical books by Stahl, where the adposition *poolt* was used thrice and *poolest* twice, e.g.

*sinna peat münno suhst se sanna kuhlema/ ninck nemmat münno pohlt mainitzema*

*du solt auß meinem Munde das Wort hören/ vnd sie von meinetwegen warnen* (Stahl, LS, 52)

‘you must listen to the words from my mouth, and they must be warned on my account’

It seems that the adposition *poolest*, which is rare in the works by Stahl, is used in the meaning of cause, e.g.

*münno Orjat peawat röhmsa süddame kahs heh meele pohlest ichkma/*

*meine Knechte sollen für gutem Muth jauchtzen* (Stahl, LS, 472)

‘My inferiors must shout for their joyful heart and good spirit’

### 1.10. *pära* → *pärra*, *päras*; *pärast*; *pärale*, *päral*; *päralt*

The word *pära* did not occur anymore as a noun in the meaning ‘the last or hind part of sth; leftover, remains’ in the documented sources of the oldest literary language. However, the adpositions that developed from this noun form a rather large group of mostly archaic words that have branched into various senses. Only the postposition *pärast* ‘after, because’ is used in contemporary Literary Estonian.

The first occurrence of the adposition derived from the noun *pära* dates back to about 1510. It is used in the spatial sense ‘after’ in a collection of sermons:

*Myna /./ tha syno pära tulla*

‘I want to come after you’

The catechism by Wanradt-Koell and a letter dating from about 1535 reveal also the archaic adposition *päras* in a more abstract sense ‘after, according to’ e.g.

*Ninck seel siddes tema Jumala sanna peresz laszke löwdade*  
(Wanradt-Koell 1535)

And in there let him find in accordance with Lord’s words’

The adposition *pärast* in the causal sense occurs three times in the catechism by Wanradt-Koell, once in a text by L.Boierus from about 1587, and in a court oath dating from about 1600, e.g.

*Sen egkewen elon ninck sen Jumala rycküden perest*  
(Wanradt-Koell 1535)

‘Because of this eternal life and this God’s Kingdom’

In the longer texts in the older literary language *pärast* is already a frequent adposition. It occurs in its causal sense in the sermons by Müller. In the works by Heinrich Stahl, too, *pärast* is a common adposition, having extended from the common meaning ‘because of’ to such more figurative meanings as ‘according to; for’ e.g.

*Rahwo ninck röhmo kahs lehn minna erra/ Jumjala tacht-  
misse perrast/*

*Mit Fried vnd Frewd ich fahr dahin/ in Gottes Willen/ (Stahl,  
HH II, 163)*

‘I will leave with peace and joy, according to God’s will’

*Issand kus sinna münnul tahaxit münno denistuse perrast  
maxma/*

*Herr wenn du mir woltest nach meinem verdienst vergelten/  
(Stahl, HH IV 191)*

‘Lord, where you would like to reward my service’

In addition to the adposition *pärast*, the authors dating from the beginning of the 17<sup>th</sup> century use extensively the adpositions *pärale* and *päral*, which were grammaticalized from the external locative cases. Both of them mean ‘according to, after, for’, which was abstracted from the sense ‘locality’ in the meaning of a sensuous process, e.g.

*meije perrale olli temmale suhr igkaw/*

*nach vns thet ihn verlangen/ (Stahl, HH II, 28)*

‘he was longing for us’

*Rahel nuttis omma lapsede perral/*

*Rahel beweinet ihre Kinder/ (Stahl, HH III, 165)*

‘Rachel wept for her children’

The adposition *päral* may also express possession, belonging to someone, which is also a rather widespread meaning. However, as a figurative combination it has already moved away from the original meaning of the noun to a considerable degree, e.g.

*Christus on meije perral/*

*Christus gehöret vns/ (Stahl, LS, 12)*

‘Christ belongs to us’

The catechism by Wanradt-Koell and a text by J.A.Völcker reveal a single instance of the grammaticalized adposition *päralt* in the fixed phrase *kellegi päralt olema* ‘to belong to someone’ that expresses possessiveness, e.g.

*Ollecksit sina seperest köhenrettin ninck pörkwhawdan  
pereld/ (Wanradt-Koell 1535)*

‘If you belonged because of this to devil and Hell’  
*Sina ei peat tachtma oma læhembe kotta, sullan, tuttrikut,*  
*döbbrat, ninck kik mea tema **peralt** om* (Völcker, ca 1585)  
 ‘You must not covet your neighbour’s house, hind, maid,  
 domestic animals and all that belongs to him’

Müller uses the same adposition in a similar manner at the beginning of the 17<sup>th</sup> century. In his sermons, too, *päralt* occurs 15 times in the combination *kellegi päralt olema* ‘to belong to someone’.

Judging by the texts in the old literary language, the adpositions that had developed from the noun *pära* had their own range of application. *Päral* and *päralt* express mostly a sensuous process or possessiveness that is abstracted from the spatial meaning, *päras* expresses spatial location, and *pärast* causality.

### 1.11. *sega* → *sekka*, *seas*

Of this group of words only the adpositions *sekka* ‘to among, to amid’ *seas* ‘among, amid’ and *seast* ‘from among, from amid’ are used in the contemporary literary language. It was difficult to find instances with a clear nominal use in the texts in the old literary language. However, there were occasional examples of the adposition that expresses location:

*istup sehl kesckil nende oppijade **sehhas**/*  
*sitzet da mitten **vnter** den Lerern/* (Stahl, LS, 151)  
 ‘he is sitting there in the middle among the learners’

There are all in all ten occurrences of the adposition *sekka* and 32 occurrences of *seas* in the sermons by Georg Müller, e.g.

*Ninck se Sana say Lihax/ ninck ellas meddy **seas*** (Müller 2/5)  
 ‘And the word became flesh, and it lived among us’

In the texts by Stahl the form *sehhas* occurred 53 times, and the shortened *sehs* in the sense ‘among, amid’ occurred 19 times.

In the works by Stahl *sekka* ‘to among, to amid’ occurred 33 times as an adposition, e.g.

*tulli temma wainlane ninck külwis umbrochto se nisso **secka**/*

*kam sein Feind/ vnd säet Vnkraut zwischen den Weitzen/*  
(Stahl, LS, 215)

‘There came his enemy and sowed some weed amid the wheat’

Taking into account the contexts where *sekka* and *seas* occur, it is likely that there may have been adpositions in the Estonian language already before the 17<sup>th</sup> century.

### 1.12. *tava* → *taval*

The word *taval* is not common in the old literary language in its contemporary sense of manner. The earliest attested occurrence dates from the text that L. Boierus wrote down at the end of the 16<sup>th</sup> century:

*ey mötlate kui palyo neit rahuit waise tawal hucka läwet?*

‘You don’t think how many nations perish miserably’

The sermons by Müller do not contain this adposition. The books by Heinrich Stahl reveal only a single occurrence as a noun that occurs together with an attribute that agrees with it:

*Eth nühdet se Lehemb ni monnesarnasel taval tappetuth sahp/*

*Wann dann der Neheste auff so mancherley Art vnd Weise getödtet wird/* (Stahl, LS, 721)

‘That now the nearest will be killed in a similar manner’

Unfortunately, the texts in the older literary language reveal too little information about this adposition.

### 1.13. *tarve* → *tarbeks*

The abstract noun *tarve* ‘need’, which is rather rare in the oldest literary language, has given rise to its grammaticalized translative form *tarbeks* ‘for the purpose of’. It is likely that thanks to the frequent use of this case form the word may have undergone adpositionalization by the period of the oldest literary language. The texts of Older Literary Estonian reveal a number of examples of the grammaticalization process. The earliest one can be found in the catechism by Wanradt-Koell, which was printed in 1535:



*Mea meyle önnis on/ synu rickus tarbix/*

‘What is blessed to us for the purpose of your kingdom’

It seems that the most frequent one in the texts of the old literary language is the fossilized construction *seks tarbeks* ‘for this purpose’. Its German equivalent is also a word that has lost its full meaning. However, such a use cannot be regarded as adpositional as yet. It seems that a combination of the pronoun and the noun that agrees in case may act as a preliminary stage towards its development into an adposition, e.g.

*Sex tarwix on sünnul heh abbi ninck nouw/*

*Darzu weistu gut hülff vnd rath/* (Stahl, HH II, 115)

‘For this purpose you have good help and advice’

By analogy with the above-mentioned and other similar combination-like constructions *tarbeks* could have been placed next to a genitival noun in the function of expressing purpose. At this the adposition has retained a close semantic link with the base word – *tarve* is an abstract noun, and the resulting adposition expresses a rather abstract relationship of purpose as well.

The older literary language reveals also examples of the construction *genitive + tarbeks*, e.g.

*Ninck hoidkut tulle tarbix sel suhrel kochto pehwal/*

*Das sie zum Fewer behalten werden/ am tage des Gerichtes vnd Verdammüs* (Stahl, HH III, 147)

And may they keep it for the purpose of the fire on this great doomsday’

The preceding example indicates that *tarbeks* has been translated into German by means of a preposition, which may confirm that Stahl was familiar with the Estonian postposition *tarbeks* because it is not a literal translation.

#### 1.14. viis → viisi

There is only a single occurrence of this manner-denoting adposition in the oldest Estonian-language texts, which can be found in an excerpt from the law on Livonian peasantry dating from the middle of the 16<sup>th</sup> century:

*Keß Roiwi wise uhest toisest Wottap the pehl ellick kirko  
ellick toa seest*

‘who in the robber’s manner takes from another on the road  
or from inside a church or a room’

According to the present-day definition, the previous example has an adposition although it would be difficult to claim it for sure, taking into account the peculiar use of forms in the texts in the old literary language. Anyway, this kind of occurrence shows that *viisi* was already used in the position of the adposition although the meaning is very close to that of the base noun.

## 2. Adposition → case ending

### 2.1. kaas → kaa → -ga

When ignoring Old Literary Estonian, one cannot ignore the most typical and interesting case of grammaticalization. During the period of the documented literary language the adposition *kaas* or *kaa* developed into the modern comitative ending *-ga*.

The initial stage in the genesis of this case can be observed in the light of the texts in the older Estonian literary language. In the texts in the older literary language the postposition *kahs* denotes abstract comitativeness (co-existence and co-presence) and also the instrumental relationship abstracted from the latter. The meanings of manner, state, and time are in turn more abstract than the previous meaning and express a higher degree of grammaticalization. This adposition is also a frequent dependent extension of some verbs (e.g. *sõdima* ‘to fight’ *täitma* ‘to fill’ *rääkima* ‘to talk’, and *kõnelema* ‘to speak’), expressing a thing at which an action is directed (cf. EKG I: 60). Thus, one could say that in the old literary language the postposition *kaas* occurs in the same functions as the modern case. Therefore, one cannot say that the case ending added new grammatical uses.

Huno Rätsep has provided a historical overview of the development of the comitative. He relates the genesis of the comitative with the *-ga* ending to the spread of the adpositional mode of expression in the earlier period of Balto-Finnic. The noun *kansa* gave rise to the *k*-lative form *\*kansak*, which came to be used as a postposition together with the head in the genitive. In the

early language the postposition was subject to a number of phonetic changes: the vocalization of *n* before *s* resulted in \**kaasak*. The word-final *k* then disappeared, and the shape of the postposition was \**kaasa*. Later the final vowel also disappeared, so that the postposition took the shape *kaas*, which is common in the older literary language. After the genitival *n* at the end of the head had been lost, the adposition could be attached more closely to the adposition. This unstressed position gave rise to various irregular changes, the first one being the loss of *s* from the end of the postposition. After the final attachment of the postposition to the head, the intervocalic fortis stop *k* was replaced by the lenis stop, and this is how the modern case ending *-ga* came into existence (Rätsep 1979: 77–78).

Martin Ehala has studied the development of the comitative from the point of view of the grammaticalization theory. He distinguishes four steps in the development of the comitative case from the noun *kansa* ‘people, crowd’ through the intermediate stage of the postposition *kaas* into the modern case ending *-ga* (Ehala 1996: 382).

One can assume the following stages in the development of the comitative meaning: belonging to a certain crowd or group → co-existence with the animate → co-existence with the inanimate (instrumental meaning) → {metaphorization} → state → manner. This chain should also indicate an increase in the degree of grammaticalization.

In the earliest Estonian texts *kaas* occurs in various forms: *kaes*, *kaeß*, *ka*, *kas*, and *kaen* (the latter form occurs in South Estonian texts). At this the same word may have a number of different grammatical functions: its occurrence as an adposition is most common, but there are also numerous examples of its use as an emphatic modal verb, e.g.

*Temmale peame meije kahs offrima/ tenno/ auwo/ ninck  
kihtuse offer/*

*Dem sollen wir auch opffern weiß/ Danckopffer/ Danck vnd  
ewign Preiß/ (Stahl, HH II, 16)*

‘We must sacrifice to him as well, the sacrifice of gratitude,  
glory, and praise’

The question word *kaas* is represented by few examples. Paul Alvre has proved convincingly that it originated from the same stem form *\*kansak* as the comitative in *-ga* and the adverb *ka* ‘also’ (Alvre 1997), e.g.

*Kahs se Issa on sünnno eddes surnud?*

*Ist der Vater auch für dich gestorben?* (Stahl, HH I, 62)

‘Did the Lord die for you, too?’

In this example the German equivalent indicates that the author has also added the emphatic *auch* to the question although in the Estonian sentence *kahs* fulfils not the emphatic but the interrogative function. It is possible that in his translation Stahl may have sometimes proceeded from Estonian, thus trying to render the same content in German as much as he could.

It is difficult to say on the basis of the texts in the old literary language which above-mentioned use is primary because its uses both as an adposition and a modal verb were very common and are represented by hundreds of examples. The studied texts provided more examples of the instrumental meaning of *kaas*. In sum, there were fewer examples of the original meaning of co-existence and co-presence. However, as sermons have been written in figurative language, then the higher representativeness of the meaning that has shifted towards a higher degree of abstraction is justified.

The following example sentence illustrates the instrumental meaning:

*lauluge ütleb, pasunie ka, Sion linna sise* (Boierus ca 1587)

‘sing, says, with trumpets, into the town of Zion’

The expression of manner and then state by means of the construction including the postposition *kaas* is more abstract than the instrumental meaning, and it is likely that the latter may have developed from the former, e.g.

*minna pallun sünd allandusse kahs/*

*Ich bit dich demütiglich/* (Stahl, HH II, 177)

‘I beg you humbly’

The previous example illustrated the meaning of manner; the following one, however, should illustrate the meaning of state, e.g.

*Babyloni jögekede jures/ itsime meije kibbedusse kahs/  
An wasserstüssen Babylon/ da sassen wir mit schmerzen/  
(Stahl, HH III, 237)*

‘By the rivers of Babylon we were sitting embittered’

There were few examples of *kaas* used in the temporal meaning, which must have been abstracted from the instrumental meaning, e.g.

*Motle Issand se rascke aja pehl/ kumba kahs se iho sahp  
wangis piddatut/  
Gedenck O HERR der schweren Zeit/ damit der Leib  
gefangen ley/ (Stahl, HH II, 180)*

‘Think Lord about the hard times when the body is in captivity’

The adposition quite often occurs also in fixed expressions, where it is difficult to identify the function of the case, e.g. *nyme kaeß, nimmy kaen* ‘with the name’ or as a dependent extension of a verb (*lepitama kellega?* ‘to reconcile with sb’ *rääkima kellega?* ‘to talk to sb’ *söötma ning jootma millega?* ‘to dine and wine with sth’), e.g.

*Nynck meyd Jumalan themen Ekewe Isen kaes leppitis  
(Wanradt-Koell 1535)*

‘And he reconciled us with Lord, his Eternal Father’

The oldest texts also reveal several examples where the case ending is written solidly with the noun. The earliest sources reveal solidly written forms only with the pronoun, and even those occur in certain translation formulae. Examples include *issand sinuga* ‘Lord with you’ and *kuidas asi sinuga on* ‘how is it with you’ (*ßwkas* in the Kullamaa MS and two instances of *sinnucasy* ‘with you’ in a letter to one’s mother, which probably dates back to the 1530s; *temekaes* ‘with him’ in the catechism by Wanradt-Koell, where the surrounding context has perished; *sinokas* ‘with you’ in the text by Völcker). However, in the text by Boierus, which dates from the end of the 16<sup>th</sup> century, *-ga* is written solidly with the

nouns *hälega* ‘with voice’ and *wehega* ‘with force’ It indicates the beginning of the use of the case ending and, interestingly, it began earlier in the area of South Estonian than in North Estonian.

In the texts by Müller the frequent adposition *kaas* was often shortened to *kz* in the manuscript. There are only five cases where it is written solidly with the noun, whereas on two occasions the shortened form *kz* is written solidly with the noun *jumal* ‘god’ and once with the nouns *isa* ‘father’ *pere* ‘family’ and *toivutus* ‘pledge’ e.g.

*Eth meye nüith olleme onsax sanuth lebby sen Vssu, sysz on meil Rawo Iumalakz* (Müller 15/6)

As we have now become blessed through the Faith, then we have peace with the Lord’

The other examples of writing as one word relate to the pronouns, whereas *sekz* or *sekaas* occurs twice and *mingkkaas* or *minckkz* seven times. As there was considerable variation in whether words were written as one word or separately, then no conclusions can be drawn on these individual examples whether the case ending had come into existence or not. These examples prove, however, that the first step towards the case ending had been taken.

In *Hand- vnd Haußbuch* by Heinrich Stahl the adposition occurs in various forms: *kahs* is the commonest (694 occurrences); *ka* occurs as a postposition nine times and *kah* three times. One should also point out that *kahs* is also very common as a modal adverb in the meaning of a contemporary sentence or phrasal extension ‘also’ In the above-mentioned work it occurs four times as a question word in the modern sense *kas* ‘do, does’ Thus, one and the same word continued to have different grammatical functions. Only on two occasions *kahs* had been written solidly with the preceding noun, but because of their exceptionality these instances could be regarded as misprints:

*Minna usckun se Pöha Waimo sisse .. üllestousmisse sesinna-se münno Lehakahs/*

*Ich gläube an den H. Geist .. aufferstehung dieses meines Fleisches/* (Stahl, HH I, 22)

‘I believe in the Holy Spirit, in the resurrection of my flesh’  
*heckiselt sai sap temma suh sisse/ ettickakahs lastut/*

*bald ward Gall in seinen Mund/ mit essig gelassen/* (Stahl, HH II, 30)

'all of a sudden was bile introduced to his mouth with vinegar'

The use of *kaas* in the second work by Heinrich Stahl *Leyen Spiegel* is, generally speaking, similar. There, too, the postposition *kahs* (2,000 occurrences) expresses the function of the instrument, co-presence, or more rarely also the manner and state; *kah* (32 occurrences) and *ka* (3 occurrences) are very rare in the same function. However, the existence of variation and the fact that there exist shorter forms side by side with the longer adpositions indicates that the adposition is developing phonologically into a case ending. There were almost no occurrences where *kaas* was written solidly with the noun. Some cases are questionable because the space between the words is so small that they can be regarded as one word. At the same time these examples, too, can be typographical errors and may not indicate the development of the case ending.

One might assume that it is the fluctuating tradition of writing solidly and separately that created the preconditions for writing the phonologically suitable shortened forms (*kaa*, *kz*) solidly with the preceding noun. Judging by the texts from the first half of the 17<sup>th</sup> century the case ending had not taken shape as yet. but the same functions were fulfilled by the adposition. Contemporary grammars point out that the comitative case expresses the function of co-presence, instrument, manner or state, time, and the thing at which the action is directed. In fact, the oldest literary language offers examples for all of them.

## Summary

The article applied the grammaticalization theory to some adpositions in Oldest Literary Estonian, which are problematic from the point of view of part-of-speech labelling (i.e. they belong to a not fully closed class). During the period of the old literary language they were borderline cases between the full noun and a relational word. Therefore, they enabled us to observe the development paths of adpositions during the initial period of Literary Estonian.

The data of Older Literary Estonian reveal that it is mostly frequent locative case forms that have been subject to adpositionalization. As a result of a smaller or greater semantic shift they have moved away from the original lexical meaning of the corresponding nouns. Their meaning has become more abstract, they have lost some of their semantics, and at the same time the occurrence conditions of the grammatical context has become more restricted, e.g. *ase* *mele* 'to in place of' *ase* *mel* 'in place of'; *kesk* 'in the middle of', *keskel* 'in the middle of'; *kõrval* 'next to'; *poole* 'towards', *pool* 'at' *poolt* 'from, on behalf'; *pärra* 'into the hind part', *päras* 'in the hind part' *sekka* 'into the middle of', *seas* 'among' Semantically the adpositions under discussion expressed mostly place (*pärra*, *päras*; *kesk*, *keskel*; *poole*, *pool*, *poolt*), manner (*kombel*; *najal*; *nimel*; *taval*; *viisi*), purpose (*päralt*; *tarbeks*), and possession (*kätte*, *käes*, *käest*). The base nouns included concrete nouns (*naga* 'peg' *pära* 'hind part'), as well as nouns denoting place (*ase* 'place' *kesk* 'middle part') or body parts (*kõrv* 'ear' *käsi* 'hand, arm') but also abstract notion words, e.g. *hea* 'a good thing' *komme* 'custom' *nimi* 'name', *tava* 'custom', *viis* 'manner, way'. *pool* 'half'

The following table highlights the general semantic groups of base nouns in the order of increasing abstraction and illustrates the semantic distribution of adpositions.

Semantics of the base noun	Semantics of the adposition
PHYSICAL OBJECT: <i>pära</i>	PLACE: <i>pärra</i> , <i>päras</i>
PHYSICAL OBJECT: <i>pära</i>	CAUSE: <i>pärast</i>
PHYSICAL OBJECT: <i>naga</i>	MANNER: <i>najal</i>
BODY PART: <i>kõrv</i>	PLACE: <i>kõrval</i>
BODY PART: <i>käsi</i>	POSSESSION: <i>kätte</i> , <i>käes</i> , <i>käest</i>
PLACE: <i>ase</i>	REPLACEMENT: <i>ase</i> <i>mele</i> , <i>ase</i> <i>mel</i>
SPATIAL CONCEPT: <i>kesk</i>	PLACE: <i>kesk</i> , <i>keskel</i>
COLLECTIVE CONCEPT: <i>kansa</i>	CO-PRESENCE, INSTRUMENT, MANNER: <i>kaas</i> → <i>-ga</i>
ABSTRACT NOTION: <i>pool</i>	PLACE: <i>poole</i> , <i>pool</i> , <i>poolt</i>
ABSTRACT NOTION: <i>pool</i>	CAUSE: <i>poolest</i>



Semantics of the base noun	Semantics of the adposition
ABSTRACT NOTION: <i>sega</i>	MANNER: <i>sekka, seas</i>
ABSTRACT NOTION: <i>komme</i> <i>nimi</i> <i>tava</i> <i>viis</i>	MANNER: <i>kombel</i> <i>nimel</i> <i>taval</i> <i>viisi</i>
ABSTRACT NOTION: <i>tarve</i>	PURPOSE: <i>tarbeks</i>
ABSTRACT NOTION: <i>hea</i>	BENEFACTIVENESS: <i>heaks</i>

Grammaticalization as a common category-replacement process may be triggered because some word form is transferred to a new morphosyntactic environment, supported by re-analysis, and by analogy it becomes to be used frequently first and foremost as a construction. As far as Older Literary Estonian is concerned, the example of German-language constructions must have been essential at the stage of analogy.

During the stage of re-analysis, the important semantic extension may be caused by the original aspiration for figurativeness and innovation, a metaphoric or metonymic mode of expression. Adpositionalization may often have been triggered by an earlier use of the word as a constituent of fixed formulae (e.g. adverbial use in periphrastic verbs). Therefore, one can speak also of lexicalization as a supposed intermediate stage in the development of some adpositions. In that case one has to assume that upon moving away from the full-word paradigm and before becoming a grammatical device the word acquires its own lexical meaning. For example, the frequent combination *aeg on käes* 'the time has come' can be traced to the periphrastic verb *käes olema* 'to be at hand'. Its extension is conceivable as an adpositional use: *aeg on meie käes* 'time is in our hand'.

Considering the partly German-style constructions of Older Literary Estonian, the examined adpositions could have come into existence as a result of the following chain: FULL NOUN → A FREQUENT CASE FORM LEAVES THE PARADIGM (in some cases lexicalization) → SEMANTIC SHIFT (re-analysis) → ASSOCIATION WITH A NOUN IN A GRAMMATICAL CASE

→ FORMULAIC USE (analogy; example of the German language) → ADPOSITION.

The lastly discussed comitative ending *-ga* is the next step of grammaticalization. It is movement towards synthetic expression, where an auxiliary word has developed into a grammatical morpheme. The development of this case ending is somewhat unusual in that all the functions of the comitative that have been described in the modern language were already present in the adposition *kaas* (or its phonetic variants) in the oldest literary language, which precede the development of the case ending.

The previous examples of adpositionalization proved that the grammatical means of Estonian are developing towards analyticity. At the same time all the previously discussed late adpositions belong to the secondary adpositions, which express concrete grammatical relations with a relatively low degree of abstraction. They are still transparent, as far as their origin is concerned. It means that the same stem occurs at the same time as a noun – such are the previously discussed words *ase* ‘place’ *hea* ‘a good thing’ *komme* ‘custom’ *käsi* ‘hand, arm’ *kõrv* ‘ear’ *nimi* ‘name’ *pool* ‘half’ *pära* ‘hind part’ *tava* ‘custom’ *tarve* ‘need’ and *viis* ‘way’. As for these adpositions, one could also say that each of them is associated with a certain semantic field: *ase* with an abstract spatial relation, *komme*, *tava*, and *viis* with the expression of manner, *käsi* with possession, *kõrv*, *pool*, and *pära* with place relations, etc. The class of secondary prepositions is an open one; new members can be added. At the same time, the primary adpositions form a closed class and the core of the adpositions as a part of speech. The primary adpositions are mostly monosyllabic and have few semantic features. In contrast to the secondary adpositions their grammaticalization is stronger.

It is fully possible that the frequent use of several adpositional constructions may have been caused by the impact of German-language structures. In order to render the numerous German adpositional relations, one had to find the suitable means in Estonian, and if they were absent, then one had to do with translation loans. In this respect the earlier Estonian texts reveal both older adpositional constructions as well as those that may have appeared in the course of translation during the early period

of the literary language. Examples of the latter include *asemel* (German *an Statt, an Ort*), *heaks* (*zu gut*), *nimel* (*im Namen*), and *viisi* (*auf Weise*). A more hidden structural impact is manifested in the fact that although Estonian may have had other means for the expression of some grammatical function, an analytic construction was often used, which retained the number of words in a phrase. The texts of the old literary language do not reveal anymore many nominal base parts of the development chains that were discussed above although they must have existed in the popular language.

Taking into account the gradual nature of grammaticalization, one could pinpoint on the basis of texts in the old literary language such adpositions that had entered the process on in the 16<sup>th</sup> and 17<sup>th</sup> centuries. Examples include *asemele*, *asemel*; *kombel*, and *tarbeks*. Judging by the texts in the old literary language, one might say that the adpositions *keskel*; *poole*, *pool*, *poolest*; *pärra*, *päras*, *pärast*, *päralt*; *sekka*, *seas* are more adpositionalized. They are represented by numerous examples in different and not only formulaic constructions. The German-Estonian parallel texts offer rather free and not word-for-word translation equivalents for them. Also the words *kõrval*; *najale*, *najal*; *nimel*; *taval*, and *viisi* seem to be rather adpositionalized. However, as the oldest texts reveal too few examples of their use, then it is difficult to assess their development and degree of grammaticalization. At the same time, all the above-mentioned adpositions are still at the stage of the so-called secondary adpositions. Therefore, their meaning is not that abstract and they do not express the grammatical relations as purely as one would assume in the case of the primary adpositions.

Thus, one could conclude that the universal grammaticalization patterns are also suitable for the description of the development of adpositions in Literary Estonian. Here, too, the important developments include re-analysis and extension of meaning, which precede the transfer of the new grammatical relation, and analogy, which extends the potential usage contexts of the new unit in a similar morphosyntactic environment.

## REFERENCES

- Alvre, Paul 1997:** Kõsisõnast *kas?* vana kirjakeele taustal. – Pühendus-teos Huno Rätsepale. (Tartu Ülikooli eesti keele õppetooli toimetised 7.) Tartu. 9–15.
- Diewald, Gabriele 1997:** Grammatikalisierung. Eine Einführung in Sein und Werden grammatischer Formen. Max Niemeyer Verlag. Tübingen.
- Duden 1997 =** Der kleine Duden. Deutsche Grammatik. Rudolf Hoberg, Ursula Hoberg, Eine Sprachlehre für Beruf, Fortbildung und Alltag. Saksa keele grammatika. Tõlkinud ja kohandanud Maie Lepp. TEA.
- Ehala, Martin 1996:** Integreeritud keeleteooria võimalikkusest tänapäeva keeleteaduses (II) – Keel ja Kirjandus, no. 6. 375–384.
- Ehasalu et al 1997 =** Epp Ehasalu, Külli Habicht, Valve-Liivi Kingisepp, Jaak Peebo, Eesti keele vanimad tekstid ja sõnastik. (Tartu Ülikooli eesti keele õppetooli toimetised 6.) Tartu.
- EKG 1995 =** Mati Erelt, Reet Kasik, Helle Metslang, Henno Rajandi, Kristiina Ross, Henn Saari, Kaja Tael, Silvi Vare, Eesti keele grammatika I. Morfologia. Sõnamoodustus. Trükki toimetanud Mati Erelt peatoimetajana, Tiiu Erelt, Henn Saari ja Ülle Viks. Eesti Teaduste Akadeemia Keele ja Kirjanduse Instituut. Tallinn.
- Heine, Bernd, Ulrike Claudi, Friederike Hünemeyer 1991:** Grammaticalization. A Conceptual Framework. The University of Chicago Press. Chicago.
- Hopper, Paul J., Elizabeth Closs Traugott 1993:** Grammaticalization. Cambridge Textbook in Linguistics. Cambridge University Press. Cambridge.
- Itkonen, Esa 1998:** Mitä on 'kieliopillistuminen'? – Anneli Pajunen (toim.), Kieliopillistumisesta, analogiasta ja typologiasta. Suomalaisen Kirjallisuuden Seura. Helsinki. 11–58.
- Jaakola, Minna 1997:** Genetiivin kanssa esiintyvien adpositioiden kieliopillistumisesta. – Kieli 12. Tapani Lehtinen ja Lea Laitinen (toim.), Kieliopillistuminen: tapaustutkimuksia Suomesta. Helsingin yliopiston suomen kielen laitos. Helsinki. 121–156.
- Laanest, Arvo 1975:** Sissejuhatus läänemeresoome keeltesse. Tallinn.
- Lehmann, Christian 1985:** Grammaticalization: Synchronic Variation and Diachronic Change. – *Lingua e Stile* 20 (3). 303–318.
- Metslang, Helle 1994:** Grammatikaliseringumisest eesti ja soome keeles. – Karl Pajusalu ja Valma Yli-Vakkuri (toim.), Lähivertailuja 7 Suomalais-virolainen kontrastiivinen seminaari Tammivalkamassa

- 5.–7.5. 1993. (Turun yliopiston suomalaisen ja yleisen kielitieteen laitoksen julkaisuja 44.) Turku. 9–32.
- Metslang, Helle 1997:** Eesti keele ja teiste soome-ugri keelte futuurumi arenguid. – Keel ja Kirjandus, no. 4. 226–231.
- Ojutkangas, Krista 1998:** Itämerensuomalainen näkökulma kieliopillistumiseen. – Sananjalka 40. Suomen Kielen Seura. Turku. 67–86.
- Onikki, Tiina 1997:** Kieliopin ja leksikon rajalla. Paikallissijaiset olotilanimaukset. – Kieli 12. Tapani Lehtinen ja Lea Laitinen (toim.), Kieliopillistuminen. Tapaustutkimuksia suomesta. Helsingin yliopiston suomen kielen laitos. 89–120.
- Rätsep, Huno 1979:** Eesti keele ajalooline morfoloogia II. TRÜ eesti keele kateeder. Tartu.

## GRAMMATIKALISEERUMISEST VANIMA EESTI KIRJAKEELE KAASSÖNADE NÄITEL

### Külli Habicht

Artiklis käsitletakse grammatikaliseerumisteooriast lähtudes vanima eesti kirjakeele tekkestaadiumis kaassõnu. Grammatikaliseerumise all mõistetakse keeles pidevalt toimuvat grammatiliste kategooriate tekke või kitsamalt leksikaalsete üksuste grammatilisteks muutumise protsessi. See on keele loomuliku süsteemi osa – mingi leksikaalne morfeem läheb üle grammatilise morfeemi staatusesse.

Et adpositsioonide kujunemist peetakse universaalseks grammatikaliseerumise ilminguks ja kaassõna kui sõnaliik sisaldab endas suhteliselt erineva grammatikalisatsioonistmega ainet, lähtutakse uurimuses mõnest vana kirjakeele adpositsioonist ja ka ühest käändest – komitatiivist –, mille teket võib jälgida eesti kirjakeele säilinud andmete põhjal. Seetõttu on valik semantiliselt küllalt kirju ning ka tekkeprotsessi seisukohast ebahühtlane, ent peaks sellest hoolimata heitma valgust küsimusele, missugused on kaassõnade kujunemise teed ja kas need protsessid on toimunud eesti kirjakeeles tüüpiliselt või on neis midagi ebaootuspärast. Samas võimaldab grammatikalisatsiooni-teooria tuua selgust adpositsioonide jaotumisse – nende eristamisse grammatikaliseerumisastme alusel.

Käsitluses püütakse mõne kaassõna kujunemise näitel osutada ka saksa keele tüpoloogilisele analüütilisuse suunas nihutavale mõjule. Põhiliselt keskendutakse vaid probleemsetele juhtumitele ega räägita

neist paljudest vanadest kaassõnadest, mis vanima kirjakeele perioodiks on jõudnud grammatikaliseerumises juba primaarsete adpositsioonide staadiumi (nt *alla*, *all*, *alt*; *sisse*, *sees*, *seest* või *peale*, *peal*, *pealt*). Valikut ja analüüsi mõjutab mõistagi tänapäevane grammatika-kompetents.

Lähtutakse arusaamast, et abisõnastumine on suuri ajalõike hõlmav pidev protsess, kusjuures eri sõnad võivad mingil ajahetkel olla erinevates arengustaadiumides, piltlikult väljendades paikneda eri kaugustel skaalal, mille alguspunktiks on täistähenduslikud ja lõpp-punktiks tähenduselt üldistunud ja kasutuselt spetsiifiliseks muutunud suhtesõnad. Aluseks võetakse üldlevinud eeldus, et grammatika-vahendeid tekib juurde leksikaalsete sõnade arvel, juhul kui viimaseid hakatakse kasutama teatud kindlas grammatilises seoses. Oluline on konstruktsiooni kasutuse automatiseerumine, teatud vormellikkus, mis loob eeldused sõna kasutuse laienemiseks ka muudesse analoogilistesse kontekstidesse. Püütakse näidata, et tänapäeva eesti keele kaassõnakasutusele omased jooned ja probleemid on suures osas olemas olnud juba ligi neli sajandit tagasi kirja pandud vanimas keeles.

Materjal on kogutud eesti keele vanimatest tekstidest: läbi on töötatud kõik teadaolevad tekstid 16. sajandist (Ehasalu jt 1997), Georg Mülleri 39 jutluse (1600–1606) originaalkäsikiri ning Heinrich Stahli kaks kirikuraamatut: “Hand- vnd Haußbuches Für die Pfarherren/ vnd Haußväter Ehnstnischen Fürstenthumbs ..” (lühendatud HH) I–IV osa aastatest 1632–1638 ning “Leyen Spiegel” (lühend LS) I–II osa aastatest 1641 ja 1649. Näidete järel esitatakse sulgudes allika nimetus või selle lühend ja lehekülje number, G. Mülleri jutluste puhul viitab esimene number jutluse ja teine lehekülje numbrile; “Hand- vnd Haußbuch’il” on lisatud ka osa number (I–IV).

Kaassõnu sünnib keelde põhiliselt kahel viisil – substantiividest ja verbidest. Eesti keele kaassõnad on enamjaolt substantiivide kivinenud kohakäändevormid, harvem on neid tekkinud verbidest, nt *hoolima* → *hoolimata*, *vaatama* → *vaatamata*. C. Lehmanni järgi läbivad adpositsioonid tekkel järgmise protsessi: relatsiooni (nt kohta või suunda) väljendav substantiiv → sekundaarne adpositsioon → primaarne adpositsioon → aglutineeriv afiks → fuseeriv afiks (Lehmann 1985: 304).

Nimisõnavormide grammatikaliseerumise käigus saab enamasti kohakäändes substantiiv tähenduse avardades juurde grammatilise funktsiooni, mis sarnaneb tüüpiliste kohakäänete funktsiooniga, ehkki vormistub analüütiliselt. Selline konstruktsioon täidab morfo-süntaktiliste struktuuride koosseisus grammatilisi ülesandeid. Keelte tüpoloogiline uurimine on avastanud rea üldisi tendentse grammatikaliseerumises, st et on olemas kognitiivselt või sotsiaalselt olulisi tähendusi, mis

grammatikaliseeruvad eriti sagedasti. Üheks selliseks on kohta või ruumi valdkond, mida peetakse grammatikaliseerumise esimeseks astmeks, järgnevad aja-, viisi- jm abstraktsemad suhted. Uue grammatikavahendi tekkekohal ei pruugi keeles olla funktsioonilist tühikut. Kujunev suhtesõna saab nii paradigmaatilisel kui ka süntagmaatilisel seotumaks kui varem, see asetub selgepiirilisesse ja väiksema liikmete arvuga morfosüntaktilisse paradigmasse, kusjuures sõna süntaktiline paindlikkus ja iseseisvus väheneb. Sellega seoses muutub tähendus spetsiifilisemaks: täistähendusliku käändsõna paradigmast eemaldunud vorm saab omaette, enamasti üldisema tähenduse kui lähtesõnal.

Eesti vanema kirjakeele andmete põhjal selgub, et suhtesõnastunud on peamiselt sagedase kasutusega kohakäänete vormid, mis on väiksema või suurema tähendusnihe tulemusel eemaldunud vastavate substantiivide esialgsest leksikaalsest tähendusest, muutunud sisult abstraktsemaks ja vaesemaks, samas ka grammatilise konteksti mõttes esinemistingimustelt piiratumaks, nt *asemele, asemel; kesk, keskel; kõrval; poole, pool, poolt; pärra, päras; sekka, seas*. Vaatluse all olnud kaassõnad väljendavad enamasti kohta (*pärra, päras; kesk, keskel; poole, pool, poolt*), viisi (*kombel, najal, nimel, taval, viisi*), otstarvet (*päralt, tarbeks*) ja omamist (*kätte, käes, käest*). Lähtesubstantiivideks on olnud nii konkreetseid asjasõnu (*naga, pära*), samuti kohta (*ase, kesk*) või kehaosa märkivaid sõnu (*kõrv, käsi*), ent ka abstraktseid mõistenimetusi, nt *hea, komme, nimi, tava, viis, pool*.

Grammatikaliseerumise kui keele muutumise käigus tavalise kategooriavahetusprotsessi taustaks võib olla asjaolu, et teatud sõnavorm kantakse reanalüüsi tulemusena üle uude morfosüntaktilisse ümbrusse ja seda hakatakse eelkõige konstruktsioonina analoogia mõjul sageli kasutama. Eesti vanema kirjakeele puhul on analoogia etapis ilmselt oluline olnud saksa keele konstruktsioonide eeskuj.

Reanalüüsi etapis olulise tähendusülekande taustaks võib olla esialgne kujundlikkuse- ja uudsusetootlus, metafoorne või metonüümne väljendusviis. Tihti võib olla kaassõnastumiseks tõuke andnud varasem kasutus püsivormelite koosseisus (nt adverbiline kasutus kas või perifrastilistes verbides), mistõttu võib mõne kaassõna kujunemisel rääkida ka leksikaliseerumisest kui arvatavast vaheetapist. Sellisel juhul tuleb eeldada, et sõna saab täistähenduslikust paradigmast eemaldudes ja enne grammatikavahendiks muutumist omaette leksikaalse tähenduse, nt sagedase kasutusega ühend *aeg on käes* taandub perifrastilisele verbile *käes olema*, mida laiendades on omakorda mõeldav adpositsiooniline kasutus: *aeg on meie käes*.

Vaadeldud kaassõnade tekkeahel võiks eesti vanema kirjakeele osalt saksamõjulisi konstruktsioone arvestades olla järgmine: TÄISTÄHEN-

DUSLIK NIMISÕNA → SAGEDASE KÄÄNDEVORMI EEMALDUMINE PARADIGMAST (+ mõnel juhul leksikalisatsioon) → TÄHENDUSNIHE (reanalüüs) → SEOS GRAMMATILISES KÄÄNDES NOOMENIGA → VORMELILAADNE KASUTUS (analoogia; saksa keele eeskuju) → KAASSÕNA.

Artiklis viimasena käsitletud komitatiivi lõpp *-ga* kujutab endast grammatikaliseerumise järgmist astet, liikumist sünteetilise väljenduse poole, kus abisõnast on arenenud grammatiline morfeem. Mõnevõrra ebatavaline on selle käändelõpu kujunemise juures asjaolu, et kõik komitatiivi tänapäeval kirjeldatud funktsioonid on olemas olnud juba vana kirjakeele kaassõnas *kaas* (või selle eri variantides), mis käändelõpu kujunemisele eelnes.

Eelnevad kaassõnastumise näited andsid tunnistust vanema eesti keele grammatikavahendite arenemisest analüütilisuse suunas. Samas kuuluvad kõik eespool käsitletud hilistekkelised kaassõnad nn sekundaarsete adpositsioonide hulka, mis väljendavad suhteliselt konkreetseid, madala abstraktsusastmega grammatilisi suhteid ning on tekkeprotsessi seisukohast veel läbipaistvad. See tähendab, et sama tüvi on kasutusel ka noomenina – sellised olid eespool käsitletutest *ase*, *hea*, *komme*, *käsi*, *kõrv*, *nimi*, *pool*, *pära*, *tava*, *tarve*, *viis*. Kõik need adpositsioonid seonduvad ühe kindla semantilise väljaga: *ase* abstraktse ruumisuhte, *komme*, *tava* ja *viis* viisi väljendamise, *käsi* possessiivsuse, *kõrv*, *pool*, *pära* kohasuhte edasiandmisega. Sekundaarsete adpositsioonide klass on avatud, sinna võib lisanduda uusi liikmeid, samal ajal kui primaarsed adpositsioonid moodustavad suletud klassi ja ühtlasi kaassõna kui sõnaliigi tuuma. Primaarsed adpositsioonid on ka enamasti ühesilbilised ja semantiliselt tunnustelt vaesed, nad on erinevalt sekundaarsetest adpositsioonidest tugevamini grammatikaliseerunud.

Mitmete kaassõnaliste konstruktsioonide sagedase kasutuse taga võib näha saksa keele struktuuri mõju – saksa rohkete adpositsiooniliste seoste edasiandmiseks oli vaja leida sobivad ainesed ka eesti keelest, ja kui neid sealt võtta polnud, tuli piirduda otsetõlgetega. Võimaliku saksamõjulisuse seisukohast on varasemates eestikeelsetes tekstides kasutusel nii vanemaid kui ilmselt ka kirjakeele algusaegadel tõlkimise käigus loodud laenulisi kaassõnakonstruktsioone. Viimastest võiks nimetada *asemel* (saksa *an Statt*, *an Ort*), *heaks* (*zu gut*), *nimel* (*im Namen*), *viisi* (*auf Weise*). Varjatum struktuurimõju ilmneb selles, et kuigi eesti keeles oli ilmselt olemas muid väljendusvõimalusi, on autorid siiski sageli kasutanud analüütilisi konstruktsioone, mis säilitavad fraasis sõnade arvu. Paljude eespool toodud arenguahelate nimisõnalisi lähteosi ei kohta enam vana kirjakeele tekstides, ehkki need on rahvakeeles olemas olnud.



Grammatikaliseerumise astmelisust silmas pidades võib eelnevast välja tuua kaassõnu, mis vana kirjakeele tekstide põhjal otsustades on 16.–17. sajandil alles grammatikaliseerumisprotsessi sisenenud. Sellised on näiteks *asemele, asemel; kombel ja tarbeks*. Mõnevõrra kaassõnastunumateks võib pidada adpositsioone *kesk, keskel; poole, pool, poolest; pärra, päras, pärast, päralt; sekka, seas*, mis on esindatud paljude näidetega erinevates, mitte vaid vormelilaadsetes konstruktsioonides ja mille kohta eestikeelne paralleeltekst pakub suhteliselt vabu, mitte sõnasõnalisi tõlkevasteid. Kaassõnastunud on ka *kõrval; najale, najal; nimel, taval ja viisi*, ent kuna nende kohta pakuvad vanimad tekstid liiga vähe näiteid, on nimetatud kaassõnade kujunemist ja grammatikaliseerumise astet raske hinnata.

Seega võib öelda, et grammatikaliseerumise universaalsed skeemid sobivad ka eesti kirjakeele kaassõnade arengu kirjeldamiseks. Siingi on oluline olnud tähenduse avardamine ja ümbertõlgendamine, mis eelneb uue grammatilise seose edasiandmisele, ja analoogia, mis laiendab tekkinud üksuse potentsiaalseid kasutusvõimalusi sarnases morfosüntaktilises ümbruses.

# REFLECTIONS ON THE DEVELOPMENT OF A PARTICLE IN ESTONIAN<sup>1</sup>

Helle Metslang

Tallinn Pedagogical University /  
Institute of the Estonian Language

The Estonian particle *las* is commonly treated as the shortened version of the imperative form *lase* of the verb *laskma* 'let, allow, have sb do sth'. For example, *las minna* – *lase minna* 'let go (of), let things slide' *las olla* – *lase olla* 'let it be, let sb be, let sb alone' (i.e. stop or refrain from molesting, disturbing, or interrupting), *las ta tuleb* – *lase ta tuleb* 'let him come'. As perceived against the background of the well-researched world languages, extensions of the imperative form of this verb are common, cf. 'to let = to more than permit: *There's a letter for you, Sir, if your name be Horatio, as I am let to know it is.* ("Hamlet" by Shakespeare) (JD). Although the phonetic shape of the word *laskma* : *lase* is strongly reminiscent of its German homologue *lassen* and its imperative *lass*, the etymological sources suggest that *lase* is a native Balto-Finnic word, possibly belonging to the Finno-Ugric wordstock. For instance, the Finnish verb *laskea* 'sink' is of the same origin (EEW: 1238–1239). The German *lassen*, however, is of Indo-European origin, cognate with the English *let* and the Swedish *låta* (Duden: 387). Rich as the polysemy of this word may be in different languages, there is a common meaning 'to cause, to enable, to allow'. In this case, the verb 'let' has a propositional complement; its meaning, underlying a range of modal spin-offs, often involves modality. All of them call for an adverb modifying the whole sentence. The word-forms reminiscent of the particle *las*, *la(a)*, *lah* have been identified in Finnish dialects and in Old Literary Finnish (the Estonian and Ingrian influences can not be ruled out in this connection either). This use has been retained in the so-called optative, in the Ingrian *laa*, Votic *la(s)*, Liv *las*, *laz*, *la*, Latvian

---

<sup>1</sup> The study was supported by the Estonian Science Foundation, Grant No.: 3858.

*la(i)*, German *laß*, Russian *pust'* (SSA: 49, SKES: 277; Leskinen 1965), as well as in the Swedish *låt* and English *let*.

We are witnessing here the grammaticalization process where a word form gives rise to a particle with a new function. There are many examples of imperative forms becoming particles, cf. the Estonian *kuule* 'listen; look (who's talking)' *vaata* 'look here; look (who's talking)' *näe* 'look (what you've done)' *anna* 'give (it to him hot)' *vat* 'look! well! (drawing sb's attention)', the borrowed Russian imperatives *davai* 'come on (tell us the story)' (cf. Tragel 1999: 51; Tauli 1983: 231–232; Hakulinen & Seppänen 1992).

### 1. Variants of use

As a particle, *las* expresses mostly the act of enabling (1), or a permission-order (2).

- (1) *Las Liina aita-b sind*  
 let Liina:NOM help-PRS:3SG you:PRTV  
 'Let Liina help you'
- (2) *Las Ott oota-b ukse taga*  
 let Ott:NOM wait-PRS:3SG behind the door  
 'Let Ott wait behind the door'

Both meanings can refer to agent-oriented modality (cf. e.g. Bybee *et al.* 1994: 177 ff.). The first meaning can be construed as causing somebody to be able to perform a respective act, i.e. it involves causality and dynamic modality (capability-possibility of the actor, cf. Uuspõld 1989). It is categorized under the concept of root modality (for the use of the term in English, cf. e.g. Hasselgård *et al.* 1999). The second meaning can be defined as deontic modality, with the *las*-sentence equal to the jussive-predicate in sentence (3).

- (3) *Ott ooda-ku ukse taga.*  
 Ott:NOM wait-JUSS behind the door  
 'Ott (should) wait behind the door'

Representative of the situation rendered as modal is the predicate verb (*aitab* 'helps' *ootab* 'waits'), related to the grammatical subject denoting the actor (*Liina*, *Ott*). The source structures of

the *las*-sentences are evidently those featuring the imperative form *lase* as the predicate, to be modified by the basic verb in the *da*-infinitive and the noun in the partitive, expressing the actor of the basic verb (Partial Object), or the noun in the adessive (4), (5).

- (4) *Lase Liina-t / Liina-l ennas-t*  
 let:IMP Liina-PRTV / Liina-ADESS (your)self-PRTV  
*aida-ta*  
 help-daINF  
 'Let Liina help you'
- (5) *Lase Otti / Oti-l ooda-ta*  
 let:IMP Ott-PRTV / Ott-ADESS wait:daINF  
*ukse taga*  
 behind the door  
 'Let Ott wait behind the door'

It seems that the displacement of roles and correspondences of participants in the situation structure, the speech situation, and the information-forwarding structure triggered grammaticalization (cf. Rätsep 1971). In those three situations one can perceive that the following participants are of material significance in the given case.

1. Activity situation of the sentence: 1) actor of the verb *laskma* 'let' – the person causing or letting by persuasive or forceful means (abbr. L), 2) actor of the main verb – the deep subject (abbr. A). Henceforth, we will designate the activity of the main verb with the abbreviation P;
2. Speech situation: 1) speaker (Sp), listener (H), the third party (O);
3. Forwarding of information (statement, order, recommendation etc.): source (So), addressee (Adr), agent viz mediator (M).

### 1.1. Imperative *lase*

In the source structure *laskma* 'let' performs its full lexical meaning (expressing causality and mostly also dynamic modality), possessing also all verbal properties (4), (5). It has its own subject; the situation to be enabled is expressed by the verb in the *da*-infinitive. The meaning of the sentence can be treated (6) as modal causality 'L causes (A has a chance (A carries on

doing P))' or (7) also as non-modal causality 'L causes (A carries on doing P)'

(6) *Mari lase-b Jüri-l maga-da*  
 Mari:NOM let-PRS:3SG Jüri-ADESS sleep-daINF  
 'Mari allows/lets Jüri (to) sleep'

(7) *Direktor lase-b Ott-i ukse taga*  
 Director:NOM let:PRS:3SG Ott:ADESS behind the door  
*ooda-ta*  
 wait-daINF  
 'Director makes Ott wait behind the door'

When *laskma* 'let' is in the imperative, the sentence expresses a direct order for L to enable or cause the situation 'A carries on doing P'. The respective roles of three participants in the situation, when using *lase*, would be as follows.

Activity situation	∅	L	A
Speech situation	Sp	H	O or Sp
Information-forwarding situation	So	Adr	∅

To illustrate, sentence (5) is acceptable in a situation where the speaker is also the source of the order, the listener both the addressee and the person enforcing Ott to wait. Ott, the person liable to wait is a third party. The actor may also be the speaker himself (8); the listener hardly fits to suit that role.

(8) *Lase mu-l rääki-da*  
 let:IMP I-ADESS speak-daINF  
 'Allow me to speak'

The verbal properties of *laskma* 'let' are revealed here in the fact that *laskma* can be negated (9), (12), the subject, which is usually elliptical, can be formalized (10), (13), the pronominalization can be controlled by the subject of the verb *lase* (11), (13).

(9) *Ära lase Otti ooda-ta*  
 IMP:NEG let Ott:PRTV wait-daINF  
 'Do not keep Ott waiting'

- (10) *Sa lase Otti oodata*  
 you:NOM let:IMP Ott:PRTV wait-daINF  
 'You should let Ott wait'
- (11) *Sa<sub>1</sub> lase Otti ennas-t<sub>1</sub>*  
 you:NOM let:IMP Ott:PRTV (your)self-PRTV  
*ooda-ta*  
 wait-daINF  
 'You should let Ott wait for yourself'
- (12) *Ära lase Liina-l ennas-t*  
 IMP:NEG let Liina-ADESS (your)self-PRTV  
*aida-ta*  
 help-daINF  
 'Do not let Liina help yourself'
- (13) *Sa<sub>1</sub> lase Liina-l ennast<sub>1</sub>*  
 you:NOM let:IMP Liina-ADESS (your)self:PRTV  
*aida-ta*  
 help-daINF  
 'You should let Liina help yourself'

In the description of the English verb *let* (Hopper & Traugott 1993: 11), the propensity for passivization served as the criterion: if *let* is a verb form, the sentence can be transformed into the passive, cf. (14), (15).

- (14) *Let us go*  
 (15) *We were let to go*

Actually, two more transformations have been performed here: the tense has shifted from the present into the preterite and the communicative sentence type has changed from an order to a statement. As for Estonian sentences (1) and (2), the same transformations would yield sentences (16) and (17).

- (16) *Liina-l las-ti sin-d aida-ta*  
 Liina-ADESS let-IMPS:PST you-PRTV help-daINF  
 'Liina was let to help you'
- (17) *Otti las-ti ooda-ta*  
 Ott:PRTV let-IMPS:PST wait-daINF  
 'Ott was made to wait'

Also, instances where the recurrent German verb *lassen* functions as an auxiliary in the second person imperative *laß* (18), (19) belong to this type (cf. the analogous Estonian expression (20)).

(18) *Laß ihn lesen!*

(19) *Laß das sein!*

(20) *Lase oll-a!*

let:IMP be-daINF

'Let it be'

## 1.2. Particle *las* with dynamic modality

The role relations and the dynamic modal meaning remain the same also in sentences (1) and (21). The person 'enabling', however, no longer functions in the structure of the meaning as a fully-fledged actor.

(1) *Las Liina aita-b sind*  
 let Liina:NOM help-PRS:3SG you:PRTV  
 'Let Liina help you'

(21) *Las ma räägi-n*  
 let I:NOM speak-PRS:1SG  
 'Let me do the talking'

The formal aspect has changed drastically. *Las* itself has shortened in form, it is in an unstressed position, and it no longer functions as a verb. The predicate verb carries the main meaning. Therefore, the structure of the whole sentence has changed; the actor of the main verb is formalized as the grammatical subject (*Liina, ma*). Because *las* is no longer a verb, it can not be negated (22), it can not take a subject (23). As the actor is absent, the reflexive pronoun *ennast* 'myself/yourself/himself etc.', of the identical reference with the actor, can not be used, as exemplified in sentence (4). The personal pronoun *sind* 'you' remains the sole option (1).

(22) \**Ära las Liina aita-b sin-d*  
 IMP:NEG let Liina:NOM help-PRS:3SG you-PRTV  
 'Don't (negation) let (positive order) Liina help you'

- (23) \**Sa las Liina aita-b sin-d*  
 you:NOM let Liina:NOM help-PRS:3SG you-PRTV  
 ‘\*You (Nominative) let (positive order) Liina help you’
- (1) *Las Liina aitab sind*  
 ‘Let Liina help you’

The intra-sentence pronominalization rather refers to the subject of the main verb (24). Passivization, too, is conceivable only if the subject (*Liina*) of the main verb is omitted (25).

- (24) *Las Liina<sub>1</sub> aita-b ennas-t<sub>1</sub>*  
 let Liina:NOM help-PRS:3SG (her)self-PRTV  
 ‘Let Liina help herself’
- (25) *Sin-d aida-ti*  
 you-PRTV help-IMPS:PST  
 ‘You were helped’

The use of *lase* ‘permit, tolerate, have’ as a particle in (26), (27) and the employment of the *da*-infinitive of the main verb in conjunction with particle *las* (28) are classified as progressive steps between the first and the second uses.

- (26) *Lase Liina aita-b sin-d*  
 let:IMP Liina:NOM help-PRS:3SG you-PRTV  
 ‘Let Liina help you’
- (27) *Lase ma räägi-n*  
 let:IMP I:NOM speak-PRS:1SG  
 ‘Let me speak’
- (28) *Las Liina aida-ta sin-d*  
 let Liina:NOM help-daINF you-PRTV  
 ‘Let Liina help you’

From among the above verbal properties, only the attachment of the subject to *lase* could be considered in marginal cases, e.g. (29).

- (29) *Sa lase Liina aita-b*  
 you:NOM let:IMP Liina:NOM help-PRS:3SG  
 sin-d  
 you-PRTV  
 ‘You should let Liina help you/you should have Liina help you’



As for the form of the actor of the main verb, the total object, as *Liina* in (30), serves as an intermediate step between the object and the subject. The nominative case is used because *lase* is in the imperative (see e.g. Tauli 1983: 52).

- (30) *Lase Liina rääki-da*  
 let:IMP Liina:NOM speak-daINF  
 Allow Liina to speak'

The forthcoming semantic developments in Estonian will be effected mainly on account of the ongoing decline of the "enabling-permitting-having-something-to-be-done" person (L), as the participant. In the second use, the role of the person in the activity situation has diminished.

### 1.3. Concessive use

The next step would be the concession for the event, with the said 'enabler' withdrawn and reduced to an indefinite, general, imaginary entity (31), (32) (33).

- (31) *Las nad vaidle-vad pealegi,*  
 let they:NOM have\_an\_argument-PRS:3PL well  
*kiüll nad ükskord ära tüüdi-vad*  
 likely they:NOM some\_time PFV get\_bored-PRS:3PL  
 'Let them have their argument out to their hearts' content,  
 they are bound to get bored some time'
- (32) *Las ma tee-n selle eksami ära,*  
 let I:NOM do-PRS:ISG that:GEN exam:GEN PFV  
*siis tee-me peo*  
 then do:PRS:1PL party:GEN  
 'Let me pass that exam, then we will have a party'
- (33) *Las tule-b suvi, kiüll ma siis*  
 let come-PRS:1SG summer:NOM surely I:NOM then  
*sõida-n su-lle külla*  
 travel-PRS:1SG you-ALL on\_a\_visit  
 'Let the summer come, I will surely travel to your place then'

Activity situation	∅	∅	L	A
Speech situation	Sp	H	?	O or Sp
Information-forwarding situation	So	Adr	∅	∅

The said *las*-sentence is used as a concessive utterance in (31), or it may express time in the sense ‘wait / you wait, until’ (32), (33) (cf. Tauli 1983: 232). Nor does the actor, embodied by the speaker, seem out of place here (34).

(34) *Las sa kasvad, küll sa ükskord neile näitad*

‘Let you grow up, you will one day prove your worth to them’

#### 1.4. Jussive use

The total absence of the role of the person enabling the action is characteristic of two uses of the particle *las* and the main verb, which are equal to the use of the jussive. In both uses the speaker has a significant role: in 4a (2) it serves as the source of the order or recommendation issued to the third party, in 4b (36)–(38) it is the mediator of an order issued from whatever source. Involved here is the commonly known grammaticalization tendency, from agent-oriented modality to speaker-oriented moods (cf. Bybee *et al* 1994: 210 ff.).

On the one hand, this use is rooted in the causative source construction, and, on the other hand, in the concessive use. The concessive use of the *las*-phrase is, accordingly, furthered by the concessive use of the grammatical jussive, e.g. (35), cf. (31).

(35) *Vaiel-gu nad pealegi küll*  
 have\_an\_argument-JUSS they:NOM well likely

*nad ükskord ära tüdi-vad*

they:NOM some\_time PFV get\_bored-PRS:3PL

‘They should have their argument out to their hearts’ content, they are bound to get bored some time’

### 1.4a. Speaker's order levelled at the third party

Activity situation	∅	∅	A
Speech situation	Sp	H	O
Information-forwarding situation	So	(M)	Adr

The speaker is communicating to the listener his wish or request, to be fulfilled, or hopefully to be fulfilled, by a third party (2), as in the case of the jussive (3). The listener may mediate the order of the speaker.

- (2) *Las Ott oota-b ukse taga*  
 let Ott:NOM wait-PRS:3SG behind the door  
 'Let Ott wait behind the door'
- (3) *Ott ooda-ku ukse taga*  
 Ott:NOM wait:JUSS behind the door  
 'Ott (should) wait behind the door'

### 1.4b. Mediation of an ambipersonal order (i.e. an order issued by an implied person)

Activity situation	∅	∅	∅	A
Speech situation	x <sub>1</sub>	Sp	H	x <sub>2</sub>
Information-forwarding situation	So	M	∅	Adr

The speaker is conveying to the listener the order that had been issued by someone to someone else. The addressee of the order coincides with the actor; however, the role of neither the source of the order nor that of the addressee in respect of the speech situation has been delineated (36), (37), (38). The sentences with the jussive (39), (40), (41) express meaning in the same way.

- (36) *Arst ütles,*  
 doctor:NOM say-PST:3SG  
*et las Ott oota-b ukse taga*  
 that let Ott:NOM wait-PRS:3SG behind the door  
 'The doctor said that Ott should wait behind the door'

- (37) *Ma ütle-si-n,*  
 I:NOM say-pst-1SG  
*et las Ott oota-b ukse taga*  
 that let Ott:NOM wait-PRS:3SG behind the door  
 'I said that let Ott wait behind the door'
- (38) *Te ütle-si-te,*  
 you:NOM say-PST-2PL  
*et las me oota-me ukse taga*  
 that let we:NOM wait-PST:1PL behind the door  
 'You said that let us wait behind the door'
- (39) *Arst ütle-s,*  
 doctor:NOM say-PST:3SG  
*et Ott ooda-ku ukse taga*  
 that Ott:NOM wait-JUSS behind the door  
 'The doctor said that Ott (should) wait behind the door'
- (40) *Ma ütle-si-n,*  
 I:NOM say-pst-1SG  
*et Ott ooda-ku ukse taga*  
 that Ott:NOM wait-JUSS behind the door  
 'I said that Ott (should) wait behind the door'
- (41) *Te ütle-si-te,*  
 you:NOM say-PST-2PL  
*et me ooda-ku ukse taga*  
 that we:NOM wait-JUSS behind the door'  
 'You said that we (should) wait behind the door'

### 1.5. Direct order

In all Estonian uses, the examined situations involve at least three participants, as in the underlying causative source structure. The roles of different aspects reveal different distribution between the participants. The second development, the use of direct order is absent in Estonian, as there as only the speaker and the listener. In such expressions as the English *let's go*, *let me see*, the German *laßt uns gehen* or the Swedish *låt se*, *låt oss gå*, the 'let'-phrase expresses a direct order, addressed by the first person to oneself without a mediator. The meaning of enabling has disappeared or is in the process of doing so.

First Person Singular (*let me see*):

Activity situation	A	∅
Speech situation	Sp	H
Information-forwarding situation	So = Adr	∅

First Person Plural (*let's go, laßt uns gehen*):

Activity situation	A	A
Speech situation	Sp	H
Information-forwarding situation	So ∈ Adr	Adr

Judging by the German sources, it looks as if the use of the construction were limited to the first person plural. However, English is witnessing the grammaticalization of the particle *lets* into a more general marker of imperativeness or, in a narrower sense, exhortation. The use is widening and has embraced other persons – in addition to the first person plural, also the first person singular (42), the second person singular (43), and the second person plural (44).

(42) *Lets give you a hand*

(43) *Lets you go first, then if we have money left I'll go*

(44) *Lets you and him fight*

Furthermore, the form is reducing, as shown in the sequence: *let us* > *let's* > *lets* > *les* > *s*. There are examples of the last link like *sgo, sfight*. (Hopper, Traugott 1993: 11–14.)

Figure 1 summarizes the above-mentioned chain of development.

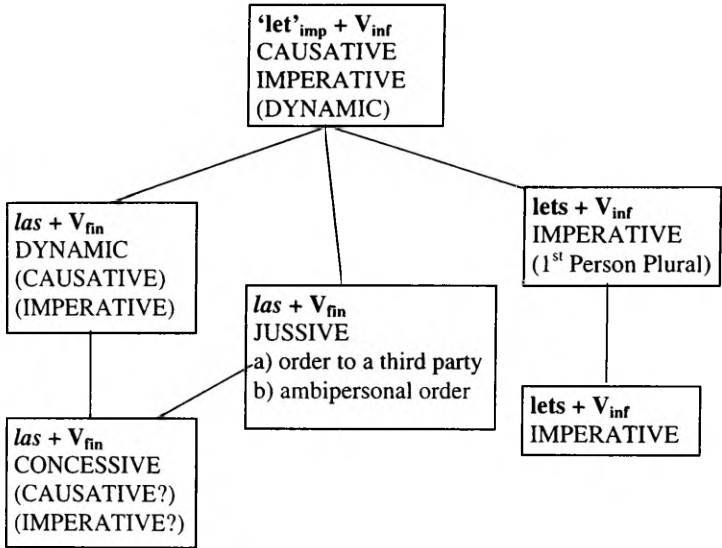


Figure 1. Trends of grammaticalization of the construction 'let' in the imperative + the second verb in the infinitive

## 2. Recurrence of variants of use in texts of Literary Estonian

Below we will focus on the expansion of constructions consisting of the word *lase* or *las* and a verb in texts of Literary Estonian in various periods. The survey is based mainly on the TÜKK corpus of Literary Estonian at the University of Tartu, embracing the period from the 1890s to the 1980s. In addition, the data of the Oldest Literary Language were cursorily examined on the basis of the WAKK corpus (1224–1600) and Ehasalu *et al* 1997, the text collection of the Old Literary Language until 1739 (Saareste & Cederberg 1992, abbr. SC), and the text corpus of the Institute of the Estonian Language (EKI), including mainly the material of the 1990s.

The periods considered belong to several stages of development of Literary Estonian, different as to their sociolinguistic circumstances and their impact. During the initial period, the development of Literary Estonian was moulded by foreign educationalists, mainly Germans. A strong German

influence still persists in Estonian and Estonian culture. In the 19th century, the Estonian *literati* took over; however the influence of German culture disappeared only as late as in the 20th century. In the context of Estonian-oriented, and in the wider Balto-Finnic-oriented corpus planning of the literary language, the syntax was tackled only during the independence of Estonia in the first half of the 20<sup>th</sup> century when Literary Estonian was intensively developed. After WWII, during the Soviet occupation, under the pressure of Russian and Soviet ideology, the nurturing, care and development of Literary Estonian, shaped during the independence, was the popular device of preservation of identity and of spiritual opposition and dissent. With Estonia becoming independent, that function of the literary language disappeared. Some other factors, too, have contributed to the liberalization of language use and to the increased western impact (cf. Hennoste *et al* 1999; Metslang 1998; Metslang forthcoming (a); Metslang forthcoming (b)).

### 2.1. The 13th–18th centuries

The Oldest Literary Language (13th–16th c.) reveals only two instances of the imperative form of the second person plural *laske*, originating from Wanradt & Koell, 1535, e.g. (45).

- (45) *ses hiwus nynck ses [weres] / laske kynny köwteda /*  
 (WAKK)  
 ‘in flesh and blood let it be bound’

The texts originating in the 17th century display mainly the use of the imperative form, the second person (and also the forms of other persons) in conjunction with the infinitive, i.e. the underlying source construction (the first attested use), equivalent to the main use of the German *laß*. There are variant spellings, equivalent to the form *lase* or *las*; there are also instances of direct German spellings *lass* or *laß* (48). (In respect of the infinitive, the confusion of the two Estonian infinitives, *da-* and *ma-*infinitive, is understandable because the writers were mostly non-Estonians.) One might assume that the German verb pattern and the similarity of the word itself contributed to the expansion of similar use in Literary Estonian, cf. the inserted heading in the Turku

manuscript dating from the first half of the 17<sup>th</sup> century (46), and the next sentence in Estonian (47).

(46) *Nun laßt vnss dem Leib begraben* (SC: 25)

(47) *Nvtt lasse meye se hiwo erra mattada* (SC: 25)

‘Now let us commit the body to earth’

(48) *Tulle tenne / Neitzikenne / Ninck heita Nalja Münno kahs / Ninck laß münd Armo nehha taas.* (Rainerus Brocman, 1638, SC: 55)

‘Come here dear maiden and have fun with me, and let my grace be seen again’

(49) *Agga lasse mind minno parrajast Roa Ossa sada*

(Johann Hornung 1694, SC: 215)

‘However let me partake of my best dish’

Besides the source structure there are also isolated examples of the use of the imperative particle (the fifth instance of use), to be found in German, however not having taken root in Estonian. In place of sentence (50), currently the exhortative “Palugem” (the imperative form of the first person plural) is used.

(50) *Laskem meit palluma* (Ristimise sõnad, 1674, SC: 158)

‘Let us pray’

In older grammars, such as the ones by Stahl (1637) or Gutsclaff (1648), the *laskma* material is not employed as a grammatical means. However, Göseken (1660: 37) presents it as an auxiliary verb to form parallel plural forms of the imperative, for instance, the forms corresponding to the modern ones *me armastagem* ‘we (should) love one another’ *te armastage* ‘you (should) love one another’ *nad armastagu* ‘they (should) love one another’, as presented under the pattern of Göseken, will be either *armastakem meije* ‘we (should) love one another’, *armastaket teije* ‘you (should) love one another’ *armastako / armastakut nemmat* ‘they (should) love one another’ or *laskem meije armastama* ‘let us love’, *lasket teije armastama* ‘let you love’, *laskut nemmat armastama* ‘let them love’

The 1739 translation of the Bible features an example of grammatical reanalysis (the second instance) – the main verb is in the finite form (49).



(51) .. *lasse ma kulen su heält* ..(SC: 349)  
 'let me hear your voice'

## 2.2. The 19<sup>th</sup> and 20<sup>th</sup> centuries

### 2.2.1. Numerical data.

The University of Tartu Corpus (TÜKK) enables us to establish the frequency of different variants of the construction in the language of the press and fiction over various decades (the 1890s, 1900s, 1930s, 1950s, 1960s, 1970s, and the 1980s). Table 1 presents the scope of the respective parts of the corpus.

Table 1.

#### The volume of component parts of TÜKK (in words)

Decade	Texts of press	Texts of fiction	Total
1890s	193,000	155,000	348,000
1900s	171,500	64,500	236,000
1930s	117,000	252,000	369,000
1950s	242,000	66,000	308,000
1960s	201,000	132,000	333,000
1970s	168,500	257,100	425,600
1980s	175,000	250,000	425,000
<b>Total</b>	1,268,000	1,176,600	2,444,600

Table 2 shows the frequency of the constructions *las* + V<sub>2</sub> and *lase* + V<sub>2</sub> in absolute figures; Table 3 shows the frequencies per 10,000 words. The imperative form *lase* seems to have been more frequent in the German-influenced language of the 1890s. It then decreased in volume and picked up anew in the 1950s and 1960s (the said form may have appeared in those particular times to be more correct, more perfect, and better-suited to fit the orthology of the literary language). The unmitigated particle *las* made the day again in the 1930s, taking the downward trend in the 1950s, whereafter it gradually became stable. The overall trend of development, though implies the decline of *lase* and the expansion of *las*; the above irregularities still call for explanation. Both constructions are more characteristic of the less departmentalized

language use, inherent in the language of fiction, among the two sublanguages.

Table 2.

*las* and *lase* + V<sub>2</sub> in texts of literary Estonian of past century (on the basis of TÜKK). Absolute figures

	1890 Press	1890 Fiction	1890 Total	1900 Press	1900 Fiction	1900 Total	1930 Press	1930 Fiction	1930 Total
<i>las</i>	0	7	7	0	2	2	3	37	40
<i>lase</i>	1	17	18	1	1	2	0	6	6
Total	1	26	27	1	3	4	3	46	49

	1950 Press	1950 Fiction	1950 Total	1960 Press	1960 Fiction	1960 Total	1970 Press	1970 Fiction	1970 Total
<i>las</i>	0	6	6	7	15	22	4	33	37
<i>lase</i>	1	6	7	0	8	8	0	3	3
Total	1	12	13	7	25	32	4	36	40

	1980 Press	1980 Fiction	1980 Total	Press total	Fiction total	Sum total
<i>las</i>	0	36	36	14	136	150
<i>lase</i>	0	3	3	3	45	48
Total	0	39	39	17	187	198

Table 3.

*las* and *lase* + V<sub>2</sub> in texts of literary Estonian of past century (on the basis of TÜKK). Rate of occurrence per 10 000 words

	1890 Press	1890 Fiction	1890 Total	1900 Press	1900 Fiction	1900 Total	1930 Press	1930 Fiction	1930 Total
<i>las</i>	0	0.45	0.2	0	0.31	0.08	0.26	1.47	1.08
<i>lase</i>	0.05	1.10	0.52	0.06	0.16	0.08	0	0.24	0.16
Total	0.05	1.55	0.72	0.06	0.47	0.16	0.26	1.71	1.24

	1950 Press	1950 Fiction	<b>1950 Total</b>	1960 Press	1960 Fiction	<b>1960 Total</b>	1970 Press	1970 Fiction	<b>1970 Total</b>
<i>las</i>	0	0.91	<b>0.19</b>	0.35	1.14	<b>0.66</b>	0.24	1.28	<b>0.87</b>
<i>lase</i>	0.04	0.91	<b>0.23</b>	0	0.61	<b>0.24</b>	0	0.12	<b>0.07</b>
Total	0.04	1.82	<b>0.42</b>	0.35	1.75	<b>0.90</b>	0.24	1.40	<b>0.94</b>

	1980 Press	1980 Fiction	<b>1980 Total</b>	Press total	Fiction total	Sum total
<i>las</i>	0	1.44	<b>0.85</b>	0.11	1.16	0.06
<i>lase</i>	0	0.12	<b>0.07</b>	0.02	0.38	0.20
Total	0	1.56	<b>0.92</b>	0.13	1.54	0.26

The form of the main verb in different periods also implies the trend towards grammaticalization (Tables 5 and 6). Uses of *lase* and *las* are diverging: in case of *lase* the main verb is more often in the *da*-infinitive, in case of *las* it is in the finite form. The *lase*-phrases predominantly express dynamic modality or the meaning of concession, the *las*-constructions having either a concessive or deontic meaning.

Table 4.

**V<sub>2</sub> form of *las* in texts of different decades  
(TÜKK, total in press and fiction; absolute figures)**

V <sub>2</sub>	1890s	1900s	1930s	1950s	1960s	1970s	1980s
<i>da-inf.</i>	2	1	4	1	2	4	4
<b>Finite form</b>	5	1	36	5	20	33	32

Table 5.

**V<sub>2</sub> form of *lase* in texts of different decades  
(TÜKK, total in press and fiction; absolute figures)**

V <sub>2</sub>	1890s	1900s	1930s	1950s	1960s	1970s	1980s
<i>da-inf.</i>	16	1	4	6	7	3	4
<b>Finite form</b>	2	1	2	0	0	0	0

The language of recent years in the EKI corpus reveals the same for all the sublanguages represented (Tables 7, 8, 9). Unlike 8 instances of *las* in TÜKK, the EKI corpus rather frequently displays the construction *laske* + V<sub>2</sub>, wherefore quite traceable is the general imperative use (variant of use 1) thereof, besides a certain use of particle (Table 9).

Table 6.

**The volume of component parts in the EKI corpus (mB)**

Dailies	75
Magazines	3
Fiction	8
Spoken language	0.5

Table 7.

**V<sub>2</sub> form of *las* in texts of different sublanguages of Modern Estonian (EKI corpus, absolute figures)**

V <sub>2</sub> form	Dailies	Magazines	Fiction	Spoken language
<i>da</i> -inf.	1	2	4	1
Finite form	3	8	24	5
<i>las</i> total	4	10	28	6

Table 8.

**V<sub>2</sub> form of *lase* in texts of different sublanguages of Modern Estonian (EKI corpus, absolute figures)**

V <sub>2</sub> form	Dailies	Magazines	Fiction	Spoken language
<i>da</i> -inf.	2	0	32	8
Finite form	0	0	2	0
<i>lase</i> total	2	0	34	8

Table 9.

**V<sub>2</sub> form *laske* in texts of different sublanguages  
of Modern Estonian (EKI corpus, absolute figures)**

V <sub>2</sub> form	Dailies	Magazines	Fiction	Spoken language
<i>da</i> -inf.	35	3	42	3
Finite form	1	0	4	0
<i>laske</i> total	36	3	44	3

### 2.2.2. Reflections on *las*- and *lase*-constructions per periods

The first instance of employment is prevalent in texts of the 1890s, where the construction *lase* + V<sub>da</sub> has a dynamic or causative meaning. It mostly takes the total object in the nominative case, which thus resembles the grammatical subject and contributes to reanalysis (i.e. the object is interpreted as the subject) (52), cf. (30). In *las*-phrases, in both forms of the main verb the meaning is dynamic (54), concessive (53) and jussive (55). On the *lase*-basis, there are mainly set phrases like *lase olla* 'let it be' *las ta olla* 'let him be', *lase aga olla* 'let it stay like that', *las käia* 'keep going' *lase kuulda* 'let us hear' *las räägivad* 'let them speak'

(52) *Lase Hilda siia tull-a!* ILU1890\pro0053  
let:IMP Hilda:NOM here come-daINF  
'Permit Hilda to come here'

(53) *Las räägiwad, kes ilma suud jõuab sulguda, arwas Mikk rahulikult.* ILU1890\pro0043  
'Let them speak, none too powerful to shut up the world'

(54) *Las' neiu ise ütleb, kumba ta meist oma kaastantsijaks soovib! rääkis Muru perepoeg pahaselt.* ILU1890\pro0023  
'Let the maid herself say whom among us she wishes to have as her partner to dances'

(55) *Kodurahwa pärast nüüd – las' oodata, löi Jaak käega.* ILU1890\pro0085  
'No need to worry for the members of the household – just let them wait, Jaak brushed off'

The construction under survey is rather rare in the language of the 1900s.

It is appropriate here to refer to the examples by H. Leskinen from Estonian folk songs, where *las* bears mainly the meaning of concession. For example, from Haljala: *Las tuleb rahu meie maale: Siis paned hära härgadeksa* (Leskinen 1965: 401) ‘Let peace reign in our land: I will set the landlord under the yoke, instead of the oxen’ This is valid proof of the process of grammaticalization in the older popular language, too.

The deontic – concessive (56) or jussive (57) *las* predominates in the *las*-constructions of the 1930s. It is a clear-cut particle that is attached to the finite verb form. In *lase*-phrases the main verb occurs in both forms, the meaning being dynamic (58). It is hard to account for the peak recurrence of *las*-constructions over this period. A certain similarity to the German word *laß* could have been a valid reason for people to reject the *las*-word during that anti-German period. It might well be that the *lase*-phrase coinciding with the German pattern was perceived as Germanic, unlike the reanalyzed *las*-construction. Proponents of the language reform, such linguists as Aavik and Loorits, failed to discuss the problem in their grammars.

(56) *las mets põleb päälegi, mis niisugused asjad temasse puutuvad.* ILU1930\lmg0010

‘Let the woods burn, such things have scarcely any bearing on him’

(57) *Peitsin enda pärna harude vahele ja mõtlesin, et kui ma suureks saan, siis ostan isale ka neli hobust ja uhke tõlla, las ta siis sõidab selles nagu Rannu härragi.* ILU1930\lmg0003

‘I hid between the branches of the linden and thought that when I grow up, I would buy to my father, too, a four-in-hand; let him then ride in it like the Lord of Rannu’

(58) “*Sa lase ometi inimesed soojendavad natuke, õlut sul muidugi ei ole!*” ILU1930\nov0049

‘You’d better let people warm themselves up a bit, lacking beer as you are!’

Texts of the 1950s reveal occasional examples of the constructions under discussion. They can be divided into two groups: the deontic particle construction with the finite verb (60) and the imperative construction, mainly dynamic (59).

(59) “*Oled kõnelnud, lase teisel ka rääkida!*” ütles Batmanov naljatleval vihaga ja lükkas peainseneri kõrvale. AJA1950\sv0005

‘You have had your say, let the others speak up, too!’ said Batmanov with jocular anger and pushed the chief engineer aside’

(60) “*Las sõjaõhutajad krutivad, kuidas oskavad,*” arvas Felipe, “*aga võidu võti on siiski meie peos.*” ILU1950\ilu0026

‘Let the warmongers twist as best as they can, reasoned Felipe, we have the key to victory in our grasp, nevertheless’

In the 1960s the use of *las* increased both in its dynamic and deontic uses, mainly with the finite verb, however (61).

(61) “*Ma lõin käega, et las rahmeldavad, nende oma asi.*” AJA1950\rh0102

‘I dismissed it, saying that let them fuss and bustle, it is their problem’

The 1970s witnessed an upturn of the *las*-particle, connected with deontic and finite verbs, e.g. (62). In the dynamic meaning *las* occurs mainly in connection with the first person (63). There are few *lase*-constructions.

(62) *Tiina jättis Tommi silitamise järele, aga ema ütles, et las mehed lähevad käivad ennem ära.* ILU1970\ilu0075

‘Tiina stopped stroking Tommi, however, Mother said that let men be the first to go’

(63) *Ma saan aru küll, kui tatine see palve on, aga vaata mind, ma olen ju endast väljas ega anna endale aru, ja las ma sellepärast palun sind, et sa mind aitaks.* ILU1970\ilu0093

‘I can fully see how mean my request is, however, look at me, I’m totally upset and cannot come to grips with the situation, and let me therefore beg you to help me’

The jussive or concessive use as a true particle continued to prevail in the 1980s (64), (65); there were few imperative constructions, occurring mainly as set phrases (*lase olla, lase käia*).

(64) *Ühel õhtul läks ta pojale külla ja ilma pikemalt keerutamata ütles välja, et tahab televiisorit osta ning las poiss tuleb ja aitab tal valida.* ILU1980\stkt0069

‘One evening he went to visit his son and said outright, without much ado that he wanted to buy a TV set, and that let the boy come and help him pick the right one’

(65) *Ei, las talved olla pehmed ja sulad.* ILU1980\stkt0101

‘No, let the winters be mild and with lots of thaw’

The number of constructions with *laskma* in the imperative form of the other persons (*laske* ‘let you’ *laskem* ‘let us’, *laseme* ‘let us’) turned out to be negligible per the whole century, amounting to only 12 instances. There was only one instance from the 1930s that had the main verb in the finite form (66).

(66) *Las-ke ma too-n.* ILU1930\nov0052  
let-IMP:2PL I:NOM bring-PRS:1SG  
‘Let me bring’

Judging by the EKI corpus, *las* is employed predominantly as a particle, in conjunction with the finite verb. The meaning is jussive (67) or concessive (68); however, there are some instances of dynamic use, too (69).

(67) *Kotka vastus oli lühike: lendaks küll, aga ei tunne teed. Lehvitas tiibu ja kadus pilvede hõlma. Siis ütles purustatud inimene, kutsugu raisakull, las ta raiub südame mul rinnast välja.*

‘The answer by the eagle was brief: I would fly; however, I don’t know the way. The eagle slapped with wings and got lost in the flap of the clouds. The crushed and desolate man then said, may the vulture be invited, let him carve out the heart from my breast’

(68) “*Kas mina olen irooniline? Ma ei tea. Tänapäeval väidetakse, et kõik on irooniline. Ei tea. Las inimesed ütlevad mida tahavad. See on ju vabadus – või kuidas?*”

‘Me ironic? I doubt it. It is claimed today that everything is ironic. I don’t know. Let people say what they will. This is freedom – or isn’t it?’



(69) "Palun, isa," anus ta peaaegu pisarsilmil. "Las ma jään siia. Mul on kahju."

'Please, father, he begged close to tears. Let me stay here. I'm sorry'

The *lase*- and *laske*-constructions in the EKI corpus are predominantly imperative source constructions (70), (71) or set collocations *lase / laske tulla*, *lase/ laske käia*, *lase / laske olla*, *lase / laske kuulda*.

(70) *Lase valmis kohvijoogil paar minutit seista!*

'Let the coffee drink made brew for a couple of minutes'

(71) *Katsuge mitte jääda liiga tõsiseks, laske ka oma tunnetel välja paista.*

'Try to not remain too serious, allow your feelings to be exposed, too'

## Conclusion

Based on the underlying second person singular imperative form *lase* 'let' the grammaticalization of the particle *las* is in progress in Estonian. It carries the agent-oriented – dynamic or deontic modal meaning. Although the process was initially affected by the use of the imperative form of the German auxiliary *laß* in conjunction with the Infinitive, the use was then channelled to the domain of concessive and jussive meanings, which can be regarded as a special development in Estonian. The semantic-pragmatic role structure serving as the background to the use of this Estonian particle has retained the integration and complexity, inherent to the source structure. The Estonian *las*-particle has no tendency to express a direct order, unlike the English word *lets*. Development of modality-oriented means of expression into mood-oriented means of expression is a common grammaticalization tendency in the world's languages. As it is, the mood-oriented particle is also a manifest phenomenon of the analytic tendency in Estonian grammar.

On the one hand, one can draw a clear demarcation line between the revealed structural centres by the source structure. The verb form *lase* governs the *da*-infinitive of the main verb, and the object or adverbial complement denotes the actor of the main

verb (*lase poisil õppida* ‘letthe boy to study’). On the other hand, the structure is derived as a result of reanalysis, where *las* acts as a particle and the main verb or the Predicate moulds the sentence (*las poiss õpib* ‘let the boy study’). In actual use there are also several mixed structures (*lase poiss õpib* ‘let the boy study’, *las poiss õppida* ‘let the boy study’ etc.).

The tentative impression of the formation of the particle in Literary Estonian is quite incomplete, reflecting rather the vernacular of the writing profession than that of the actual speakers. Starting with the 17<sup>th</sup> century there are examples of the German-patterned source structure, however, also of the particle *laß*, identical with its German counterpart. In the period from the 1890s to the 1990s *las* as a particle and *lase* as an imperative have continually striven to differentiate themselves from each other. The use of the particle became more recurrent. The 1890s still evidenced the domination of the Germanic imperative phrase, the 1930s witnessed an increase in the use of the particle, evidently to counterbalance the Germanic use. During the post-war period the use of the particle declined in the literary language, conceivably due to its spoken-language flavor. Later the use of the particle became more frequent again. Both phrases form set collocations (*las olla*, *lase tulla*); however, in all periods the *lase*-phrases have outnumbered the *las*-phrases in that function.

## ABBREVIATIONS OF GRAMMATICAL TERMS

1	1st person	NOM	nominative
2	2nd person	PFV	perfective
3	3rd person	PL	plural
ADESS	adessive	PRS	present
ALL	allative	PRTV	partitive
daINF	da-infinitive	PST	past
GEN	genitive	SG	singular
IMP	imperative	V <sub>da</sub>	verb in da-infinitive
IMPS	impersonal	V <sub>fin</sub>	verb in finite form
JUSS	jussive	V <sub>inf</sub>	verb in infinitive form
NEG	negation		

## REFERENCES

- Bybee, Joan; Perkins, Revere; Pagliuca, William 1994:** The evolution of grammar. Tense, aspect and modality in the languages of the world. The University of Chicago Press. Chicago and London.
- Duden, Bd. 7** Etymologie. Herkunftswörterbuch der deutschen Sprache. Bibliographisches Institut Mannheim/Wien/Zürich. Dudenverlag. 1963.
- EEW = Julius Mägiste, Estnisches etymologisches Wörterbuch IV.** Helsinki. 1955
- Ehasalu, Epp; Habicht, Külli; Kingisepp, Valve-Liivi; Peebo, Jaak 1997:** Eesti keele vanimad tekstid ja sõnastik. (Tartu Ülikooli eesti keele õppetooli toimetised 6.) Tartu Ülikool. Tartu.
- EKI = Eesti keele Instituudi tekstikorpus.** (Text Corpus of the Institute of the Estonian Language.) <http://www.eki.ee/corpus/>
- Göseken, Heinrich 1660:** Manuductio ad Linguam Oesthonicam. Anführung zur Östnischen Sprache. Reval.
- Gutslaff, Johannes 1648:** Observationes grammaticae circa linguam Esthonicam. Dorpat.
- Hakulinen, Auli & Seppänen, Eeva-Leena 1992:** Finnish *kato*: From verb to particle. – Journal of Pragmatics 18. 527–549.
- Hasselgård, Hilde; Johansson, Stig; Lysvåg, Per 1999:** Glossary of grammatical terms used in English Grammar: Theory and Use. <http://www.hf.uio.no/~hhasselg/terms.html>
- Hennoste, Tiit; Keevallik, Leelo; Pajusalu, Karl 1999:** Introduction. – Estonian Sociolinguistics. International Journal of the Sociology of Language, vol. 139. 1–16.
- Hopper, Paul J., Traugott, Elizabeth Closs 1993:** Grammaticalization. Cambridge University Press. Cambridge.
- JD = Johnson's Dictionary. A Modern Selection** by E.L. McAdam, Jr. & George Milne. Pantheon Books. NY 1963.
- Leskinen, Heikki 1965:** “Lahka Lalloi kotihin saapi” – Virittäjä, vol. 69. 399–402.
- Metslang, Helle (forthcoming a):** Analytism and synthetism in the development of the tense and aspect systems of literary Estonian. – Grammaticalisation aréale et sémantique cognitive: Finnic and Sami. Paris.
- Metslang, Helle (forthcoming b):** On the Developments of the Estonian Aspect: the Verbal Particle *ära*. – Östen Dahl, Maria Koptjevskaja-Tamm (eds.), Circum-Baltic volume. Benjamins. Amsterdam.
- Metslang, Helle 1997:** On the use of the Estonian past tense forms during the last century. – Mati Ereht (ed.), Estonian: Typological

Studies II. (Publications of the Department of Estonian of the University of Tartu 8.) Tartu. 98–145.

**Rätsep, Huno 1971:** Kas kaudne kõneviis on kõneviis? Verbivormide situatsioonianalüüsi. – Keel ja struktuur 5. Tartu Riiklik Ülikool, eesti keele kateeder. Tartu. 43–69.

**SC = Albert Saareste, A.R Cederberg:** Valik eesti kirjakeele vanemaid mälestisi a. 1524–1739. (Akadeemilise Emakeele Seltsi Toimetised XVI.) Tartu. 1925–31.

**SSA = Suomen sanojen alkuperä.** Etymologinen sanakirja. 2, L–P. SKS, KTKK. Helsinki 1995.

**Stahl, Henricus 1637:** Anführung zu der Ehstnischen Sprach. Revall.

**Tauli, Valter 1983:** Standard Estonian grammar. Part II. Syntax. (Acta Universitatis Upsaliensis. Studia Uralica et Altaica Upsaliensia 14.) Uppsala.

**Tragel, Ilona 1999:** Verbi *andma* kognitiivlingvistiline analüüs. Tartu Ülikool. (MA thesis, University of Tartu.)

**TÜKK = Tartu Ülikooli eesti kirjakeele korpus.** (Corpus of Estonian written texts.) <http://www.cl.ut.ee/ee/corpusb/>

**Uuspõld, Ellen 1989:** Modaalasusest ja modaalsest predikaadist eesti keeles. – Keel ja Kirjandus, no 8. 468–477.

**WAKK = Wana kirjakeele korpus.** (Corpus of Old Estonian written texts.) <http://ee.www.ee/filosoft/wakk/>

## ÜHE EESTI PARTIKLI ARENGUID

### Helle Metslang

Imperatiivivormi *lase* põhjal on eesti keeles grammatikaliseerumas partikkel *las*, mis kannab agendile orienteeritud – dünaamilist või deontilist modaal tähendust, nt *Las Liina ootab sind*, *Las Ott ootab ukse taga*. Kuigi partiklistumine on võinud saada algmõjutusi saksa abiverbi *lassen* imperatiivivormi *laß* kasutusest koos infinitiiviga (*Laßt uns gehen!*), on edasine kasutus suundunud möönvasse ja jussiivsesse tähendusesse, mis on juba eesti eriareng. Eesti partikli kasutuse taustaks olev semantilis-pragmaatiline rollistruktuur on säilitanud lähtestruktuurile omase kompleksuse; eesti *las*-partikliil ei ole tendentsi väljendada otsest käsku nagu näiteks inglise sõnal *lets* (*Let's go! Lets you go first, then if we have money left I'll go* – Hopper, Traugott 1993: 13). Modaalasuse väljendusvahendite areng kõneviisivahenditeks on üks paljudest maailma keeltes tuntud grammatikalisatsioonitendentsidest. Kõneviisitähendust kandev partikkel on ühtlasi üks eesti grammatika analüütilisuseilminguid.

Kuigi selgelt eristuvad ühelt poolt struktuursete tsentritena lähte-  
struktuur, kus imperatiivivorm *lase* alistab põhiverbi *da*-infinitiivi ja  
põhiverbi subjekti tähistava objekti või adverbiaali (*lase poissi / poisil  
õppida*), ja teiselt poolt reanalüüsi tulemusena saadud struktuur, kus *las*  
on modaalne või jussiiivne partikkel ja põhiverb lauset kujundav predi-  
kaat (*las poiss õpib*), on tegelikus kasutuses ka mitmeid segastruktuure  
(*lase poiss õpib*, *las poiss õppida* jne).

Esialgne mulje partikli kujunemisest eesti kirjakeeles on üsnagi eba-  
täiuslik ja peegeldab ehk rohkem kirjaseadjate kui tegelike keelekõnele-  
jate keelepruuki. XVII sajandist on näiteid saksamallilisest imperatiiv-  
sest lähtestruktuurist, samas ka saksa imperatiivivormiga samakujulisest  
partiklist *laß*. Ajavahemikul 1890ndaist 1990ndate aastateni on *las*  
partiklina ja *lase* imperatiivina teineteisest järk-järgult eristuda püüdnud,  
kusjuures partikli kasutus on sagenenud. 1890ndail aastail domineeris  
veel saksapärase imperatiivitarind, 1930ndail ehk saksapärasusele vastu-  
kaaluks partiklikasutus. Sõjajärgsel perioodil partikli kasutamine kirja-  
keeles vähenes, ehk selle kõnekeelse nüansi tõttu. Hiljem on partikli-  
kasutus taas tõusma asunud. Kinnisühendeid moodustavad molemad  
tarindid (*las olla*, *lase tulla*), siiski esineb neid kõigil perioodidel rohkem  
*lase*-tarindeina.

# INDEFINITE DETERMINERS *mingi* AND *üks* IN ESTONIAN<sup>1</sup>

Renate Pajusalu

University of Tartu

## 1. General remarks

When we speak we refer to people, things, and phenomena. I will use the term 'entity' for all of them regardless of whether they exist in the real world or in the world of discourse. Thus, an entity is any discrete unit that can be referred to (Lyons 1977: 442–445, see also Pajusalu 1999: 19–23). Referring to the world's entities (referential act) is the possibility to relate one's speech to the speech-external world.

In addition to a referential use, a linguistic expression may have an attributive use. An attributive use does not refer to the entity itself but provides its description. The difference is first and foremost pragmatic, i.e. it depends on the speaker's intentions and need not be unambiguously distinguishable on the basis of the linguistic form. Yule (1996) explains the pragmatic difference between the referential and attributive uses by means of an example sentence (1) as follows:

(1) He wants to marry **a woman with lots of money**.

*The expression 'a woman with lots of money' can designate an entity that is known to the speaker only in terms of its descriptive properties. The word 'a' could be replaced by 'any' in this case. This is sometimes called an attributive use, meaning 'whoever/whatever fits the description'. It would be distinct from a referential use whereby I actually have a person in mind and, instead of using her name or some other description, I choose the expression 'a woman with lots of money, perhaps because I think*

---

<sup>1</sup> This work was supported by the Estonian Science Foundation, Grant No.: 4405. I would like to thank Mati Ereht and Liina Lindström for their helpful comments on a previous draft of this paper.

*you'd be more interested in hearing that this woman has lots of money than that she has a name.* (Yule 1996: 18.)

On the one hand, the referential use of an expression can be specific or non-specific, on the other hand, it can be either definite or indefinite. In the case of a specific use the speaker refers to some concrete entity. A non-specific NP, however, refers to any representative of the mentioned quantity. The specificity or non-specificity can be judged only on the basis of the context of the utterance; the Estonian language has no special means for this. In other languages of the world the determiners that indicate specificity are fairly widespread (Lyons 1999: 59).

In the case of a definite use the speaker assumes that the entity is unambiguously identifiable to the interlocutors, either because it had been mentioned earlier or by the means of common knowledge shared by them. The familiarity of an entity is often mentioned as a definite meaning, but that is not quite true. A definite NP can also refer to an entity that is unfamiliar at the moment of speaking but can be identified. For example, when hearing the sentence *Give me the hammer!* the hearer need not even know at the moment of speaking that there is a hammer somewhere nearby, but NP *the hammer* serves as an indication that it can be found, and thus it is identifiable. In certain cases definiteness would rather carry the meaning of uniqueness or wholeness. For example, NP *the car she just brought* refers to the car as the only one and *the shampoo I put here* as a whole, whereas neither of them has to be identified as an exemplar. In such cases the definite NP indicates that there is a concrete car and an amount of shampoo that match the descriptions (Lyons 1999: 2–15).

An indefinite expression points to the novelty of an entity at the moment of reference, whereby it is often the first mention of the entity in the context. In some English contexts it is better to consider non-uniqueness as the meaning of the indefinite article because the English *a* may signal that the same NP can also refer to other entities in the given context, for example, in the sentence *pass me a hammer* (Lyons 1999: 12–13). Estonian makes no formal difference between the numerical and indefinite uses, therefore I would regard *üks* in the sentence *anna mulle üks haamer* as a numerical use.

Many languages of the world distinguish definite and indefinite expressions by means of different articles. Estonian, as well as some other Finno-Ugric languages, has been traditionally regarded as non-article languages (with the exception of Erzya-Mordvin and Moksha-Mordvin, where the definite article has developed independently in the form of inflectional material on the head noun, Lyons 1999: 49). However, they have their own means for the expression of (in)definiteness as well. The function of the definite article can be performed by the demonstrative pronouns: in Estonian the pronoun *see* (Pajusalu 1997), in Finnish *se* (Laury 1991, 1997), and in South Estonian *tuu* (Pajusalu 1998: 167–168).

The linguistic shape of the indefinite NP has not been studied in Estonian as yet. It is, however, a well-known fact that in Standard Estonian the numeral *üks* 'one' can occur in the position of the indefinite article. In this respect Estonian is similar to many other languages where the numeral 'one' and the indefinite article either fully coincide (German *ein*, French *un*, Turkish *bir*, Tagalog *isa*) or the latter has developed from the former (English *a*, Kurdish *-ek*) (Lyons 1999: 95). The Estonian numeral *üks* is likely to become an article because of its phonological shape as well because the majority of articles in the world's languages are monosyllabic morphemes (Lyons 1999: 64).

The recordings of ordinary conversations enable us to consider also the pronoun *mingi* 'some, a certain' as a potential indefinite determiner. In principle, the pronoun *keegi* 'some(body)' is another possible candidate for an indefinite marker. However, it rarely occurs in spoken language and therefore it will not be discussed in the present article. Saareste's conceptual dictionary, which is one of the most important lexical collections, does not explain the non-numerical senses of the numeral *üks* in greater detail, but it points out synonymy with the indefinite pronouns *keegi* (*Kuule, üks tuleb!* 'Listen, someone's coming! – Saareste 1958: I, 671) and *keegi/mingi* (*üks mees tuli talle vastu* 'a man came approached him from the opposite direction'; *nägi kaugel üht maja* 'he saw a house in the distance' – Saareste 1958: I, 200). In the previous two example sentences *üks* does not perform as much the function of the numeral as the function of the indefinite pronoun. "A Grammar of Estonian"



(1995) lists *üks* and *mingi* as indefinite pronouns, whereas there is a crucial difference between them. *Üks* can be either a pro-substantive (see *üks tuleb* ‘someone’s coming’ above) or a pro-adjective, but *mingi* can only function as a pro-adjective (EKG I: 31). The Dictionary of Written Estonian provides two senses for the word *mingi*: indefinite pronoun and a particle that reinforces negation (*selles polnud mingit kahtlust* ‘there was no doubt about it’ – EKS III: 444). As in recent years linguists have taken an interest in spoken language, they have noticed that the pronoun *mingi* has become a colloquial particle that can indicate any degree of vagueness and approximation (Erelt 1998: 821).

Below I will analyze the functions of the pronouns *üks* and *mingi* as a part of the noun phrase on the basis of everyday conversations in the corpus of spoken Estonian of the University of Tartu. For comparison I will study also the indefinite determiners *üks* and *mingi* in the corpus of written Estonian of the University of Tartu. The main emphasis, however, will be on their pragmatic functions in conversation because the dynamics of the written text would additionally require a textual approach. I will call a noun phrase that includes the pronoun *üks* as *üksNP* and the noun phrase that includes the pronoun *mingi* as *mingiNP*.

## 2. *mingi* ‘some, a certain’

### 2.1. *mingi* in everyday conversations

There were 133 occurrences of the pronoun *mingi* in the everyday conversations of the spoken corpus. Four groups of functions can be distinguished in its usage: 1) reinforcement of negation in the negative sentence (as in written language), 2) expression of the vagueness of the amount expressed by the numeral (together with the numeral), 3) presentation of the referential NP as an indefinite one and de-concretization of the non-referential NP. 4) particle. To begin with, we will discuss the first two groups that can grammatically be delimited more easily, and we will then discuss Groups 3 and 4 that have fuzzier boundaries. For the present article the most important one is Group 3, which expresses indefiniteness and approximation.

**2.1.1.** The pronoun *mingi* can reinforce negation both in written and spoken language, and it has no important differences in

comparison with the written language within the framework of this function. *Mingi* can be an obligatory component of a negative fixed expression (e.g. *ta ei tule (mitte) mingil juhul* ‘there’s no way he will come’, or it can be omitted freely (*tal ei ole (mingeid) probleeme* ‘he has no problems whatsoever’). Line 5 of example (2) is a case between these two extremes: in the exchange *see pole mingi iga* ‘that’s no age at all’ there is no need for the pronoun *mingi* from the point of view of formal logic. The sentence would be negative anyway and it would be understood without the pronoun. In this context, however, such negative reinforcement is almost obligatory, resulting in a fixed phrase that expresses the paucity of age in the given context. In more general terms one could state that the construction *see pole mingi X* ‘it’s no X whatsoever’ expresses a negative evaluation (in the meaning ‘little’ or ‘bad’) of the conceptual norm of the grammatical subject that is expressed by predicative X in the given context (*see pole mingi raha / söök / pidu* ‘that’s no sum / food / party at all, that’s bad or not enough’). The spoken corpus revealed 13 instances of such adjectival uses in the negative sentence.

- (2)<sup>2</sup> 1 E: ää:h (.) Mustamäel millal seal ehitama hakati. (4.5)  
 2 kuuekümnendatel ikka. (0.8)  
 3 Ka: ma ütsin=ju et kuuekümnendatel.  
 4 Kä: **mingi nelikend aastat** peaaegu=ja.(.)  
 5 Ka: no kuule, see pole **mingi iga** maja jaoks.  
 1 ‘E: Aah. In Mustamäe when they started to build there.  
 It must have been  
 2 in the sixties.  
 3 Ka: I told you that it was in the sixties.  
 4 Kä: **some forty years** ago almost and  
 5 Ka: well, that is **no age** at all for a house.’

2.1.2. The same example (2) includes also another pronoun *mingi* that is used differently from the written standard. In line 4 it

<sup>2</sup> The examples from the spoken corpus follow the same transcription (based on Jefferson’s transcription for CA) that was used for the corpus at the time of writing the article. Only some instances of false starts and feedback that were irrelevant to the observed phenomena were omitted. The bold face has been added to the phrase under discussion.

makes the numeral phrase vaguer. The spoken corpus included eight such occurrences. In these examples *mingi* is a grammatically independent adverb or particle because it does not agree with the numeral phrase. Example (3) includes two such uses.

(3) The topic of conversation is publishing books

- 1 L. [aga miks] miks =miks peab olema **mingi**  
 2 **neljasaja või seitsmesaja leheküljene.**  
 3 J: ei mitte neljasaja aga noh tavaline niukeste  
 4 raamatute paksus on kakssada viiskend kolm  
 5 pakun.  
 6 M: no aga [matemaatika]  
 7 A: [KUI PAKS] praegu tuleb see Kas jumal mängib  
 8 täringuid. (0.5)  
 9 J: aa muide tulebki: kolmsada viiskend kaks.  
 10 L: aga sa aga sa peaksid=ee katsuma kompressida  
 11 **mingi kahesajale.**

- 1 'L: But why, why, why should it have **some four**  
 2 **hundred or seven hundred pages**  
 3 J: No not four hundred pages but well usually such  
 4 books have two hundred and fifty three  
 5 I suggest  
 6 M: well but [mathematics]  
 7 A. [HOW THICK] will this be now Does the Lord  
 8 play dice.  
 9 J: Aah, by the way it will be three hundred and  
 fifty-two.  
 10 L: But you should eeh try to compress it into  
 11 **some two hundred.**

In Lines 1 and 11 the approximate number of pages is expressed by means of the particle *mingi*. The pragmatic purpose of numeral phrases 4 and 9 is to be accurate, therefore no de-concretizing word is added.

**2.1.3.** Indefiniteness need not consist in the inexactness of an amount, but consists in the vagueness of other circumstances as well. At first we will examine referential NPs to which a de-

concretizing determiner *mingi* has been added (43 occurrences). Example (4) includes two vague entities: the child and the job of the person under discussion.

(4) The conversation is about Aire, a longtime joint acquaintance.

- 1 J: jah. ta abiellus siin.  
 2 A: ei ma ütlen tal on **mingi: laps** juba:  
 3 J: jah.  
 4 A: nagu meie vanune varsti.  
 5 J: heh (0.8) ma nägin teda. Palta oli ka ballil. (1.0)  
 6 A: ta töötab E-kaubamajas vist. (.)  
 7 J: töötab=vä. (0.5)  
 8 A: jah. (1.0)  
 9 J: mm ta=enne=  
 10 A: =pere pere (.) ma=ei=mäleta, ta ütles küll **mingi**  
 11 vist on üleväl kuskil **mingi perepood**. on seal  
 12 **mingi perepood**.

[5 turns in the course of which the third speaker R recognizes the person being discussed]

- 13 issand selle laps on küll, (.) selle laps on  
 14 kümne=aastane minu=meelest (.) ma olen teda  
 15 üks=kord **mingi lapsega** näinud, täitsa noh  
 16 **mingi nii suur**.

- 1 'J: Yes, she got married here  
 2 A: No what I'm saying is that she's got a **child** already  
 3 J: Yes  
 4 A: Will be of our age soon  
 5 J: Heh, I saw her. Palta attended the ball as well  
 6 A: She must be working in the E-department store I think.  
 7 J: Is working, huh  
 8 A: Yeah.  
 9 J: Mm, before she  
 10 A: Family, family. I don't remember, she told me **some**  
 11 perhaps upstairs somewhere, **some family store**,

- 12        there is **some kind of family store**.  
 13    R:    Gosh, it must be her child, her child is  
 14        ten years old, I think I've seen her once together  
 15        with **some child**, well quite  
 16        **As big as this.**'

Line 2 informs us that Aire has a child, but relatively little is known about this child, as can be seen from the subsequent conversation. It is likely that because of the scarcity of information the child is referred to by means of *mingi*NP. Line 15 discusses the same child again, but it is not referred to by the expected definite NP. There are two possible reasons for this: on the one hand, the event that R is talking about (meeting Aire and her child) had happened earlier and its information structure is independent. At the moment of meeting the child was unfamiliar. In the case of this interpretation we are dealing with a deictic projection: the speaker takes the narration to the spatiotemporal axis of the scene of action. The data include other examples, where the beginning of the narrative refers to the person with *mingi*NP, thus emphasizing his or her obscurity at the moment of the event. However, there is another possible interpretation for the phrase *mingi lapsega* in line 15: R does not know whether it was actually Aire's child or not, and therefore she uses the indefinite determiner *mingi*.

The other vague entity is the shop where Aire works. In line 10 A begins to recollect which shop had it been in the department store. As she does not remember it exactly, she adds the pronoun *mingi* to the name of the shop in line 11. The independent *mingi* in line 10 could be either a pro-substantively used constituent or simply an incomplete NP constituent. In line 15 of example (4) *mingi* is a de-concretizing extension of a deictic adverbial phrase of degree. It is likely that the phrase may be accompanied by a gesture that points to the height of the child (the presence of the gesture is confirmed by the change in the tone of voice when this phrase is uttered).

In example (5) the word *mingi* has been used altogether six times. None of them occurs in a negative sentence or is an extension of a numeral phrase or an adverbial phrase of degree.

(5) [an everyday exchange discussing a visit to the police station]

1 käisin mina vahepeal Tartus, võtsin selle kviitungi sis  
 2 ära sealt, lähen=s mina tagasi politseijaoskonda, (.)  
 3 sel õhtul oli seal **mingi hullumaja täielik**. istusin  
 4 seal niimodi vaatasin, (.) **mingi rahvas** marsib  
 5 mööda nagu **mingi** ma ei tea  
 6 **hullumaja täielik** (.) ja aru ei saa, kes on politsei, kes  
 7 on pätt. (.) mõnest kohe  
 8 vaatad no se=on se=on küll mingisugune  
 9 **mingi kohutav**, se on **mingi kurjategija**.  
 10 (.) võtab võtmed taskust välja ja hakkab  
 11 **mingit kabinetiust** lahti tegema. politsei raisk. vahet  
 12 ei ole. **mingid niuksed tüübid** kakerdasid ringi,

1 'I went to Tartu in the meanwhile, I collected this  
 2 receipt then, I return to the police station,  
 3 **a total madhouse**, I was sitting there,  
 4 looking around like this, **some kind of crowd** is  
 5 marching by like **a** I don't know  
 6 **total madhouse** and you can't tell a cop from  
 7 a crook. When you look at  
 8 somebody, you'll realize that he must be  
 9 **something awful**, he must be **some criminal**.  
 10 Takes the keys from his pocket and starts to unlock  
 11 **some office door**. The damned police. There is no  
 12 difference. **Some types** were staggering around.

The expression *mingi hullumaja täielik* is repeated twice in lines 3–6 (although on the second occasion it is interrupted by the hesitation *ma ei tea* 'I don't know' there is no good reason to analyze the expression as two separate phrases). The phrase describes the referent under discussion *politseijaoskond* 'police station' and is therefore not referential. In line 9, too, the pronoun *mingi* is used in the non-referential phrases *mingi kohutav* and *mingi kurjategija*, which semantically form a single indefinite descriptive unit 'some awful criminal' However, the phrases *mingi rahvas* and *mingit kabinetiust* in lines 4 and 11 have a referential character because in the described situation they have their own referents that they refer to. These referents are new and unfamiliar, therefore we can say that these phrases are indefinite.

They also share the feature that there is no intention to talk about them in the future. They will not become characters in the story, they only refer to entities that are important from the point of view of describing the entire situation. Although the phrase *mingid niuksed tüübid* 'some kind of types' in line 12 refers to the previously mentioned people at the police station, their number is still uncertain and they are presented as indefinite. From the point of view of the narrative they are as unfamiliar and unimportant persons as *mingi rahvas* 'some crowd' in line 4. Thus, the people at the police station are not characters of the story but details that can be called indefinite independently of their real status on the unfamiliar/familiar axis.

The vagueness and unfamiliarity of the entity are revealed when the pronoun *mingi* is used together with a personal name. A proper noun in its primary use is always specific because one has in mind a concrete person. People, however, know each other to varying degrees. There are good friends and acquaintances but also people whom you do not know so well, and there are even those who are familiar because they have been mentioned by somebody. Examples (6) and (7) mention persons with *mingi*NP whose degree of familiarity we are unable to establish. It is clear that *mingi Tiit* or *mingi Heli* do not count as good friends of the speaker. In the case of Tiit we can learn from the conversation that his person is identifiable to both speakers. Heli, on the other hand, must be unknown to everybody with the exception of the story's character Priit.

(6) [the episode is about a birthday party at a student dormitory]

- 1 H: ei olnud. (.) puntratants oli. (.) tulen mina (.)  
 2 Mustlase toast vist tulin mina tuppaja (.) sis  
 3 **mingi altkorruse psühholoog mingi Tiit**, \*on ta  
 4 psühholoog=võ.\*  
 5 K: minu=arust on ta ajaloolane.

- 1 'H: No, it wasn't. There was a tangle dance. Then I  
 arrive.  
 2 I must have come from the Gypsy's room, I must  
 have entered the room, and then  
 3 **some downstairs psychologist some Tiit**,  
 4 is he a psychologist or

5 K: To my mind he's a historian.'

(7) [a young woman is talking about a young man with whom she had previously had intimate relations]

1 läksin kirikusse=ja (.) ta=isegi tere mulle=ei öelnud.

2 (.) j=sis=ku=ma=ütsin=kule kas ma saan suga paar

3 sõna rääkida ta=hakkas sinna taga ruumi minema.

4 (0.5) **mingi Heli** tahtis seal midagi, **mingeid asju**

5 nad seal koos ajasid. (.)

1 'I went to church, and he even didn't say hello to me.

2 And when I said listen could I have a word with you

3 he started to walk there to the back room,

4 with **a certain Heli** he wanted **something** there,

5 they were engaged in **some business** there.

In the referential phrases *mingi Tiit* (6) and *mingi Heli* (7) the pronoun *mingi* must have some other meaning than simple indefiniteness. As *Tiit* is a person known to everyone, the speaker may show her superiority by de-concretizing him. On the other hand, taking into account the fact that the interlocutors are female students, it may indicate a concealed interest in *Tiit*. In the case of *Heli* (7) one can perceive jealousy or just scorn on the part of the speaker. At the same time one cannot rule out the possibility that *Tiit* and *Heli* may be simply rather unfamiliar people. The common feature shared by both uses is that the pronoun *mingi* is not used in its primary meaning but various pragmatic meanings are added through the context, the final interpretation of which is impossible without knowing the speakers and the situation. It is of course impossible to draw up a complete list of such additional meanings because, first, the interpretation is subjective, and, second, the number of different contexts is infinite.

The previously discussed examples (4–7) were mostly specific indefinite phrases *mingi laps* 'some child' *mingi perepood* 'some family store' *mingi rahvas* 'some kind of crowd', *mingi kabinetiuks* 'some office door', *mingid niuksed tüübid* 'some types' With some reservations we could consider as indefinite also the phrases *mingi Teet* and *mingi Heli*. They all have a concrete referent that is presented for some reason as



indefinite. Example (8) contains a non-specific *mingi*NP (there were altogether 34 such occurrences).

(8) [An episode describing a visit to the police station, the speaker renders the police officer's words.]

1 üts mulle=et tal tõesti pole täna aega, et ta ilgelt

2 vabandab onju. aga et äkki ma saan homme

3 **mingil ajal** tulla, et sis ta teeb selle tõendi valmis

1 'He told me that he didn't have any time today, he

2 apologizes profusely you know, but perhaps I could

3 come **some time** tomorrow, he could then issue this certificate.'

The time expression in line 3 is non-specific because it refers to some moment that has not been determined as yet. The non-specific *mingi*NPs occurred mostly in those turns that described future events (*teeks mingi etenduse* 'would make a performance' *lubas teha mingi [kontroll]töö moodi asja* 'promised to give something like a test') or when the entire situation described by the sentence is generic and does not refer to any concrete event (*et võib mingi erinevus olla* 'that there can be some difference' *et mingi vahe on tehtud* 'that some difference has been made' *mingi isegi väike asigi võib palju maksta* 'even a small item may cost a fortune').

There were altogether 22 occurrences of non-referential *mingi*NPs. One of them occurred in line 3 of example (5): *mingi hullumaja* 'a total madhouse' does not refer to any referent but describes the entire situation. In example (9), too, *mingi omadussõna* 'some kind of adjective' performs a descriptive and not a referential function.

(9) [the episode is about a woman whose name is Marelle, but her nickname is Maru Ella]

R: kõik räägivad Maru Ella. (.) noh (.) ja väga loogiline, maru on **mingi omadussõna** ja Ella on nimi. (.) eile ma seal juba kaks korda ütlesin Ella.

'R: Everybody calls her Maru Ella. Well, and very logical, *maru* is **some kind of adjective** and *Ella* is a name. Yesterday I uttered there Ella twice already.'

2.1.4. The particle *mingi* in spoken language has developed from the uses that express indefiniteness. The particle bears no relation to the grammatical structure of the sentence, and it can be regarded as a de-concretizer or simply as a hesitation with the purpose of winning time. Our data contained 13 occurrences of the particle *mingi*.

(10) [The topic of conversation is hitchhiking, the feedback turns of the other interlocutors have been omitted ]

- 1 meeletu, mis eelmine reede oli, no=see=oli (.) ropp
- 2 lausa. (.) **mingi küm- kahe kilomeetri** ulatuses iga
- 3 (.) iga kümne meetri tagant **mingi** üks ilusam kui
- 4 teine beib, tead seisavad kõik onju=tead, (.)
- 5 ja **mingid hälvikud** ju sõitsid mööda
- 6 sealt=et=umbes=nii=et vaatavad välja, sõidavad veel
- 7 tagasi vaatab veelkord onju=no
- 8 **mingi täielik perverssed** vanamehed ausalt. /.../
- 9 kusjuures me saime see leeduka peale=ju.
- 10 (.) tead umbes niuke et gavarim parusski onju.
- 11 ja=minu vene keel on ju=seal=tead ma ütlen seal hlep
- 12 ja markovka juures onju.
- 13 no **mingi täitsa lõpp** lihtsalt noh.

- 1 'That was crazy last Friday, well that was in fact
- 2 filthy. Within a distance of **some ten two kilometres**
- 3 after every ten metres **a babe** of sorts, one more
- 4 beautiful than another, everybody's standing
- 5 you know and **some disturbed persons** drove past
- 6 there something like this that they look out, then
- 7 return and take another look
- 8 you know, well, **total** perverse old men honestly.
- 9 Whereas a Lithuanian gave us a lift, you know the
- 10 kind of person who let's speak Russian you know.
- 11 And my Russian is in fact there you know I can say
- 12 hlep and markovka
- 13 you know. **Some ultimate** simply you see.'

In line 2 *mingi* is grammatically unrelated to the rest of the sentence, and if it has any meaning at all, then it could be "everything that I say, is almost" Such a use is characteristic of

particles. Although *mingi* can occur alone as well, in spoken language people often prefer to use an absolutely vague phrase together with some other word as in lines 8 and 13 (*mingi täielik* 'some complete' *mingi täitsa lõpp* 'some ultimate') or with a pro-form. The combinations *mingi sihuke* 'some kind of' and *mingi nimodi* 'something like this' (11) that express indefiniteness occurred in our data a number of times.

(11) [the interlocutors speak badly of a person who claims to know everyone; his words are being quoted]

- 1 jah=et (0.5) kui mina seal nõukogust tulin, onju
- 2 **mingi nimodi**=et=sinna nõukogusse ma
- 3 nüüd varsti lähen onju, koosolekud ja kõik mis tal on.

- 1 'Yes, so when I came from that board, isn't it
- 2 **something like this** that I'll go to that board soon,
- 3 you see, the meetings and everything that he's got.

Sometimes *umbes* 'about' which is more traditional and closer to the written language, is used in such contexts instead of *mingi* (example (10), line 10: *umbes niuke* 'the kind of').

## 2.2. *mingi* in written language

There were 703 occurrences of *mingi* in the one-million-word basic corpus of written Estonian. When compared with spoken usage it is immediately clear that there is no particle-like use and de-concretizing use of the numeral in written language. All the occurrences were divided into four groups on the basis of their function:

- 1) reinforcement of negation in the negative sentence (220);
- 2) de-concretization of the non-referential NP (79);
- 3) de-concretization of the referential NP
  - a) in the non-specific NP (238);
  - b) indefinite determiner in the specific NP (166).

In written language the pronoun *mingi* has a stronger meaning of vagueness or lack of knowledge than in spoken language. Also, the specific *mingi*NPs include the overwhelming majority of such instances, where the speaker refers to a clearly existent and specific entity, which is indefinite in the sense that it has not been

mentioned before. At the same time indefiniteness is not the primary meaning of *mingi*, it would rather reflect the speaker's inability to define the entity also for himself. In such cases the referred entities are, for example, states of mind (12) or 'unleashers' of intuitive decisions (13) or previously unknown phenomena that the speaker tries to place in an existing category (14).

- (12) **Mingis ebamäärases hirmus** seisis ta inspektori ees.  
 'He was standing in front of the inspector **in some vague fright**.
- (13) Seal istusid teksaspükstes noormehed ja neiud, keda **mingite märkide** järgi arvasin Moskva turistideks.  
 'Some young men and women in jeans were sitting there whom I took **by some signs** for tourists from Moscow.'
- (14) Jorma tantsis rõõmu pärast **mingit veidrat karutantsu**.  
 'Jorma was dancing **some kind of odd bear dance** with joy.'

Examples (12–14) refer to entities that are specific for the speaker because *mingi ebamäärane hirm*, *mingid märgid* and *mingi veider karutants* do exist in the space and time of the story, and NPs do not refer to them as members of any category as in the case of a non-specific NP. At the same time they are difficult to define and indefinite because they have not been mentioned before.

In spoken language *mingi*NPs often refer to persons; in written language they usually do not. However, we could find some sentences of this kind in the corpus of written language, for example, (15).

- (15) Ronisin, et see pall alla tuua, aga **mingid tüdrukud** hüüdsid all üksteisele: "Kas elus ahvi oled näinud?"  
 'I climbed to retrieve this ball, but **some girls** shouted below to each other, "Have you ever seen a living monkey?"'

In example (15) *mingi* means 'unknown to all of us, i.e. indefinite'. There is no good reason to think that it expresses only vagueness.

Thus, the pronoun *mingi* indicates indefiniteness as well, but in most cases it is not its primary meaning. Indefiniteness would rather accompany vagueness. At the same time it would be of

course possible to express the definite NP by using the article *see*, so that formally there are in Estonian semantic pairs, where one member is indefinite and the other one is definite, for example, *mingid märgid – need märgid* ‘some signs – these signs’ The corpus of written Estonian also included an example where such a pair occurred in the same sentence (16), so that the first mention was expressed as indefinite and the second one, as expected, definite.

- (16) Kogu see periood on mu mälusse jäänud väga lünklikult ning segaselt, ma elasin nagu omaenda lapseelu edasi ja samal ajal liikusin otsekui **mingis hoovuses** või liikus **see hoovus** minus, igatahes tean ma selgesti, et mind jälitas kogu aeg aimus millestki suurest ja ootamatust, mis varsti juhtub, kuigi mul polnud vähimatki ettekujutust, mis see peab olema.

‘The whole period has remained in my memory very sketchily and vaguely, I as if continued to live my own child’s life and at the same time I was moving as if **in some current**, or **this current** was moving in me. Anyway I know well enough that I was constantly followed by a premonition of something big and unexpected that would happen soon although I didn’t have the slightest idea what it should be.

A large number of non-specific *mingi*NPs in the corpus of written Estonian refer to a member of any category in general statements (17). Such examples were especially numerous in popular scientific texts.

- (17) Turusituatsioonis läheb **mingi väärtuse** edasiandmise eel lahti kauplemiseks: väärtust valdav pool püüab selle võimalikult kallilt loovutada.

‘In a market situation bargaining will begin before handing over **some value**: the party that possesses the value tries to get the highest price for giving it up.’

The corpus of written Estonian also included an example where a non-specific *mingi*NP was followed by an explanation *mingi* in this context (18). May it sum up the use of the pronoun *mingi* in written language.

- (18) Meis oli liiga vähe peenust, et toetuda **mingitele ideaalsetele väärtustele** (näete, ütlen mingitele –, sest ei oska neid praegugi nimetada!) ja vitaalsus mandus eneseirooniaks.  
 ‘We weren’t sophisticated enough to rest on **some ideal values** (I’m saying some, you see, because I cannot name them even now!) and vitality degenerated into self-irony.’

### 3. *Üks* ‘one’

#### 3.1. *üks* in everyday conversations

Another word that expresses indefiniteness is the indefinite pronoun *üks*. However, it was far less frequent in the everyday conversations of the spoken corpus than the pronoun *mingi*. There were in all 38 occurrences of *üks* in the position of the attribute, but some of the uses had a clearly numerical meaning. It is of course difficult, if not impossible, to make a distinction between the numerical and non-numerical meanings. It is true that in my data all the *üks* NPs corresponded to a single entity. It must be possible in spoken language, too, to use the pronoun *üks* as the attribute of a plural entity (*Ühed poisid tulid vastu* ‘certain boys were coming from the opposite direction’), but there were no such examples in the corpus. I considered as numerical use such *üks* NPs where *üks* was stressed or where the referent was ‘one of several’ (*veel üks mees* ‘another man’ in example 19). It is easier to identify such *üks* NPs where *üks* means ‘the same’ (e.g. *ühel ajal* ‘at the same time’ in example 19, also in the sentence *elan temaga ühes trepikojas* ‘I share **the same staircase** with him’).

- (19) minuga **ühel ajal** tuli kabinetti **veel üks ee mees**, kelle nimi oli Peeter Sink.

**At the same time** with me **another man** entered the office whose name was Peeter Sink.

Nor did I consider as indefinite an instance of *üks* NP where *üks* had the meaning ‘about’ (20). It is of course a very interesting parallel with the pronoun *mingi*, which indicates that *üks* and *mingi* share some other functions in addition to indicating indefiniteness. In written language this use is not considered to be acceptable, and it seems that in spoken language, too, this kind of vague *üks* can occur only together with a numeral.

- (20) 1 H: niivisi ja palju=s seda tailiha siis nii kangesti on ja  
 2 sääne noh ja mis on sis  
 3 M: kuule mamma niipalju  
 4 [ei jõuagi teha kui nad]  
 5 H: [**üks üks kaheksakümmend kilo liha**]  
 6 süüa siis suure perega. (.)
- 1 'H: this way and how much of this lean meat is there  
 and  
 2 well and what is then  
 3 M: listen, Ma so much  
 4 [you cannot make than they]  
 5 H: [**some some eighty kilograms** of meat]  
 6 for such a big family.

Typical cases of indefiniteness include first and foremost those referential specific *üks*NPs, where the pronoun *üks* indicates that the referred item is a new referent in the conversation. My data revealed 22 *üks*NPs of this kind. Usually, this kind of *üks*NP initiates the narrative, as in examples (21) and (22), where the story that follows describes a person who was introduced to the space-time of the narrative by means of the *üks*NP.

(21) [the speaker works in a hospital]

A: mõtle meil **üks=ee õde** käis öö=m Taanis suvel. oli tööl seal.

'A. think, at our hospital **a nurse** visited Denmark this summer. She worked there.'

- (22) 1 L: tahad ma räägin sulle tollest (.) lollist mehest,  
 2 tollest Pae Sassist.  
 3 T: no räägi. (1.0)  
 4 L: no et **üks selline** (.) **vana mees** elab seal (.)  
 5 Valgesoo külas.

1 'L: If you wish I can tell you about that stupid man,  
 2 about that Pae Sass

3 T: Go ahead.

4 L. Well, there lives **a certain old man** in that  
 5 Valgesoo Village.

In example (21) they actually begin to talk about an accident in a Danish water park. It is associated with the current topic of the entire conversation because the dangerousness of water parks has been under discussion. The person referred to with the phrase *üks õde* 'a nurse' is introduced as the source of the next item of information because the story originated from her. However, in example (22) the character of the narrative is actually mentioned. It is interesting to note that the same person was referred to in line 1 by means of the definite phrases *tollest lollist mehest, tollest Pae Sassist* 'about that stupid man, about that Pae Sass' However, the narrative that begins in line 4 is a whole for the narrator that has to be told in accordance with the 'rules' of the narrative, thus at first the character has to be introduced as indefinite.

While the majority of specific *üks*NPs refer to a new person (usually at the beginning of the narrative), then my data included also occasional specific indefinite *üks*NPs that referred to inanimate objects (23).

- (23) 1 T: kule=aga=kule mis kell on, ma pean tahan  
 2 televiisorit vaadata. ah aega veel on.  
 3 L: mis sealt kella või televiisoris on.  
 4 T: ah on üks multifilm jälle. (0.5)
- 1 'T: well, but, well, can you tell me the time, I'll have  
 to, I must  
 2 watch TV Ah, there's some time left  
 3 L: What's there the time or on TV  
 4 T: Ah, there'll be a **cartoon** again.

*Üks*NP can be non-referential as well. In example (24) Alma is described as an old-time schoolteacher, *üks* signals here belonging to elderly schoolteachers. In the case of a descriptive predicative it must be a common sentence pattern in spoken language: *X is a Y* The latter differs from the sentence *X is Y* in that the shade of meaning 'typical' is added (24).

- (24) 1 /.../ noh keegi neid kodus ei kasvata aga no=lihtsalt  
 2 vanematel ei=ole nh (.)  
 3 aega või tahtmist või on taa vabakasvatus nii aga Alma  
 4 on **üks vanaaegne kooliõpetaja**, tema tahab sedasi=et  
 5 (0.5) et laps midagi õpiks



- 1 'I.../ well, there's nobody to educate them at home,
- 2 but, well, the parents
- 3 simply have no time or will or it could be free education  
so, but Alma
- 4 is **an old-time schoolteacher**, she wants a child to learn
- 5 something.'

Another *üks*NP that describes typicalness of the same type occurred in the syntactic construction *Siis on üks X* (25) 'then there is a certain X' which is evidently the same syntactic pattern for spoken language although, according to traditional grammar, in one case X functions as the predicative and in the other case as the subject.

- (25) no=ja=sis s=on jälle **üks tüli ja karvade kitkumine**  
'Well and then there will be again a **kind of quarrel**'

### 3.2. *üks* in written language

In written language the frequency of *üks*NPs is roughly the same as that of *mingi*NPs. At the same time the one-million-word corpus included in all 705 *üks*NPs and numeral phrases including the numeral *üks* (and 45 occurrences of *üks* used as a pre-substantive). However, if we exclude 72 occurrences of *ki*-suffixed *ükski*NPs, which have a totally different meaning and are not discussed in the present article, and 9 prepositions *ühes* 'together' then we get 624 phrases. Three larger groups included *üks* in the numerical meaning (altogether 462 phrases): 1) 189 as a pure number (e.g. 26), 2) 190 in the meaning 'one out of many' (see 28–30 below), and 3) 83 in sentences where two or more entities were contrasted by means of the pattern *one X and another X (and a third X)* (e.g. 27).

- (26) Kolmest viimasest partiist vajab Kasparov üldvõiduks vaid **ühte punkti**.  
'Of the three final games Kasparov needs only **one point** for the overall victory.'

- (27) **Üks õpetaja** kirjutab märkusi kergekäeliselt, teine veenab õpilast suusõnaliselt, kolmas karistabki.  
 ‘**One teacher** writes complaints eagerly, another convinces a student orally, a third would even punish.’

The semantic group ‘one out of many’ is actually a transition between the numerical and article-like uses. Here *üks*NP refers to a single entity, but emphasis is not so much on the number as belonging to some larger quantity. In cognitive terms one could explain this difference so that the figure of the expression is the same, but the background is different. In the first case it is undetermined, in the second case it is in some way determined although it need not be identifiable as far as each detail is concerned. In the meaning ‘one out of many’ *üks*NP highlights its entity in the determined quantity.

One part of this group consists of *üks*NPs of a certain structural type:

- (28) *üks N (elat pl.)* ‘One N(elat pl.)’, ‘one of them’  
**Üks neist** on Mihhail Ülenurm.  
 ‘One of them is Mikhail Ülenurm.’
- (29) *veel üks N* ‘another N’  
 Lõpuks tahaksin kõnelda **veel ühest algatusest**.  
 ‘Finally I would like to talk about **one more initiative**.’
- (30) *üks N (pl. part)* ‘one N (part pl.)’, ‘one of them’  
 Teema poolest peaks see tulema **üks viimase aja probleemikamaid**.  
 ‘As far as the topic is concerned it should be **one of the most problematic ones** in recent times.’

However, the structure of the phrase is only one factor that speaks for the interpretation of numericalness. Usually the pragmatic function of *üks*NP becomes clear from the context. For example, in example (31) there is formally NP, which could express indefiniteness as well, but the syntactic context makes us realize that the author must have used *üks*NP because it is not the only entity of its kind, but there are more centres of this kind in Afghanistan.

(31) [1980s newspaper]

See toimus hiljuti Peshavari lähistel, **ühes Afgaani kontr-revolutsionääride diversiooni- ja terroriõppe keskuses Pakistani territooriumil.**

‘It happened recently in the vicinity of Peshawar, in a **subversion and terrorist training centre of Afghan counterrevolutionaries in Pakistani territory.**

The data of written language included 116 occurrences of *üks*NPs that allowed an indefinite interpretation. These are typically new persons who are unfamiliar to the reader (32) or other entities (33).

(32) **Ühe kohaliku elaniku** sõnade kohaselt oli heitlus niivõrd äge, et mürskude ja reaktiivrakettide killud tungisid läbi talumajade katuste kuni kilomeetri raadiuses.

‘According to **a local inhabitant** the fighting was so fierce that the fragments of shells and jet rockets penetrated the roofs of farm houses within the range of a kilometre.’

(33) **Ühes lahingus** sai neiu haavata.

‘In **a battle** the maid was wounded.

In case *üks*NP is not referential but descriptive, *üks* would rather add the meaning ‘typical’ (see examples 24–25 of everyday conversations) and thus make the meaning of the entire sentence generic.

(34) Matemaatika on **üks konkreetne asi**, millest kõik on teada.

‘Mathematics is **a concrete thing** about which everything is known.

(35) Kui **üks mees** niimoodi ütleb, siis arvatavasti püüab ka ise nõnda talitada.

‘If **a man** talks like this, then he is probably trying to act this way himself.’

There is another semantic dimension of the word *üks* that stays aside from the primary numerical meaning, where *üks* means identity (36). The longer and more idiomatic attribute *üks ja sama* ‘one and the same’ may have the same meaning.

(36) D. B. Belzoni (1778–1823) oli koloriitsemaid Egiptuse-uurijaid, kes tuli püramiidide maale **ühel aastal** Richteriga.

'D.B. Belzoni (1778–1823) was one of the more colourful Egyptologists who arrived in the Land of Pyramids **in the same year** as Richter.

Apparently, the somewhat outdated preposition *ühes* (37), which expresses a joint activity, became grammaticalized from the same meaning of identity. It can occur independently or in combination with the preposition *koos* 'together'. There were no occurrences of this preposition in the spoken corpus.

(37) Elasin täie hingega talle kaasa, põlesin ja kannatasin **ühes temaga**.

'I followed his life with all my heart, I burned and suffered **together with him**.'

The meaning 'same' has also given rise to a number of idiomatic expressions. The corpus of written language included the following examples: *ühel häälel vastu võtma* 'to adopt unanimously' *üks X puha* (the example was *üks küünla rasv puha* 'one and the same tallow anyway').

#### 4. *üks* and *mingi* – some similarities and differences

Table 1 presents statistical data about the comparable part of *mingi*NPs and *üks*NPs, that is the areas where these phrases have a similar meaning.

The cases that cannot be compared include:

- 1) *mingi*NPs in the negative context
- 2) *mingi*NPs as the particle;
- 3) referential non-specific *mingi*NPs;
- 4) *üks*NPs in the numerical meaning;
- 5) *üks*NPs in the meaning 'same'

Apparently referential non-specific phrases require a comment of their own. I identified as such 34 *mingi*NPs in everyday conversations and 238 in written language (see examples 8 and 17–18 above). In these cases the meaning of vagueness is very strong and indefiniteness is not expressed as clearly as in the case of specific NPs. The corresponding *üks*NPs, however, have such a clearly numerical meaning that there is no good reason to speak about definiteness. In example (38) *üks punkt* 'one point' has of

course a non-specific character, but does not provide enough ground to regard it an NP because the numerical meaning is prevalent here.

(38) Kolmest viimasest partiist vajab Kasparov üldvõiduks vaid ühte punkti.

'Kasparov needs only a **single point** from the three last games to gain an overall victory

Table 1.

**Occurrence of comparable semantic areas of *mingi*NPs and *üks*NPs in spoken and written corpora**

	mingi		üks	
	everyday conversations	written language	everyday conversations	written language
non-referential	22	79	4	8
referential specific	43	166	22	116
Rough with a numeral	8	–	1	–
Total	73	245	27	124

*Üks* and *mingi* are similar in their indefinite function, whereas in many cases they are apparently interchangeable. As the meaning of *mingi* carries more vagueness, there are no clearcut borders between the indefinite and the approximate meanings of *mingi*. The indefinite function can be seen more clearly in those contexts where it is known that *mingi*NP refers to a fully specific and self-evident entity, such as a person (*mingid tüdrukud* 'some girls' in example 14).

There is another shade of meaning where *mingi* and *üks* differ from each other by indefiniteness. By using *mingi*NP the speaker leaves an impression that the referred entity is unfamiliar to him as well, not only to the hearer. Thus, *mingi* is a determiner that implies total unfamiliarity (see *mingi lapsega* 'with some child' in example 4, line 15). *Üks* NP cannot communicate such an additional meaning. In this case indefiniteness really means

only that in the given discourse the entity is unidentifiable to the hearer.

As the pronoun *mingi* implies a strong tendency to vagueness and, accordingly, a certain degree of evaluation in spoken language, then the use of *mingi* enables the speaker to indicate his or her attitude toward the referred person (*mingi Heli* 'some Heli' and *mingi Tiit* 'some Tiit' in examples 6 and 7). Through emotionality the pronoun *mingi* has developed into a spoken particle. The pronoun *üks* has no evaluative connotations.

There is an interesting similarity in that both pronouns are used as adverbs meaning 'about'. By de-concretizing the numeral they can act as synonyms in spoken language (see *mingi nelikend aastat* 'some forty years' in example 2 and *üks kaheksakümmend kilo liha* 'some eighty kilograms of meat' in example 19). In other contexts only *mingi* can de-concretize the meaning.

On the one hand, the meanings of *üks* and *mingi* can be regarded in general terms as a semantic spectrum (Cruse 1982: 71–72) with the numerical meaning of *üks* and the vague meaning of *mingi* in the centres. Both centres branch into fields with fuzzy transitions.

*Üks* branches into two directions: first, in the meaning 'same' and, second, in the meanings 'one out of many' and 'indefinite'. On the other hand, assuming that the numerical meaning is primary, the same group of meanings could be examined by proceeding from the grammaticalization process. The part of speech of the stem changes first of all from the numeral into the pro-adjective, and then it becomes more grammatical: on one side it becomes the preposition and on the other side it becomes the indefinite adverb through the determiner. The final links of both ends of the chain have clear boundaries in the sense that the preposition *ühes* and the adverb *üks* are non-inflectional words.

PREP	← PROADJ	← NUM →	PROADJ	→ INDEF. DET	→ ADV
<i>ühes</i>	<i>üks</i>	<i>üks</i>	<i>üks</i>	<i>üks</i>	<i>üks</i>
'together'	'same'	'1'	'one out of many'	'a'	'about'

The semantic spectrum of *mingi* would rather branch into three directions: from the center 'vague' first, towards the

strengthening of negation, second, towards indefiniteness, and, third, towards the pro-adverb 'about' Both indefiniteness and the adverb 'about' are associated with the particle *mingi*. Here, too, the semantic transitions are fuzzy, but the part-of-speech difference sets the grammatical boundaries. However, whether the semantic field, as depicted in the scheme, corresponds to the historical grammaticalization process calls for additional research.

PROADJ← <i>mingi</i> 'strengthening of negation'	<b>PROADJ.</b> <i>mingi</i> 'vague, unclear'	→ INDEF DET→ <i>mingi</i> 'a'	PARTICLE <i>mingi</i>
		→ PROADV → <i>mingi</i> 'about'	

By way of conclusion one could say that there are two pronouns in Estonian – *üks* and *mingi* – one of their functions is to act as an indefinite determiner. The existence of two indefinite determiners with somewhat different meanings indicates once again that definiteness and indefiniteness do not stand in polar opposition. There can be various kinds of indefiniteness. Chesterman (1991: 182–183) has depicted definiteness as a scale, where one end is the most definite NP type and the other end is the most indefinite NP: null – the – a – some – zero. The indefinite article *a* remains in the middle of this scale. According to the same logic, the Estonian *mingi* would be more indefinite than *üks* because *üks* refers to a more concrete entity, which is specific at least for the speaker. However, conversational analysis reveals that we are dealing with a different type rather than a degree of indefiniteness. *Mingi* is indefinite for the speaker first and foremost through vagueness, the speaker does not know exactly how to describe *mingi*NP more precisely (see example 18). *Üks*, in contrast, is a determiner that creates the conversational structure that indicates the speaker's opinion concerning the identifiability of the entity for the speaker that is referred to by means of NP. It seems that *mingi*NP and *üks*NP often differ from each other by the status of the introduced entity: while *mingi*NP would rather introduce

insignificant details, *üksNP* would introduce the characters of the future narrative.

Another important theoretical conclusion is that sometimes definiteness/indefiniteness would rather act as carriers of pragmatic intention than the expression of the actual status of definiteness of an entity Laury has claimed (in an unpublished handout) that an entity can be presented as definite although it is actually not the case. In such a case the speaker's strategy could be worded roughly as follows, "I actually know that you cannot identify X, but I hope that you agree with this and believe me that it is a concrete, specific referent" The same can be said about indefiniteness. The speaker may present an entity as an indefinite one although he or she knows that the speaker is able to identify it. In such a case the referring strategy could be worded as follows, "I actually know that you can identify X, but nevertheless I'm going to present it as an indefinite one because it's only a detail of the story that I'm not going to concentrate on" or "...because it is the first mention of X in the present narrative" (see examples 5 and 22 and the following discussion, respectively).

The data of the present article do not enable us to consider either of them as established indefinite articles, first and foremost because of their low frequency (the same section of the spoken corpus revealed 375 occurrences of the pronoun *see* as the definite article).

## REFERENCES

- Chesterman, Andrew 1991:** On definiteness. Cambridge University Press. Cambridge.
- Cruse, D. A. 1986:** Lexical Semantics. Cambridge University Press. Cambridge.
- EKG I = Erelt, Mati; Kasik, Reet; Metslang, Helle; Rajandi, Henno; Ross, Kristiina; Saari, Henn; Tael, Kaja; Vare, Silvi 1995:** Eesti keele grammatika I. Morfoloogia, sõnamoodustus. Eesti Teaduste Akadeemia Eesti Keele Instituut. Tallinn.
- EKS III = Eesti kirjakeele seletussõnaraamat. III köide. Tallinn, 1992–1994.**



- Erelt, Mati 1998:** Mõnest tähendusnihkest tänapäeva eesti keeles. – Keel ja Kirjandus, no.12. 816–822.
- Laury, Ritva 1991:** On the development of the definite article *se* in spoken Finnish. – Suomen kielitieteellisen yhdistyksen vuosikirja. Helsinki.
- Laury, Ritva 1997:** Demonstratives in Interaction. John Benjamins PC. Amsterdam/Philadelphia.
- Lyons, John 1977:** Semantics 1,2. Cambridge University Press. Cambridge.
- Lyons, Christopher 1999:** Definiteness. Cambridge University Press. Cambridge.
- Pajusalu, Renate 1997:** Is There an Article in (Spoken) Estonian. – Mati Erelt (ed.), Estonian Typological Studies II. (Publications of the Department of Estonian of the University of Tartu 8.) Tartu. 146–177
- Pajusalu, Renate 1998:** Eesti pronoomeneid 2. Võru *sjoo, taa* ja *tuu*. – Keel ja Kirjandus, no. 2. 159–172.
- Pajusalu, Renate 1999:** Deiktikud eesti keeles. Dissertationes Philologiae Estonicae Universitatis Tartuensis 8. Tartu.
- Saareste, Andrus 1958–1963:** Eesti keele mõisteline sõnaraamat. Vaba Eesti. Stockholm.
- Yule, George 1996:** Pragmatics. Oxford University Press. Oxford.

## INDEFINIITSED MÄÄRATLEJAD *mingi* JA *üks* EESTI KEELES

### Renate Pajusalu

Eesti keelt on sarnaselt muude läänemeresoome keeltega traditsiooniliselt peetud artiklita keeleks. See aga ei tähenda, et eesti keeles puuduksid võimalused märkida NP definiitsust või indefiniitsust. Definiitse määratlejana funktsioneerib demonstratiiv *see* (Pajusalu 1977), lõuna-eesti keeles *tuu* (Pajusalu 1998). Käesolev artikkel püüab vastata küsimusele, kuid võrd väljendatakse eesti keeles indefiniitsust. Konkreetselt on vaatluse all indefiniitsed pronoomenid *mingi* ja *üks*. Materjalina on kasutatud eelkõige Tartu Ülikooli suulise kõne korpust, võrdluseks ka Tartu Ülikooli kirjakeele baaskorpust.

### *mingi*

Asesõna *mingi* esines suulise kõne korpuse argivestlustes 133 korda ja selle kasutuses võib eristada viit gruppi funktsioone:

- 1) eitavas lauses eituse tugevdamine (nagu kirjakeeleski), sealhulgas konstruktsioonides *see pole mingi X* – vt *see pole mingi iga* näite (2) real 5;
- 2) arvsõnaga väljendatud hulga umbmäärastamine, st selle ebatäpsuse väljendamine – vt *mingi nelikend aastat* ‘umbes 40 aastat’ näite (2) real 4. vt samuti näite (3) ridu 1 ja 11;
- 3) referentsiaalse NP indefiniitsena esitamine – vt näites (4) *mingi laps* ‘laps, keda me ei tunne’ real 2 ja 15 ja *mingi perepood* ‘pood, mida me ei tea’ real 11 ja 12, mõnikord ka koos pärisnimega nagu näidetes (6) ja (7), ja mittereferentsiaalse NP umbmäärastamine – vt näites (5) *mingi hullumaja* ‘umbes nagu hullumaja’ real 3 ja *mingi kohutav kurjategija* ‘enam-vähem nagu kurjategija’ ridadel 9, samuti *mingi omadussõna* ‘vist omadussõna’ näites (9),
- 4) partikkel – vt näite (10) rida 3, mõnikord kombinatsioonis teiste partiklilaadsete sõnadega nagu näites (10) *mingi täielik* real 8 ja *mingi täitsa lõpp* real 13.

Eesti kirjakeele baaskorpuses (1 miljon sõnet) esines pronoomen *mingi* 703 korral. Kõnekeelse kasutusega võrreldes on kõigepealt selge, et kirjutatud keeles puudub täielikult partikliline ja arvsõna ebatäpsust väljendav kasutus, palju on aga eituse tugevdamist. Teiselt poolt on kirjakeeles pronoomenil *mingi* tugevam ebamäärasuse või teadmatus tähendus kui kõnekeeles ja puhtalt indefiniitsust näitab see harvemini – vt näiteid (12–14). Pigem on kirjakeele *mingi*NPdes indefiniitsus ebamäärasusega kaasnev tähendus.

### üks

Kõnekeele korpuse argivestlustes leidis seda siiski tunduvalt vähem kui pronoomenit *mingi*. Kokku oli 38 *üks*NP-d, mõned neist kasutustest aga olid selgelt arvulises tähenduses. Arvulist ja mittearvulist tähendust on muidugi raske kui mitte võimatu eristada, minu materjalis vastasid kõik *üks*NP-d tõesti ühele entiteedile. Kindlasti on kõnekeeles võimalik pronoomenit *üks* kasutada ka mitmusliku entiteedi atribuudina (*Ühed poisid tulid vastu*), kuid korpuses neid ei esinenud. Arvuliseks kasutuseks pidasin selliseid *üks*NP-sid, kus *üks* oli rõhutatud või mille referent oli ‘üks mitmest’ (*veel üks mees* näites 19). Suhteliselt kergem on eristada selliseid *üks*NP-sid, milles *üks* on kasutatud tähenduses ‘sama’ (näit *ühel ajal* näites 19, samuti *elan temaga ühes trepikojas* jms).

Indefiniitseks ei pidanud ma ka ühte *üks*NP-d, milles *üks* oli kasutatud tähenduses ‘umbes’ (vt näide 20 *üks kaheksakümmend kilo liha*). See on huvitav paralleel pronoomeniga *mingi*, mis näitab, et *üks* ja *mingi* omavad ka muid sarnaseid funktsioone peale indefiniitsuse osutamise. Indefiniitsuse (kokku oli selliseid *üks*NP-sid 22) tüüpilised juhtumid on

eelkõige referentsiaalsed spetsiifilised üksNP-d, milles pronoomeni üks funktsioon on näidata, et viidatav on vestluses uus referent. Tavaliselt alustab selline üksNP narratiivi, nagu näidetes (21) ja (22), milles järgneb lugu sellest inimesest, kes on üksNP-ga narratiivi aegruumi sisse toodud.

üksNP võib olla ka mittereferentsiaalne. Näites (24) kirjeldatakse Almat kui vanaaegset kooliõpetajat, üks vanaaegne kooliõpetaja näitab siin kuulumist tüüpiliste vanade kooliõpetajate hulka.

Kirjakeeles on üksNP-de sagedus umbes sama kui mingiNP-del: samas miljonisõnalis keskkorpus leidus kokku 705 üksNP-d ja numeraali üks sisaldavat arvsõna fraasi (ja 45 prosubstantiivina kasutatud sõna üks). Kui nende hulgast aga välja arvata 72 ki-liitega ükskiNP-d, mille tähendus on teistest hoopis erinev ja mida siinkohal ei vaadelda, ja 9 prepositiooni ühes 'koos' jääb järele 624 fraasi. Kolm suuremat rühma nende hulgas sisaldasid sõna üks numeraalses tähenduses (kokku 462 fraasi): 1) puhtalt arvuna 189 (näiteks 26), 2) tähenduses 'üks paljudest' 190 (vt 28–30) ja 3) lausetes, kus vastandavalt viidati kahele või enamale entiteedile malli üks X ja teine X (ja kolmas X) abil 83 (näiteks 27).

Indefiniitse tõlgendusega üksNP-sid oli kirjakeele materjalis 116. Need on tüüpiliselt uued, lugejale tundmatud isikud (vt 32) või muud entiteedid (vt 33).

Tabel 1 esitab arvandmed mingiNP-de ja üksNP-de võrreldavast osast, st valdkondadest, kus nendel fraasidel on lähedane tähendus.

Tabel 1.

	mingi		üks	
	argivestlused	kirjakeel	argivestlused	kirjakeel
mitte-referentsiaalne	22	79	4	8
Referentsiaalne Spetsiifiline	43	166	22	116
Umbkaudne arvuga	8	–	1	–
Kokku	73	245	27	124

üks ja mingi on sarnased oma indefiniitses funktsioonis, kusjuures paljudel juhtudel võiks neid ilmselt vabalt üksteisega asendada. Kuna mingi tähendus on tugevamalt umbmäärastav, pole siiski selgeid piire indefiniitse ja umbmäärase mingi vahel. Selgemalt on indefiniitne funktsioon näha kontekstides, kus on teada, et mingiNP-ga viidatakse täiesti spetsii-

filisele ja iseenesest selgele entiteedile, näiteks isikule (*mingid tüdrukud* näites 14).

On veel üks nüanss, milles *mingi* ja *üks* erinevad ka indefiniitsena. *mingi*NP-ga viidates jätab kõneleja mulje, et viidatav entiteet on tundmatu ka talle endale, mitte ainult kuulajale. *mingi* on seega totaalse tundmatuse determinaatore (vt *mingi lapsega* näites 4 real 15). *üks*NP-ga viidates sellist lisatähendust ei edastata, indefiniitsus seisneb tõesti ainult selles, et entiteet on antud diskursuses kuulajale identifitseerimatu.

Kuna pronoomeniga *mingi* kaasneb tugev umbmäärastav tähendus ja sellega koos kõnekeeles ka teatud hinnangulisus, lubab *mingi* kasutus näidata kõnelejal oma suhtumist viidatavasse (*mingi Heli* ja *mingi Tiit* näidetes 6 ja 7). Selle emotsionaalsuse kaudu on pronoomenist *mingi* tõenäoliselt kujunenud kõnekeele partikkel. Pronoomenil *üks* hinnangulisi lisatähendusi ei ole.

Sõnade *üks* ja *mingi* tähendust võib vaadelda üldistavalt kahe tähendusspektrina, mille keskmes on numeraalne tähendus sõnal *üks* ja umbmäärane tähendus sõnal *mingi*. Mõlemast keskmest hargnevad hägusate üleminekutega väljad.

PREP	←	PROADJ	←	NUM	→	PROADJ	→	INDEF	DET	→	ADV
<i>ühes</i>		<i>üks</i>		<i>üks</i>		<i>üks</i>		<i>üks</i>			<i>üks</i>
'koos'		'sama'		'1'		'üks		'kaasvestlejale			'umbes'
						paljudest'		tuvastamatu'			

PROADJ	←	PROADJ	→	INDEF	DET	→	
<i>mingi</i>		<i>mingi</i>		<i>mingi</i>			
'eituse		'umbmäärane,		'kaasvestlejale		PARTIKKEL	
tugevdamine'		segane'		tuvastamatu'		<i>mingi</i>	
			→	PROADV	→		
				<i>mingi</i>			
				'umbes'			

Käesolevas artiklis kasutatud materjali põhjal ei saaks kumbagi siiski pidada grammatikaliseerunud indefiniitseks artikliks eelkõige väikese kasutussageduse tõttu (samas kõnekeele korpuse osast leidsin 375 definiitse artiklina kasutatud pronoomenit *see*).

# **BASIC TERMS AND BASIC VOCABULARY**

**Urmaz Sutrop**

*University of Konstanz*

## **1. Introduction**

There are no clear definitions of what a basic word, term or vocabulary are. They are ambiguous concepts in linguistics. The most often used attempt to define basic terms is to define them based on their frequency in the language. One could then infer that circa one hundred grammatical and lexical entities make up to 50% of the lexicon of some simple texts. If one would take this view and look in a dictionary where all the entries are ordered by frequency, he or she would find that all basic words would fall in the first two or three pages. All such frequent terms would then be considered to form what is known as basic vocabulary.

One could also attempt to define basic terms requiring that they be a native and morphologically simple word. From a psycholinguistic point of view, a basic term must be psychologically salient. One could also claim that there are basic words or terms for every semantic sphere – for aristocracy, for seafaring, for body-parts, for kinship relations, and for colours. Of course there are authors who say that the terms for aristocracy and seafaring are not basic, but rather cultural words. There is no good argument against the fact that cultural words are both basic and non-basic words.

Evidently all these views are correct but one must consider that there are basic words in the language as a whole and in the special semantic fields as well. In some cases the basic status in some semantic fields may coincide with the basic status in the language as a whole but this is not always the rule.

One assumes that there is a basic vocabulary in every language. In this paper the concepts of basic term and of basic vocabulary will be attempted to be clarified in the typological tradition in which Berlin and Kay defined the basic colour term

(1969) and Viberg analysed the basic vocabulary (verbs) (1983, 1998). First the concept of the basic colour term will be revised, the list task in defining a basic colour term will be criticised, some linguistic and cognitive criteria of basicness will be presented, a cognitive salience index will be offered, basic objects at a basic level will be investigated, and then an attempt to offer a more general definition of the basic term that will cover domains other than colours will be made. At the end of this paper the concept of basic vocabulary will be clarified and the basic Estonian vocabularies suggested by several authors will be considered. Last the structure of the basic vocabulary will be attempted to be found and the core of the Estonian (basic) vocabulary will be sketched.

## 2. What is a basic colour term?

Although there are important critical papers discussing the original concept and characteristics of the basic colour term, there is no real progress in defining the basic colour term (Crawford 1984, Moss 1989, Ratner 1989).

Some authors claim that “the exact meaning of basic colour terms has never been spelled out” (Saunders & Brakel 1997: 168). For this reason the original concept of the basic colour term (Berlin & Kay 1969: 5–7) will be reviewed. The original characteristics will be rearranged, be given them different weights, and be adapted, for example, to the concepts of the basic temperature term, basic smell term, and basic taste term. In the work of Berlin and Kay, the original characteristics were divided into two groups. The first four characteristics of the Berlin and Kay’s definition of basicness were most important. Only if the term was still questionable after applying the four primary characteristics, would the four subsidiary criteria be applied.

The original four primary characteristics for defining the basic colour term are:

- (i) It [the term] is *monolexemic*; that is, its meaning is not predictable from the meaning of its parts.
- (ii) Its signification is not included in that of any other colour term.
- (iii) Its application must not be restricted to a narrow class of objects.

- (iv) It must be psychologically salient for informants. Indices of psychological salience include, among others:
- (1) a tendency to occur at the beginning of elicited lists of colour terms,
  - (2) stability of reference across informants and across occasions of use, and
  - (3) occurrence in the idiolects of all informants.

And the four subsidiary criteria were:

- (v) The doubtful form should have the same distributional potential as the previously established basic terms.
- (vi) Colour terms that are also the name of an object characteristically having that colour are suspect.
- (vii) Recent foreign loan words may be suspect.
- (viii) In cases where lexemic status is difficult to access, morphological complexity is given some weight as a secondary term (Berlin & Kay 1969: 6–7).

A. E. Moss (1989) argues that Berlin and Kay's concept of basicness is an indivisibly psycholinguistic concept, but it may be useful to distinguish between psychological – or perceptual – basicness and linguistic basicness. According to Moss, neither physiological nor psychological basicness is likely to differentiate primary and secondary colours on one hand, from remaining colours on the other. Moss suggests that in this case the concept of linguistic basicness could be more useful. He points out that on purely linguistic grounds one can distinguish between more basic and less basic terms, rather than between basic and non-basic terms. Finally Moss argues that one level of basicness cannot be reduced to another.

Bearing in mind Moss's distinction between linguistic and psychological characteristics of the definition of basic colour term, one can better understand Crawford's revised criteria, for basicness. One notices that he has rejected all linguistic criteria, mentioned in Berlin and Kay's definition. His own revised definition has left out all such linguistic criteria. He states: "A basic colour term occurs in the idiolects of all informants. It has stability of reference across informants and across occasions of use. Its signification is not included in that of any other colour

term. Its application is not restricted to a narrow class of objects” (Crawford 1982: 342).

### 3. The critique of the list task

Discussing Berlin and Kay’s characteristic (iv), T. D. Crawford also rejects the condition (1) – a tendency to occur at the beginning of elicited lists of colour terms, arguing that its application requires a large number of informants, whereas Berlin and Kay often only had access to one informant for a language. This disadvantage has now been overcome with the help of Davies and Corbett’s field method, which they used to interview more than 50 informants (Davies & Corbett 1994, 1995). This field method is also applied in (Sutrop 1995a, 1998a, 1998b, 1998c) where 80 people were interviewed in investigating Estonian colour, smell, taste, and temperature terms. The above mentioned condition (1) – to occur at the beginning of elicited lists – is a fundamental one.

The second condition (2) – stability of reference across informants and across occasions of use – is very important for part two – the tile naming task of Davies and Corbett’s field method. Crawford accepts this condition but adds that this condition also requires the use of at least a small group of informants (1982: 340). This requirement is satisfied by Davies and Corbett’s field method using a representative group of informants, rather than a single informant. Crawford also rejects the third condition (3) – occurrence in the idiolects of all informants – if the number of informants is small. He includes this condition into his own definition of basic colour term. Condition (3) implies that every subject must name a term in order to be basic in the list task; the corresponding measure in the list task is the frequency of the term in question.

Of the three conditions of psychological salience, conditions (1) and (3) are applicable to the list task, conditions (2) and (3) to the tile naming task. Condition (2), which is very important in colour surveys, is not adequately applicable to other semantic fields for there are no standard stimuli available which represent, e.g., the smells, tastes, and temperatures as precisely as colour tiles, defined in colour space, represent the colours. This means



that the modified list tasks must be used in such cases where there are no standard stimuli available.

The idea of the list task goes back to the works of Eric H. Lenneberg who writes that “we might start with a collection of colour words in use by English speakers (for instance, by asking a sample of speakers to write down all the words for colours they can think of) and then to try to assign each word a region in the colour space” (1967: 339–340).

Lenneberg calls this procedure *Approach A*. He also theoretically grounded *Approaches B* and *C*, but *Approach A* already contains the idea of the list task. Lenneberg’s approaches are, without being directly quoted, incorporated into Berlin and Kay’s methodology for establishing basic colour terms. It is also a part of the Davies and Corbett’s field method.

Davies and Corbett (1994, 1995, Corbett & Davies 1997) set two parameters – the term frequency, and the mean position the term is named – from the list task data. The term frequency represents Berlin and Kay’s condition (3) – the basic term is to be found in the idiolects of all informants; and the mean position satisfies the condition (1) – the basic term occurs at the beginning of elicited lists of colour terms – only when the mean position is small.

#### 4. A cognitive salience index

J. Jerome Smith has independently developed a list task – the free-listing exercise – integrating the two list task parameters – frequency and mean position – into one integral parameter – free-list salience index. He and his colleagues explain: “The free-list salience index ranges from 0 to 1, with higher figures denoting higher salience. For a given term in a given list in which it occurs (ranged according to its order of appearance in that list), an index score is calculated as follows:

$$\text{term index score} = (\text{list-length term rank}) \times (1 / (\text{list length} - 1))$$

The term’s index score for each list in which it occurs is thus calculated. A term’s mean score across all lists (even those in which the term does not appear) is its free-list salience index” (Smith *et al* 1995: 206).

The salience index introduced in this paper is different from the free-list salience index. If the free-list salience index is calculated over individual lists, then salience index for a term is calculated from its frequency and mean position. This index can be used for discriminating basic terms from non-basic terms (see Sutrop 1998a, 1998b, 1998c).

Salience index introduced in this paper is based on the important characteristics (iv) of the basic colour term according to which the basic term must be psychologically salient for informants. This index combines the tendency of a basic term to occur at the beginning of the elicited lists (mean position) and its occurrence in the idiolects of all informants (term frequency) into one integral parameter. Salience (S) is a product of two factors:

$$S = (F / N) \times [(L - mP) / (L - 1)]$$

The first factor  $F / N$  considers the frequency (F) a term is named in the list task, N is the number of informants. If all informants have named a term, then this factor is for that term one. The second factor  $(L - mP) / (L - 1)$  considers the weight of the mean position (mP) the term is named; L is a parameter which takes into account the lengths of the lists (L is the mean length of the individual lists). If the mean position is one, this factor is also one.

The ideal basic term that is psychologically most salient, has the value one for both factors, so the product S would also be one. If the parameter mean position (mP) for some term is equal to the mean length of the list (L), i.e.  $mP = L$ , then the value for our salience index  $S = 0$ ; and if the mean position of a term is greater than the mean length of the individual lists, then the cognitive salience index has negative figure ( $S < 0$ ).

If the value of the parameter L is changed, the salience index's selectivity can be tuned and the second component – the mean position – can be given greater or smaller weight in comparison with the first frequency factor.

The frequency, mean position and integral salience are all good criteria for discriminating basic terms from non-basic terms. Sometimes the discrimination must be made between more and less basic terms. In such cases the linguistic criteria can be well applied. Unfortunately the development of basic colour term

research is moving in the direction of accepting only psychological salience, rather than linguistic criteria.

### 5. What is a basic level and a basic object?

The linguistic criteria – morphological complexity (both Berlin and Kay's original characteristics (i) and (viii)), as well as distributional potential may be useful as a frame for comparative analysis of the basic or non-basic terms. This applies also to the etymological criteria (vi) and (vii). For establishing basic status of terms other than colours, criterion (iii) – its application must not be restricted to a narrow class of objects – has great importance. But this criterion must be expanded – its application must not only be restricted to a narrow class of objects, but also to some of all possible semantic fields. For example, one may discriminate between three semantic fields for temperatures – animate, inanimate, and weather. To be basic, a temperature term must be applicable to all these three semantic fields. The terms, e.g. restricted only for speaking about weather, are not basic. Also recent loan-words (vii) may be suspected, if there are some doubts about their basicness according to salience or some other criteria.

Keeping in mind the concept of basic level and basic objects, one may refine the definition of the basic term. This approach is developed in the work of Eleanor Rosch and her colleagues. She shows that basic objects should be at the level of abstraction at which it is generally most useful to refer to objects (Rosch *et al* 1976: 406–407). George Lakoff recites the basic-level effects from the paper of Rosch and her colleagues:

“Rosch and others (1976) have found that the basic level is:

- The highest level at which category members have similarly perceived overall shapes.
- The highest level at which a single mental image can reflect the entire category
- The highest level at which a person uses similar motor actions for interacting with category members.
- The level at which subjects are fastest at identifying category members.
- The level with the most commonly used labels for category members.

- The first level named and understood by children.
- The first level to enter the lexicon of a language.
- The level with the shortest primary lexemes.
- The level at which terms are used in neutral contexts.
- The level at which most of our knowledge is organized” (Lakoff 1987: 46).

John R. Taylor also discusses basic level terms (1989). He points out that basic level terms can often be distinguished from non-basic terms on purely linguistic, language internal grounds. For basic level terms it is characteristic that (Taylor 1989: 49):

- 1) there is a high frequency of occurrence,
- 2) they are generally short, and
- 3) they are structurally simple, i.e. monomorphemic.

Here the use of the concept ‘basic level term’ is confused with ‘basic term’ (especially in respect to defining a basic level term as opposed to a non-basic term). With high probability a basic level term is a basic term at the basic level but there may be some basic terms also on superordinate or subordinate levels. If so, the non-basic level basic terms form a minority group.

The discussion about similarities between basic level categories and Berlin and Kay’s description of basic colour terms is summed up by David R. Andrews (1994: 12). He states:

- 1) they are the most salient both cognitively and linguistically;
- 2) they are the most neutral in discourse; and
- 3) they are comprised of the shortest and structurally simplest words.

This discussion is very intriguing because it brings the linguistic criteria back into discussion of basicness. The thesis of Taylor, that basic level terms can often be distinguished from non-basic terms on purely linguistic, language internal grounds, contradicts *prima facie* with the opinions of Moss that, based on purely linguistic grounds, one can distinguish between more basic and less basic terms. If one discriminates basic terms from non-basic terms, then he can hierarchically order the basic terms according to their more-or-less-basicness in the sense of Moss. Such hierarchical ordering of basic terms especially makes sense if one attributes that to the term’s numerical values, e.g. salience index.

The higher the salience index, the more basic the term is, and the lower the salience index is, the less basic the term is. On purely language internal grounds one can distinguish between basic and non-basic terms but is not able to decide whether there is an internal hierarchy in the basic term groups or not.

## 6. What is a basic term?

Recently Frans Plank revised the original Berlin and Kay's criteria for basicness and called these criteria for defining basic vs. non-basic term, for example number terms, a general lexical principle (1999).

Plank began his analyses of the basic term with a "precondition: being a term of the language concerned

- (a) being in general rather than only in a specialist use in the language community
- (b) those who know and use it agree rather than disagree on its meaning
- (c) being salient rather than marginal (e.g., comes to mind quickly, without much reflection, when asked to list colour terms)
- (d) literally denoting a colour, and if also used for something else, these meanings must be secondary and metaphorical /.../
- (e) being unrestricted in its applicability to objects of a given colour rather than being restricted to certain objects /.../
- (f) not being included in the meaning of another BCT /.../
- (g) being morphologically simple rather than complex (derivative or compound) /.../ (g') being non-compositional, i.e. either morphologically simple or, if complex, with the meaning and perhaps form of the whole not resulting regularly from the meaning and form of the parts (it is less strict a requirement than (g))
- (h) being of the word class prototypically expressing property concepts in attributive and predicative function rather than being formally specially marked for these functions
- (i) having the full grammatical (e.g., inflectional and derivational) potential of prototypical members of the class of property concept words" (Plank 1999).

Although this definition retains references to colour, it is a great step forward towards the general definition of basicness. As a new characteristic it contains the demand (h) that a word under consideration must belong to the word class prototypically expressing property concepts in attributive and predicative function rather than being formally and specially marked for these functions. In addition the last paragraph (i) sounds linguistically clearer than the original Berlin and Kay's formulation (v) that a word under consideration must have the full grammatical (e.g., inflectional and derivational) potential of prototypical basic terms of the concrete semantic field (e.g., colours, temperatures, numbers, etc.).

The discussion of basic terms brings into consideration five linguistic criteria for basicness. Firstly, basic terms are neutral in discourse, i.e. they are semantically unmarked. Secondly, they are short. This correlates with the fact that native words are normally shorter than loan-words, and short words are usually structurally simpler than longer words. The third point, that basic words are structurally (morphologically) simple, is also included in the second point – shortness. Fourthly, a basic term must belong to the prototypical word class of that domain and last, a basic term must have the same grammatical (e.g., inflectional and derivational) potential of prototypical basic terms of that concrete semantic field.

Now a new definition of the basic term can be given:

**A basic term is a psychologically salient, in most cases morphologically simple and native word, which belongs to the same word class and has the same grammatical potential as the prototypical member(s) of its semantic field. It is a term which generally denotes an object, a quality, or a phenomenon at basic level, and which is applicable in all relevant domains.**

Now this general definition of the basic term can be applied to the concrete semantic fields. For example, the definition of the basic temperature term is as follows:

*Basic temperature term is a psychologically salient, in most cases morphologically simple and native word, which belongs to the same word class and has the same grammatical potential as the prototypical temperature term(s). That term generally denotes*

*the quality or a phenomenon of temperature at basic level, and is applicable in animate, inanimate, and weather domains.*

The definitions for the other domains are quite similar.

*Basic colour term is a psychologically salient, in most cases morphologically simple and native word, which belongs to the same word class and has the same grammatical potential as the prototypical colour term(s). That term denotes a quality of colour at basic level, and is applicable in all relevant domains.*

*Basic taste term is a psychologically salient, in most cases morphologically simple and native word, which belongs to the same word class and has the same grammatical potential as the prototypical taste term(s). That term generally denotes a quality or a phenomenon of taste at basic level, and is applicable in all relevant domains.*

*Basic smell term is a psychologically salient, in most cases morphologically simple and native word, which belongs to the same word class and has the same grammatical potential as the prototypical smell term(s). That term generally denotes a quality or a phenomenon of smell at basic level, and is applicable in all relevant domains.*

From the definition of the basic term it follows that the basic term must not be derived. Only non-derived words are morphologically simple. One must consider this property in the synchronous and diachronous analyses, but in the diachronous analyses only in some arbitrary depth of time. For example, the English colour term *green* (German *grün*) should be diachronously parsed as *gree-n*. The term is derived with *ni*-suffix from the Germanic verbal base *\*grō-a-* 'grow'; *grass* is also derived from the same base. One can thus define the colour term *green* shortly: 'green is the colour of the objects which grow, particularly the colour of grass'. If one recounts Berlin and Kay's characteristics, then it can be seen that the term *green* is not monolexemic (i). The term *green* is morphologically complex (viii) while its meaning is predictable from its parts – the base *grow-* plus the *ni*-suffix. The term *green* is derived from the same root like the name of the object – *grass* – characteristically having that colour (vi). On purely linguistic grounds one must suspect basicness for this term. Green and its cognates in West-Germanic languages are basic colour terms while they are psychologically

salient. On the other hand, synchronously they are opaque (e.g. in German the verbal base *grow-* itself as a verb 'to grow' is absent) and the adjectives *green* and *grün* seem morphologically simple. For such reasons one must restrict his diachronous parsing to limited period, if he will establish basic terms in modern languages.

## 7. What is a basic vocabulary?

As stated previously the concept of basic vocabulary is extremely ambiguous in linguistics. For many linguists basic vocabulary associates with the hundred words proposed by Morris Swadesh which are said to be most stable in any language. It is thought that body parts, kinship terms, personal pronouns, some basic actions, plant and animal names are most conservative. This list actually contains English pronouns (e.g., I, you, this that), question particles (who, what), numerals (one, two), adjectives, (big, long, hot, cold), nouns (man, person, fish, louse, foot, hand), and verbs (eat, drink, kill) (Swadesh [1960] 1971: 283, cf. Raun 1956). The word class of these concepts may be quite different in different languages. This list also includes some colour terms –red, green, yellow, white, and black (nos. 87–91). But in Uralic languages these terms are not at all stable. (Sutrop 1996, 1998c).

Sometimes it is thought that the basic vocabulary consists of the elementary expressions and words that people use, for example, in shopping or hotel reception. In this paper such a primitive approach is ignored.

Another widespread sphere of use of the concept of the basic vocabulary is applied linguistics. The so-called basic dictionaries (Estonian *baassõnastik*) usually present most frequent words with some thematic additions. This approach is based on the frequency lists and dictionaries (see, e.g., Kühn 1979, Krohn 1992). Another approach for determining the basic vocabulary has been innovated in Klett Publishing House in Germany (Ernst Klett Verlag, Stuttgart). This approach is also based on the former frequency lists. In Klett series *Basic and Continuation Vocabulary* (Grund- und Aufbauwortschatz) the words' usability and dispersion are added to the words frequency parameter. The written and colloquial speech are considered equally (see, e.g., Amorim-



Braun 1987, Heupel 1991, Giovanelli 1989, Nickolaus 1990, Vogt 1988, Weis 1991).

Let one take basic and structural English vocabulary into further consideration. It is assumed that the vocabulary structure is essentially the same for all industrial nations. In Figure 1 English vocabulary structure is presented after Weis (1991). The whole vocabulary of the English language is divided in three main classes. Basic vocabulary (*Grundwortschatz*) contains ca. 2000 words, continuation vocabulary (*Aufbauwortschatz*) ca 2500 words, the rest contains more than 100 000 words.

With the help of the basic vocabulary it is possible to understand 85% of a normal standard text, next 2500 structural words add only 10% to reader's understanding, and ca. 2500 words from the remaining hundred thousands or even millions of words make up the final 5% of the understanding. It is important that the core of the basic vocabulary – circa 100 words – helps one understand 50% of a whole simple standard text with medium difficulty.

It is characteristic to all such basic vocabularies, based on a word's frequencies, usability, and dispersion, that all the linguistic parameters (morphologically simple word structure, word class, morphological distribution, nativeness) are ignored. There is no discussion about the structure of the basic vocabulary or lexicon. For that reason the structure of the basic vocabulary is attempted to be found. But first the basic vocabularies of the Estonian language will be looked at.

## 8. Basic Estonian vocabularies

There are two primary Estonian basic vocabularies. One was published in Sweden by Raimo Raag ([1986] 1994) and another in Germany by late professor Wolfgang Veenker (1992). Basic vocabularies published in Estonia are all based on Raag's vocabulary (e.g. Pärn & Simm, koost. 1989, 1990). A pictorial vocabulary of Estonian is an exception (Salmin, Ussisoo 1994).

Raag explained that his *Basic Estonian Vocabulary* "contains in all 517 most frequently used Estonian words. The selection of words is based on an on-going statistical investigation of the vocabulary of Estonian which is carried out at the University of

Tartu /.../ to the 500 most frequently used words according to this investigation, any missing cardinal numbers between one and ten, the word for ‘hundred’ and the words for the days of the week have been added” (1994: 5). So the principle of composing this basic word list is primarily statistical – to the 500 most frequent words some thematic extensions (17) are added.

Veenker selected words for his *Minimal Estonian Vocabulary* using slightly different principles. First the original statistical word lists composed at the University of Tartu (the same that were used by Raag) and the Raag’s word list (Raag 1994) were assembled and extended with the Estonian terms from the other sources (e.g., Pärn & Simm 1989) and with the terms primarily from German and Russian corpora. Then to those German or Russian terms the Estonian equivalents were looked up from German-Estonian or from Russian-Estonian dictionaries. The resulting corpora contained ca. 2800 Estonian words.

Second, two experts in Estonian language (Hans-Hermann Bartens and Cornelius Hasselblatt) were both asked to select 1000 most basic words. After that Veenker revised the Estonian word list and chose final minimal word list of 1000 words. Veenker used the following three criteria.

- 1) Word frequency parameter (rank) ,
- 2) Importance of the concepts and equivalents in other languages (Russian and German),
- 3) Estimated word lists by two experts.

In the process of revision the structure of the minimal word list was considered. It was controlled whether the so-called small or structural words (postpositions, most important pronomina, important numerals, certain adverbs, conjunctives, etc.) substantives, adjectives, and verbs were equally chosen. Contrary to the traditional way of composing basic word lists (most frequent words with thematic extensions) here the unrevised word list was minimised subtracting words not belonging to the basic vocabulary.

Basic vocabularies composed in Estonia which stem from the Raun’s vocabulary contain 734 words. To the original list of 517 words 217 words were added. The guiding principle of choosing these additions was thematic. Some words denoting body parts,

time (months), physical parameters, etc. were added through the method of thematic extension (Pärn & Simm 1989: 3, 1990: 3).

In these basic Estonian vocabularies the words were grouped into substantives, adjectives, pronomina, numerals, verbs, adverbs, postpositions, conjunctions, and interjections. After that all the entries were listed alphabetically and supplemented with short grammatical (derivational) information and glosses in Russian or Swedish (Pärn & Simm 1989, 1990).

For the *Estonian Pictorial Dictionary*, the late Uno Ussisoo pre-chose the terms divided into 59 thematic groups, for example, house, inside the house, family, human body, buildings, birds, red-letter days and events, trades and professions, etc. (Salmin & Ussisoo 1994). The choice of terms is not linguistically grounded in this dictionary. For example, it contains such terms as *periskoop* 'periscope' *telefoniputka* 'call box', *mälestussammas* 'monument' and *sangpomm* 'weight with a handle' which are not basic linguistically or cognitively.

## 9. The structure of the basic vocabulary

Bearing in mind the discussion on the properties of a basic term above and the short look at the basic vocabularies in general and in Estonian in particular, the structure of the basic vocabulary is sketched in Figure 2. The structure of the basic vocabulary is sketched in the tradition in which Berlin and Kay defined a basic colour term (1969) and Åke Viberg analysed the lexicon (1983, 1998).

The basic vocabulary (500–2000 words) is divided into basic and non-basic words. Basic words are divided into cognitively and linguistically basic words. Cognitively basic words are most frequent (frequency parameter) and/or psychologically most salient in the list task (salience index). Linguistically basic words are morphologically simple and native words. Every linguistically basic word belongs to the same word class and has the same grammatical potential as the prototypical member(s) of the corresponding semantic field. Basic terms, both cognitive and linguistic generally denote objects, qualities or phenomena at basic level, and are applicable in all relevant domains.

Non-basic words in basic vocabulary are thematic extensions to the semantic fields covered by cognitively and linguistically basic words (e.g., some numerals, colours, body parts, weekdays, months, etc.). Not all thematic extensions are considered as members of the basic vocabulary. It is also possible that some cognitively basic words fail to be included in the basic vocabulary. For example, in the case of arbitrary subtraction from a longer list some cognitively based words may fall from the basic vocabulary (cf. Veenker 1992). Both possibilities are shown by the overlapping of basic and non-basic vocabulary in Figure 2.

The linguistically basic terms that are most salient (i.e., also cognitively basic) constitute the core of the basic words and the language. There is good reason for one to believe that circa one hundred of the most frequent words make up the core of the Estonian language. According to the word frequency lists composed by the Research Group of Computational Linguistics at the University of Tartu the most frequent words in Estonian are *olema* 'to be', *ja* 'and', and *see* 'this, that'

## 10. The core of the Estonian vocabulary

In this paper a preliminary attempt was made to find out the core of the Estonian language. The core of the Estonian language is thematically presented in Table 1. This corresponds nicely to the cores of other languages (cf. Figure 1). The peculiarity of the Estonian language is its very rich inflectional morphology (14 standard cases in singular and plural). This complicates the composing of the frequency lists, especially when trying to find out the correct lemmas. There are many homonymies of the different inflectional forms. For example, the original frequency list contains among the one hundred most frequent terms the noun *siga* 'pig'. With high probability the lemma *siga* is wrongly derived from the inflectional form *seal* 'demonstrative there but also adessive of *siga* 'pig' (pig-on)'

Unfortunately the Tartu corpus covers only registers from the written language (newspapers, fiction, and academic prose) and colloquial speech is totally ignored in the corresponding word frequency list. For example, there are important differences in the register distribution of the most common verbs across semantic

domains in English. The most common English lexical verb *say* is the most frequent verb in fiction and in news, second most common in conversation but its use is quite low in academic prose. The distribution of frequencies of the most common verbs in various registers in English is quite different (Biber *et al* 1999: 373–375).

The Tartu corpus represents the eighties of this century. For that reason it contains many Soviet references. For example, among the one hundred most frequent terms were such words as *nõukogu* ‘council’ and *liit* ‘alliance, coalition, league’ from *Nõukogude Liit* ‘Soviet Union’. The original frequency list contains also some parts of compounds. All such parts of compounds, apparently wrong lemmas and ideological terms (sovietisms) that no one uses since after the collapse of the Soviet Union are left out from our preliminary list of core words in Estonian.

In Table 1 the Estonian core words are grouped principally in the same way as the basic vocabulary in (Pärn & Simm 1989, 1990). In every group (word class) the terms are ordered according to their frequency. In this paper any closer analysis of the Estonian core vocabulary is not attempted to be given. It is evident that the words are quite short, native, and morphologically simple. The main exception being adverbs and ordinal numerals. Many adverbs are grammaticalised inflectional forms which are, for the majority of the speakers of Estonian, opaque. For example an adverb *juures* ‘at, by near’ is actually an inessive of *juur* ‘root’ (root-in). The full linguistic analysis of the Estonian core vocabulary would be covered in a following paper. But first a corrected frequency list of the Estonian words is needed which is based upon a corpus of all important registers – conversation, fiction, news, and academic prose.

## 11. Summary

In this paper the concept of a basic term and a concept of basic vocabulary has been analysed. It began with the analyses of the concept of the basic colour term. The criticism of the concept of the basic term and methods for defining a basic term (list task) was investigated. After that an original cognitive salience index

(S) was proposed which combined two list task parameters – the tendency of a basic term to occur at the beginning of the elicited lists (mean position) and the occurrence in the idiolects of all informants (term frequency) – into one integral parameter. The concept of a basic level and of a basic object was also clarified. After that a new definition of a basic term was offered:

A basic term is a psychologically salient, in most cases morphologically simple and native word, which belongs to the same word class and has the same grammatical potential as the prototypical member(s) of its semantic field. It is a term which generally denotes an object, a quality, or a phenomenon at basic level, and which is applicable in all relevant domains.

What defines a basic vocabulary was also asked. In all languages the whole vocabulary can be divided into three main classes – basic vocabulary, continuation vocabulary and the remaining hundred thousands of words. Basic vocabulary contains 500–2000 words and can be divided into basic and non-basic terms (words). Basic terms can be divided into cognitively and linguistically basic terms. Those linguistically basic terms that are also concurrently cognitively basic (most frequent and salient) form the core of that language. These one hundred words are short, morphologically simple and native. The non-basic words in the basic vocabulary are thematic extensions to the semantic fields covered by the basic words (some absent body parts, numerals, colours, etc.).

The existing basic Estonian vocabularies were also investigated and a preliminary attempt to sketch the core of the Estonian language was made.

### **Acknowledgements**

I am grateful to Prof. Frans Plank for comments and helpful discussion and to Ms. Shannon Stout for reading the manuscript and correcting my English.

## REFERENCES

- Amorim-Braun, Maria Luisa (Bearb.) 1987:** Grund- und Aufbauwortschatz: Portugiesisch. Klett. Stuttgart.
- Andrews, David R. 1994:** The Russian color categories *sinij* and *goluboj*: An experimental analysis of their interpretation in the standard and emigré languages. – *Journal of Slavic Linguistics* 2. 9–28.
- Berlin, Brent, Kay, Paul 1969:** Basic Color Terms: Their Universality and Evolution. University of California Press. Berkeley. Reprinted in 1991 with an additional bibliography by Luisa Maffi.
- Biber, Douglas, Johansson, Stig, Leech, Geoffrey, Conrad, Susan, Finegan, Edward 1999:** Longman Grammar of the Spoken and Written English. Longman. London.
- Corbett, Greville G., Davies, Ian R. L. 1997:** Establishing basic color terms: Measures and techniques. – C. L. Hardin, Luisa Maffi (eds.), *Color Categories in Thought and Language*. Cambridge University Press. 197–223.
- Crawford, T.D. 1982:** Defining ‘basic color term’ – *Anthropological Linguistics* 24. 338–343.
- Davies, Ian R. L., Corbett, Greville G. 1994a:** The basic colour terms of Russian. – *Linguistics* 32. 65–89.
- Davies, Ian R. L., Corbett, Greville G. 1994b:** A statistical approach to determining basic color terms: An account of Xhosa. – *Journal of Linguistic Anthropology* 4. 175–93.
- Davies, Ian R. L., Corbett, Greville G. 1995:** A practical field method for identifying probable basic colour terms. – *Languages of the World* 9. 25–36.
- Giovanelli, Paolo (Bearb.) 1989:** Grund- und Aufbauwortschatz: Italienisch. Klett. Stuttgart.
- Heupel, Carl (Bearb.) 1991:** Grund- und Aufbauwortschatz: Spanisch. Klett. Stuttgart.
- Krohn, Dieter 1992:** Grundwortschätze und Auswahlkriterien. Metalexigraphische und fremdsprachendidaktische Studien zur Struktur und Funktion deutscher Grundwortschätze. (Göteborg germanistische Forschungen, 34.) Acta Universitatis Gothoburgensis. Göteborg.
- Kühn, Peter 1979:** Der Grundwortschatz. Bestimmung und Systematisierung. (Germanische Linguistik, 17.) Niemeyer. Tübingen.

- Lakoff, George 1987:** *Women, Fire, and Dangerous Things. What Categories Reveals about the Mind.* University of Chicago Press. Chicago, London.
- Lenneberg, Eric H. 1967:** *Biological Foundations of Language.* With appendices by Noam Chomsky and Otto Marx. John Wiley & Sons, Inc. New York.
- Moss, Anthony E. 1989:** Basic color terms: Problems and hypotheses. – *Lingua* 78. 313–320.
- Nickolaus, Günter (Bearb.) 1990:** *Grund- und Aufbauwortschatz: Französisch.* Klett. Stuttgart.
- Pärn, H., L. Simm (koost.) 1989:** *Базовый словарь эстонского языка. Eesti keele baassõnastik.* Tallinn.
- Pärn, H., L. Simm (koost.) 1990:** *Eesti keele baasõnastik. Estniskt basordförråd.* Tallinn.
- Plank, Frans 1999:** *The implications of colour.* Manuscript.
- Raag, Raimo 1994:** *Estniskt basordförråd. Basic Estonian vocabulary.* 6<sup>th</sup> printing. Uppsala.
- Ratner, Carl 1989:** A sociohistorical critique of naturalist theories of color perception. – *The Journal of Mind and Behaviour* 10. 361–372.
- Raun, Alo 1956:** Über die sogenannte lexicostatische Methode oder Glottochronologie und ihre Anwendung auf das Finnisch-Ugrische und Türkische. – *Ural-Altäische Jahrbücher* 28. 151–154.
- Rosch, Eleanor, Mervis, Carolyn B., Gray, Wayne D., Johnson, David M., Boyes-Braem, Penny 1976:** Basic objects in natural categories. – *Cognitive Psychology* 8. 382–439.
- Salmin, Aleksander, Ussisoo, Uno 1994:** *Eesti keele piltsõnastik. Koolibri.* Tallinn.
- Saunders, B. A. C., Brakel, J. van 1997:** Are there nontrivial constraints on colour categorization? – *Behavioral and Brain Sciences* 20. 167–228. (This article contains also 31 open peer commentaries, the authors' response, and a bibliography.)
- Smith, J. Jerome, Furbee, Louanna, Maynard, Kelly, Quick, Sarah, Larry Ross 1995:** Salience counts: A domain analysis of English colour terms. – *Journal of Linguistic Anthropology* 5. 203–216.
- Sutrop, Urmās 1995:** *Eesti keele põhivärvinimed.* – *Keel ja Kirjandus*, no. 12. 797–808.
- Sutrop, Urmās 1996:** *Eesti keele värvussõnavara arengu põhijooni.* – *Keel ja Kirjandus*, no. 10. 661–674.
- Sutrop, Urmās 1998a:** Basic temperature terms and subjective temperature scale. – *Lexicology* 4. 60–104.
- Sutrop, Urmās 1998b:** *Eesti keele maitsenimetustest.* – *Keel ja Kirjandus*, no. 7. 446–458.



- Sutrop, Urmas 1998c:** The Vocabulary of Sense Perception in Estonian: Typology, History, and Cognition. PhD dissertation. Manuscript. Konstanz.
- Swadesh, Morris 1971:** What is glottochronology? – Morris Swadesh, The origin and diversification of language, ed. by. Joel Sherzer. Aldine-Atherton. Chicago. 271–284. (Spanish original of this paper Mexico City 1960).
- Taylor, John R. 1989:** Linguistic Categorization. Prototypes in Linguistic Theory. Clarendon Press. Oxford.
- Veenker, Wolfgang 1992:** Minimalwörterverzeichnis der estnischen Sprache. EST BAS LEX MINI WV Die 1000 wichtigsten Wörter. Mare Balticum. Köln.
- Viberg, Åke 1983:** The verbs of perception: A typological study. – Brian Butterworth, Bernard Comrie, Östen Dahl (eds.), Explanations for language universals. (Linguistics, 21 [263].) Mouton. Berlin, Amsterdam, New York. 123–162.
- Viberg, Åke 1998:** Polysemy and differentiation in the lexicon: Verbs of physical contact in Swedish. – Jens Allwood, Peter Gärdenfors (eds.), Cognitive semantics: Meaning and cognition. (Pragmatics & Beyond. New Series, 55.) Benjamins. Amsterdam & Philadelphia. 87–129.
- Vogt, Helger Oleg (Bearb.) 1988:** Grund- und Aufbauwortschatz: Russisch. 2. Aufl. Klett. Stuttgart.
- Weis, Erich (Bearb.) 1991:** Grund- und Aufbauwortschatz: Englisch. Klett. Stuttgart.

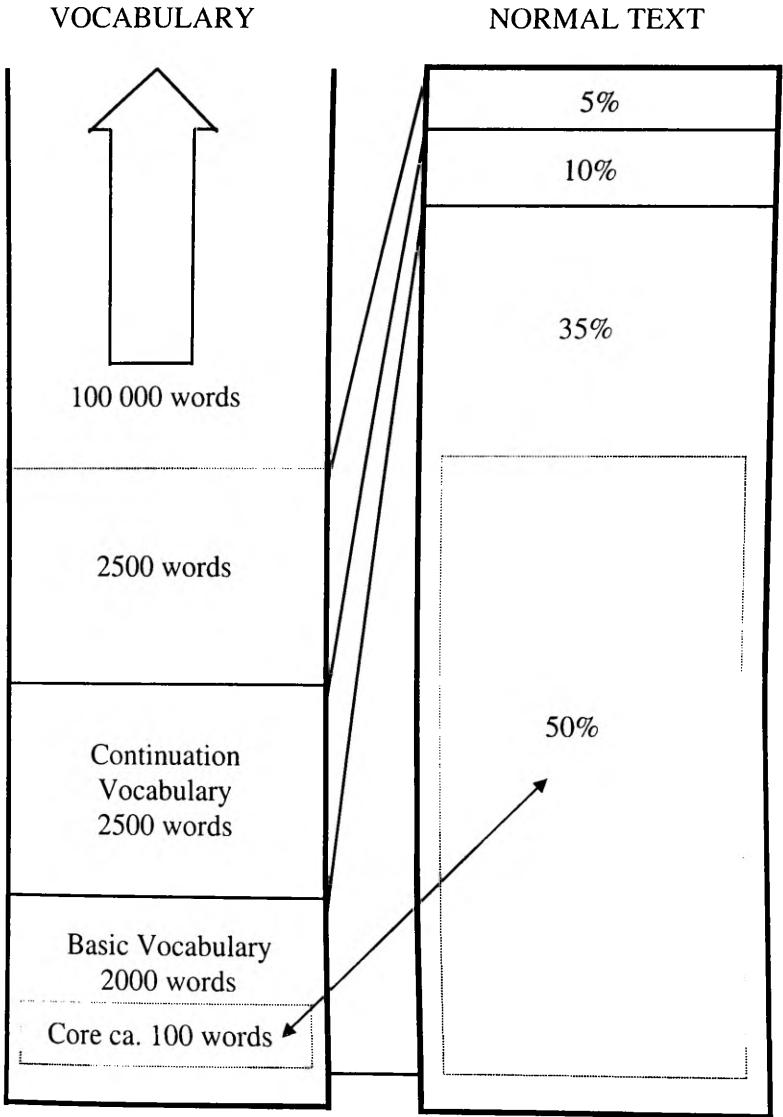


Figure 1. The structure of the (English) vocabulary modified after Weis (1991).

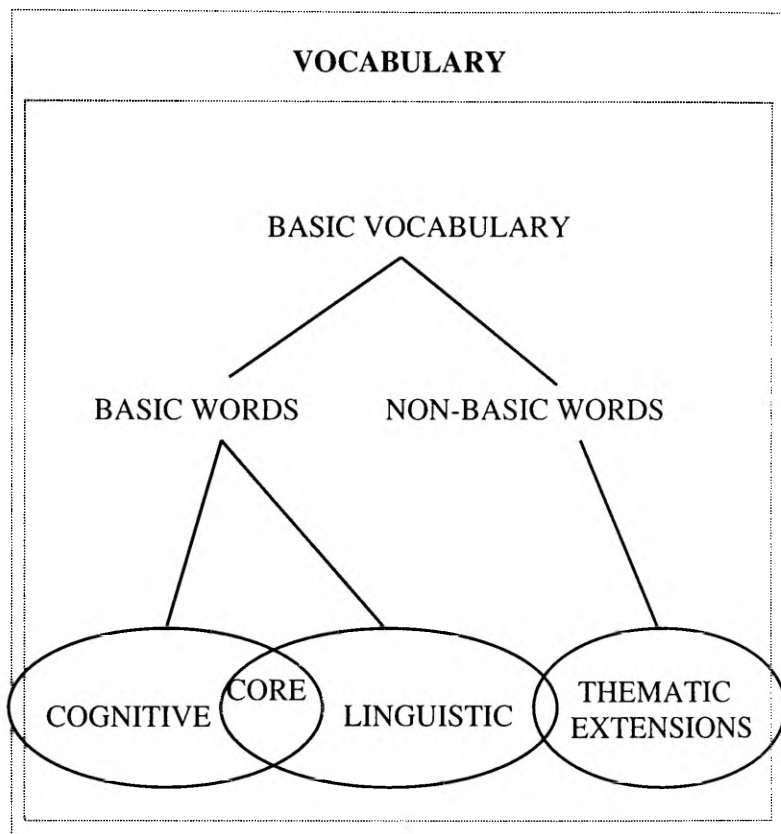


Figure 2. The structure of the basic vocabulary.

**Table 1. The core of the (basic) Estonian vocabulary**

One hundred most frequent words selected from the frequency list of the corpus of the written Estonian of the Research Group of Computational Linguistics at the University of Tartu ([http://www.cl.ut.ee/ee/tulemusi/sag\\_lem\\_1000.kogu](http://www.cl.ut.ee/ee/tulemusi/sag_lem_1000.kogu)) In every word class the terms are ordered according to their frequency.

**1. Pronouns**

<i>see</i>	'this, that'
<i>mis</i>	'what, which, that'
<i>mina</i>	'I'
<i>ta</i>	'he, she, it (short form)'
<i>oma</i>	'own'
<i>kõik</i>	'all, everybody, everything'
<i>kes</i>	'who'
<i>ise</i>	'self'
<i>ma</i>	'I (short form)'
<i>iga</i>	'every'
<i>sina</i>	'you (sg.)'
<i>kogu</i>	'all, whole'
<i>mõni</i>	'some'

**2. Numerals**

<i>teine</i>	'second'
<i>üks</i>	'one'
<i>esimene</i>	'first'
<i>kaks</i>	'two'

**3. Nouns**

<i>aasta</i>	'year'
<i>vesi</i>	'water'
<i>töö</i>	'work'
<i>kuu</i>	'month, moon'
<i>aeg</i>	'time'
<i>pea</i>	'head'
<i>inimene</i>	'human, person'
<i>kord</i>	'order, time'
<i>eesti</i>	'Estonia, also adj. Estonian'
<i>koht</i>	'place'
<i>maa</i>	'earth, land, country'
<i>osa</i>	'part'
<i>tuli</i>	'fire'
<i>tee</i>	'way, road'
<i>mees</i>	'man'
<i>päev</i>	'day'

<i>vahe</i>	'space, distance, difference'
<i>käsi</i>	'hand'
<i>elu</i>	'life'
<i>näide</i>	'example'

**4. Adjectives**

<i>suur</i>	'big, great, grand, large'
<i>uus</i>	'new'
<i>noor</i>	'young'
<i>hea</i>	'good'

**5. Verbs**

<i>olema</i>	'be'
<i>võima</i>	'be able, capable'
<i>saama</i>	'get'
<i>omama</i>	'own, possess, have'
<i>pidama</i>	'keep, hold'
<i>tulema</i>	'come, arrive'
<i>tegema</i>	'do, make'
<i>minema</i>	'go'
<i>võtma</i>	'take'
<i>jääma</i>	'remain, stay'
<i>andma</i>	'give'
<i>ütleva</i>	'say, tell'
<i>teadma</i>	'know'
<i>hakkama</i>	'begin, start'
<i>nägema</i>	'see'
<i>tahtma</i>	'want, wish'
<i>koguma</i>	'collect, gather'
<i>tooma</i>	'bring'
<i>tundma</i>	'feel, sense'
<i>panema</i>	'put, lay, set'
<i>kohtama</i>	'meet'
<i>viima</i>	'carry'
<i>vaatama</i>	'look'

**6. Adverbs**

<i>ka</i>	'too, also'
<i>siis</i>	'then'

<i>nii</i>	'so', thus'
<i>veel</i>	'still, yet'
<i>pool</i>	'at, to, towards, also numeral half'
<i>juba</i>	'already, yet'
<i>kiüll</i>	'enough'
<i>välja</i>	'out'
<i>üle</i>	'over, above'
<i>kus</i>	'where'
<i>palju</i>	'many'
<i>ära</i>	'away, off, also neg. do not'
<i>ainult</i>	'only'
<i>nüüd</i>	'now'
<i>vastu</i>	'against, to, towards'
<i>eest</i>	'from, before'
<i>pärast</i>	'after, later'
<i>enam</i>	'more'
<i>siin</i>	'here'
<i>juures</i>	'at, by, near'
<i>juurest</i>	'from'

### 7. Conjunctions

<i>ja</i>	'and'
<i>et</i>	'that'
<i>kui</i>	'if, when'
<i>ning</i>	'and'
<i>aga</i>	'but'
<i>või</i>	'or'
<i>nagu</i>	'as'
<i>kuid</i>	'but, yet'
<i>kas</i>	'whether, also question particle'
<i>vaid</i>	'but, merely, only'
<i>sest</i>	'for, because'

### 8. Negations

<i>ei</i>	'no, not'
<i>ära</i>	'do not'
<i>mitte</i>	'not, no'
<i>ega</i>	'nor, neither ... nor,
<i>not'</i>	

## PÕHISÕNAD JA PÕHISÕNAVARA

### Urmas Sutrop

Keeleteaduses kasutatakse nii põhisõna kui ka põhisõnavara mõistet ähmaselt, ilma defineerimata. Tihti arvatakse, et põhisõnad on keele kõige sagedasemad sõnad, või siis hoopis, et igal semantilisel väljal on oma põhisõnad, näiteks kehaosade, sugulussuhete, merenduse, jahinduse, hotellinduse või värvuste nimetused. Kõige primitiivsemal juhul peetakse põhisõnavaraks turismiteatmikes leiduvaid sõnu ja väljendeid, nagu *ei, jaa, palun kutsuge arst, appi!, kust ma leian restorani?* jne.

Käesolevas artiklis püütakse määratleda põhisõna ja põhisõnavara lähtudes tüpoloogilisest traditsioonist, milles Brent Berlin ja Paul Kay defineerisid põhivärvinimed (1969) ning Åke Viberg analüüsis taju- verbe ja põhisõnavarasse kuuluvaid verbe (1983, 1998).

Põhisõna defineerimist alustatakse Berlini ja Kay põhivärvinime mõiste ja selle kriitika analüüsiga. Näidatakse, et see määratlus sisaldab nii psühholoogilisi (põhivärvinimi on psühholoogiliselt esiletulev) kui ka lingvistilisi elemente (põhivärvinimi on morfoloogiliselt lihtne ning täieliku morfoloogilise (nii tuletus- kui ka sõnamuutelise) potentsiaaliga. Lühidalt: põhivärvinimi on psühholingvistiline mõiste.

Põhivärvinime psühholoogiline esiletulek ilmneb ühelt poolt põhivärvinime üldises sageduses antud keeles ning teiselt poolt selle sõna kasutamise sageduses ja kiires meenumises nimetamiskatses. Pärast nimetamiskatse kriitilist vaatlust esitas käesoleva kirjutise autor sõna kognitiivse esiletuleku indeksi nimetamiskatses.

Nimetamiskatse puhul palutakse katseisikutel nimetada näiteks nii palju värve, kui neile meenub. Katses peab osalema vähemalt 25 informanti (käesoleva töö autoril osales värvide, temperatuuride, lõhnade ja maitsete testides 80 isikut; vastavaid andmeid pole siin töös esitatud). Seejärel arvutatakse iga nimetatud värvinime sagedus ( $S$ ) ja mitmendana seda keskmiselt nimetati ( $kP$  = keskmine positsioon). Kognitiivse esiletuleku indeksis ( $E$ ) ühendatakse need kaks näitajat ( $S$  ja  $kP$ ):

$$E = (S / N) \times [(X - kP) / (X - 1)]$$

Vaatleme sõna kognitiivse esiletuleku indeksi värvide nimetamiskatse näitel. Sõna kognitiivses esiletuleku indeksis on esiletulek ( $E$ ) kahe teguri korrutis. Esimene tegur  $S / N$  väljendab seda, mitu isikut ( $S$  = sagedus) kõigist katseisikutest ( $N$  = katseisikute koguarv) antud värvi nimetas. Ideaalselt esiletulev värvinimi on selline, mida nimetasid kõik katseisikud, seega siis  $S = N$  ja  $S / N = 1$ . Teine tegur  $(X - kP) / (X - 1)$  väljendab seda, kui kiiresti mingi värv katseisikutele meenus (kui üldse).  $X$  on võetud katseisikute poolt keskmiselt nimetatud värvide arvuks. Kui mingi värvinimi meenub kõigile katseisikutele esimesena, siis on selle värvinime keskmine positsioon  $kP = 1$  ja avaldise teine tegur  $(X - 1) / (X - 1) = 1$ .

Ideaaljuhul võrduvad mõlemad tegurid ühega ja seega ka  $E = 1$ . Niisugune on ideaalne kognitiivselt kõige esiletulevam värvinimi. Kui aga mingi värvi nimetamise keskmine positsioon on võrdne keskmiselt nimetatud värvide arvuga ( $kP = X$ ), siis  $X - kP = 0$  ja seega ka  $E = 0$ . Kui aga mingit värvi nimetatakse päris viimaste meelde tulevate värvide seas, st  $kP > X$ , siis on esiletuleku indeks negatiivse väärtusega  $E < 0$ .

Järgnevalt vaadeldi psühholoogiast pärit põhitaseme (basic level) ja põhiobjekti (basic object) mõisteid. Näiteks tasemete *kassitõug* – *kass* – *loom* puhul on põhitasemeks *kass* (näeme ju kassi, aga mitte looma). Sedastati, et põhivärvinimede uurimises on lingvistilised kriteeriumid jäämas tagaplaanile ning põhivärvinimi on muutumas psühholoogiliseks mõisteks. Mõned autorid, näiteks Frans Plank, on püüdnud põhivärvinime mõistesse siiski lingvistilisi kriteeriume tagasi tuua. Nii psühholoogilistest kui lingvistilistest kriteeriumidest lähtuvalt on siinses töös esitatud põhisõna üldine psühholingvistiline definitsioon:

**Põhisõna on psühholoogiliselt esiletulev, enamikul juhtudel morfoloogiliselt lihtne omasõna, mille sõnaliik ja grammatiline**

**potentsiaal on sama, mis selle sõna semantiliselt välja prototüüpsetel esindajatel. Põhisõna tähistab üldiselt põhitaseme objekti, kvaliteeti või nähtust ning on kasutatav kõikides olulistes valdkondades.**

Näiteks kui rakendame seda definitsiooni temperatuuridele, siis saame:

*Temperatuuri põhisõna on psühholoogiliselt esiletulev, enamikul juhtudel morfoloogiliselt lihtne omasõna, mille sõnaliik ja grammatiline potentsiaal on sama, mis temperatuurisõnade prototüüpsetel esindajatel. Põhisõna tähistab üldiselt põhitaseme temperatuuri, kvaliteeti või nähtust ning on kasutatav rääkides inimestest, asjadest, ilmast, veest jne.*

Põhisõnavara määratlemist alustasime Ernst Kletti kirjastuse põhisõnastike analüüsist. Need lähtuvad eeldusest, et industrialiseeritud rahvaste keelte sõnavara struktuur on enam-vähem sarnane. Joonisel 1 on kujutatud inglise keele sõnavara struktuur ning selle seos keskmise keerukusega teksti mõistmisega. Kogu sõnavara jaguneb kolme suurde rühma: põhisõnavara, jätkusõnavara ning kogu ülejäänud sõnavara. Põhisõnavara (umbes 2000 sõna) abil saab mõista 85% keskmise keerukusega tekstist, jätkusõnavara (2500 sõna) abil veel 10% ning kogu ülejäänud sõnavarast 2500 sõna aitavad mõista viimast 5% tekstist.

Põhisõnavara sisaldab omakorda tuumosa ehk tuumsõnavara (umbes 100 sõna), mille abil on kogu tekstist võimalik aru saada 50% ulatuses. Sinna kuuluvad abisõnad ning mõned omadus-, nimi- ja tegussõnad.

Seejärel võtsime vaatluse alla olemasolevad eesti keele põhisõnastikud – Rootsis Raimo Raagi koostatud eesti keele põhisõnastiku ([1986] 1994) ning Saksamaal Wolfgang Veenkeri koostatud eesti keele miinimumsõnastiku (1992). Eestis välja antud baassõnastikud (nt Pärn ja Simm 1989, 1990) on Raagi sõnastiku edasiarendused.

Raag on võtnud Tartu Ülikooli sõnavaraaurijate tööd oma sõnastikku 500 eesti keele kõige sagedasemat sõna ning täiendanud seda nimestikku 17 sõnaga (puuduvad kardinaalarvud ühe ja kümne vahel, sõna 'sada' nädalapäevad jt). Seega on nimekirja laiendamisel lähtunud sellest, et teatud semantiliselt väljad (teemad) oleksid esindatud terviklikult. Nimetame seda meetodit temaatiliseks laienduseks (thematic extension).

Veenker valis sõnu oma miinimumsõnastiku jaoks veidi teistmoodi. Võtnud aluseks Raagi sõnastiku ja sellele Pärna ja Simmi poolt tehtud täiendused, lisas ta neile arvukalt mõisteid vene ja saksa keele korpusest, tõlkides need eesti keelde saksa–eesti ja vene–eesti sõnaraamatute abil. Nii sai ta kokku 2800 eestikeelset sõna. Siis palus ta kahel eksperdil (Cornelius Hasselblatt ning Hans-Hermann Bartens) kummalgi valida välja 1000 põhisõna. Seejärel kaalus Veenker kogu korpuse

(2800 sõna) uuesti läbi kolme näitaja alusel: sõna sagedus, mõiste ja selle ekvivalentide tähtsus nii eesti kui vene ja saksa keeles ning kolmandaks kahe nimetatud eksperdi hinnangud. Lõpptulemuseks sai ta 1000 sõnast koosneva eesti keele miinimumsõnastiku. Siin on metoodiliselt tähtis, et põhisõnavara ei pruugi leida mitte üksnes temaatilise laienduse, st sõnade nimestikku lisamise teel, vaid ka sõnade nimestikust kustutamise abil.

Pärna ja Simmi sõnastikud sisaldavad 734 sõna. Raagi põhisõnastikule (517 sõna) on lisatud temaatiliste laiendustena veel 217 sõna (kehaosade, kuude, füüsiliste parameetrite jms nimetused).

Eelneva alusel on käesolevas töös esitatud uus skeem põhisõnavara struktuuri kohta (vt joonis 2). Põhisõnavara moodustab osa kogu antud keele sõnavarast. Põhisõnavara jaguneb põhisõnadeks ning mitte-põhisõnadeks. Põhisõnad, mis vastavad ülalesitatud põhisõna definitsioonile, jagunevad omakorda kognitiivseteks põhisõnadeks ning lingvistilisteks põhisõnadeks. Kognitiivsed põhisõnad on niisugused sõnad, mis on antud keeles (selle korpuskes) kõige sagedasemad või osutuvad katse põhjal (nimetamiskatse) kõige esiletulevamateks (sõna kognitiivse esiletuleku indeks). Lingvistilised põhisõnad on niisugused, mis on põhisõnad lingvistiliste kriteeriumide alusel (morfoloogiliselt lihtsad, täieliku morfoloogilise potentsiaaliga omasõnad).

Põhisõnavara mittepõhisõnad on kognitiivsete ja lingvistiliste põhisõnade temaatilised laiendused (puuduvad arvsõnad, nädalapäevi, värve, sugulussuhteid, kehaosi jms väljendavad sõnad). Kognitiivsete ja lingvistiliste põhisõnade ühisosa moodustab nii põhisõnavara kui kogu antud keele tuuma, mille abil peaks olema võimalik aru saada umbes poolest keskmise keerukusega tekstist.

Töö lõpus on esitatud eesti keele tuum (tuumsõnavara), mis on välja valitud Tartu Ülikooli eesti kirjakeele kaheksakümnendate aastate korpusel. Eesti keele kõige sagedasemateks sõnadeks on *olema*, *ja* ning *see*. Tuumsõnavara on rühmitatud sõnaliikidesse põhimõtteliselt samadel alustel, nagu Pärn ja Simm liigitasid eesti keele põhisõnavara.

Eesti keele tuumsõnavara leidmist raskendas eesti keele muutevormide rohkus. See oli takistuseks õigete algvormide leidmisele. Teiseks raskuseks oli see, et korpusel puudub suuline kõne, ning kolmandaks, et korpus kannab oma aja pitsert. Mis viimasesse puutub, siis olid Tartu kirjakeele korpusel saja sagedasema sõna hulgas näiteks sõnad *liit* ja *nõukogu*, mis ilmselt pärinesid ajakirjanduses ja tollal mujalgi sagedasest väljendist *Nõukogude Liit*, mida praeguses eesti keeles kasutatakse suhteliselt harva. Kasutades Veenkeri sõnade kustutamise printsiipi jätsime ideologismid ja muidu kahtlased vormid (nt liitsõnaosad) eesti (kirja)keele tuumsõnavara hulgast välja.



# USAGE OF THE VERB *andma* 'to give' IN THE DOMAIN OF VERBAL COMMUNICATION<sup>1</sup>

Ilona Tragel

*University of Tartu*

## 1. Introduction

The concept GIVE plays a central role in human relations. Our experience of giving relations has become organized into a system in our cognition, which enables us to describe and recognize the descriptions of giving acts by means of language. That we can perceive the similarity between the giving of concrete things and the 'giving' of abstract units (comparison processes of perception) belongs to this system, too.

The carrier of the concept GIVE is usually a verb (or a grammatical element that has developed from this verb), which prototypically indicates that an entity THING passes from the possession of one entity GIVER to the possession of another entity RECIPIENT. At this the transferred entity is smaller than GIVER and RECIPIENT who are in prototypical relations. GIVER is usually AGENT, and RECIPIENT is PATIENT. The following relations between the participants in the act of giving can be observed: at the beginning of the act THING is in the possession of GIVER and/or under his or her control – in some kind of contact with GIVER; at the end of the act, however, THING is in the possession of RECIPIENT and/or under his or her control – in contact with RECIPIENT. THING is the movable entity of the process/relation, and therefore it is in the focus of the interpreter's attention, the direction of the movement of THING is from GIVER to RECIPIENT. The relations between the participants in the act can be illustrated with the following figure

---

<sup>1</sup> This work was supported by the Estonian Science Foundation, Grant No.: 4405.

(the concept of partner object comes from Birute Klaas (1996: 53)).

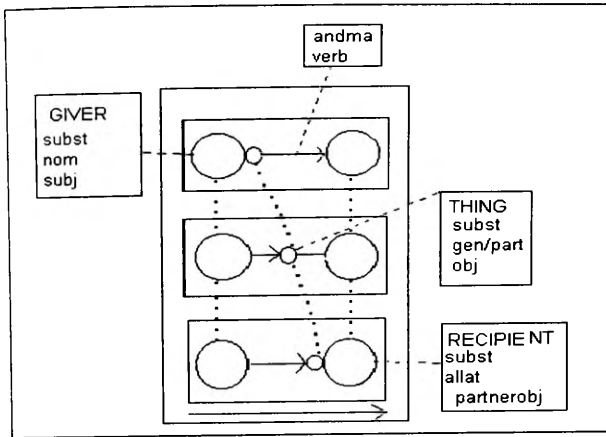


Figure 1.

### Participants of GIVE together with the construction schema

In a typical case, the giving relation is expressed by a verb, for example Estonian *andma*, English *give*, Russian *dat*, German *geben*, French *donner*, Spanish *dar*, Italian *dar*, etc. There are languages where the verb that expresses this relation has developed into a grammatical element. For example, in Ewe<sup>2</sup> the giving verb *ná* has a prepositional meaning 'FOR' or 'TO' in certain contexts.

(1) *me-wo do ʼvévié ná dodókpo lá*

'1SG-do work hard give exam DEF' (Heine *et al* 1991: 1–2)

In Estonian the giving relation is expressed by the verb *andma* 'to give', which usually retains its verbal character. The Estonian verb *andma* can express meanings that belong to the following cognitive domains:

1. POSSESSION RELATIONS (*Ma annan sulle õuna* 'I'll give you an apple') (the nucleus of this domain is the literal meaning of *andma*);

<sup>2</sup> A language belonging to the Kwa branch of the Niger-Congo language family that is spoken in eastern Ghana and southern Togo and Benin (Heine *et al* 1991: 1).

2. PROPERTY (*Ta andis mulle öömaja* 'He gave me a shelter for night');
3. CAUSATION and/or CREATION (*Õunapuud annavad sel aastal head saaki* 'The apple-trees will yield a good crop this year');
4. VERBAL COMMUNICATION (*Ta andis Hommikuteleviisioonile intervjuu* 'He gave an interview to Morning TV');
5. MANIFESTATION (*Lamp annab valgust* 'The lamp gives off light');
6. MANIPULATION (*Ta andis oma tütre naabrimehele naiseks* 'He married his daughter to his neighbour');
7. EVENT FOR THE PUBLIC (*Vanemuine andis külalisetendusi Tallinnas* 'The Vanemuine Theatre gave some guest performances in Tallinn');
8. BEATING (*Jüri andis vennale peksa* 'Jüri beat up his brother');
9. Subjectified uses (*Temaga annab rääkida* 'He can be talked into' *Rongi annab oodata* 'The train won't arrive soon').

In this article I will analyze the use of the verb *andma* in the domain of VERBAL COMMUNICATION, e.g.

- (2) *Pean Teie ees vabandama, kuid Teile usutavasti mõistetavatel põhjustel ma ei anna praegu ajakirjandusele intervjuusid.* (aja90)

'I have to apologize to you, but for reasons that are presumably understandable to you I'm not giving interviews to the press now'

- (3) *.. ja (--) politsei andis allkirju seal.* (kõk)

'.. and the police gave signatures there'

One of the aims of this article is to contrast the uses of the communicative domain of the Estonian verb *andma* with the corresponding uses of other languages.

According to Rudzka-Ostyn (1988), it is possible that the 'moving' of 'linguistic objects' by means of expressing moving and movement of concrete objects is a universal linguistic

phenomenon, which is caused by various cognitive processes (Rudzka-Ostyn 1988: 507, 550).<sup>3</sup>

The article is based on the analysis of almost 400 sentences that contained the verb *andma*,<sup>4</sup> the syntactic object of which or some other argument that is defineable as the landmark of the relation in the terminology of cognitive grammar (Langacker 1987: 231–243) establishes the domain of verbal communication. The data concerning those languages that I am unfamiliar with come from dictionaries and interviews with native speakers. In addition, example sentences from two books are used: Newman 1996 and Newman (ed.) 1997 (I'll try to retain as far as possible the original transcription in the example sentences that come from these sources). The choice of languages is random. When establishing the correspondence of the lexeme that designates the object of GIVE, I will mostly proceed from prototypicalness. At the corresponding sets of expressions I will discuss also those Estonian examples that suit this concept group and the use of which may be motivated by the main expression under discussion (e.g. under *käsku andma* 'to give a command' will be provided also examples illustrating *korraldust andma* 'to give an order' and *juhiseid andma* 'to give instructions').

## 2. Verbal THINGS

The main body of the article (this section) discusses predications where:

- 1) the process of 'sending' a verbal THING is expressed by the verb *andma*;

---

<sup>3</sup> The Estonian language has a large number of examples to illustrate this point: *jutuga edasi minema* 'to go on with the story' *kellegi kaitseks välja astuma* 'to stand up for sb' *argumente välja laduma* 'to come up with the arguments' *jutulõnga üles võtma* 'to pick up the thread of a conversation' *ideega välja tulema* 'to come up with an idea' *detailidesse laskuma* 'to go into detail' *keerukamatest küsimustest üle hüppama* 'to skip more complicated questions'; compounds *ettekanne* 'paper' and *ettekandja* 'speaker'; the use of space adverbs in these expressions is also a topic in its own right.

<sup>4</sup> The data come from my MA thesis (Tragel 1999).

- 2) the verbal THING is audible and/or visible;<sup>5</sup>
- 3) GIVER and RECIPIENT are usually animate entities, mostly humans.

### 2.1. *sõna andma* 'to give the floor'

The word is the prototypical verbal unit. This observation can be clearly proved with help of the expression *sõna andma* 'to give the floor' in various languages, where the word combination means 'to give the conversational turn to sb' Lakoff and Johnson (1980: 10–12) follow Michael Reddy's idea of *conduit metaphor*, according to which how we speak about language is structured by three metaphors: IDEAS (OR MEANINGS) ARE OBJECTS (*Sain selle idee Hellelt* 'I got this idea from Helle'); LINGUISTIC EXPRESSIONS ARE CONTAINERS (*Võtan selle lõigu ka artiklisse sisse* 'I will include this passage in the article as well'), and COMMUNICATION IS SENDING (*Andsin sõna järgmisele esinejale* 'I gave the floor to the next speaker'). Some aspect of structuring may be concealed in the case of some expressions. Of the following examples the last one has a clearly active character.

- (4) Estonian: *annan talle sõna* 'I'll give him the floor'
- (5) Spanish: *dar la palabra*
- (6) Russian: *Ja daju jemu slovo*
- (7) German: *Ich werde ihm das Wort geben*

In Papuan Rumu, however, the same expression means giving an order or a command:

- (8) Rumu: *hei ta-ra*  
'word give-Purposive' (Newman 1996: 141)

In the giving process that is expressed by the verb *andma* RECIPIENT is PATIENT whose activity in the process is usually minimal (but not without exception, see 2.7.2.). In the pattern with an active RECIPIENT the perspective is contrary to giving, RECIPIENT is TAKER and AGENT, and the process is usually

---

<sup>5</sup> Usually in this material a visible entity is verbal as well (e.g. written text), but sometimes it can have a sign-like or symbolic character (e.g. gestural command).

expressed by the verb *võtma* ‘to take’ (*Järgmine kord ma võtan ise sõna* ‘Next time I will take the floor myself’). Estonian has even compounds *sõnavõtt* (word + taking) ‘speech’ and *sõnavõtja* (word + taker) ‘speaker’. The verb *andma* and its equivalents in other languages have rather universally a causal meaning as well (Newman 1996: 172): GIVER causes RECIPIENT to get THING. In the examples of this group of expressions GIVER (in the prototypical situation a person who has control over what is happening) causes the transfer of the conversational turn to RECIPIENT.

## 2.2. *sõna andma* ‘to make a promise’

In a large number of languages the word combination *sõna andma* means also to make a promise.

(9) Estonian: *Annan sulle oma sõna* ‘I’ll give you my word’

(10) Finnish: *Annan sinulle sanani*

(11) English: *I give you my word*

(12) French: *donner sa parole*

(13) German: *Sie gab ihr Wort*

(14) Bulgarian: *Davam ti дума-та си* (Newman 1996: 141)

(15) Russian: *Daju tebe slovo*

(16) Cora: *n<sup>y</sup>a=tiih n=i t<sup>y</sup>ú-hu -u-tá- a í n<sup>y</sup>e-n<sup>y</sup>uuka*

‘then I gave my word’ (Casad 1997: 146)

In this schema, too, at least in Estonian there is an opposite. We can say *Ma võtsin talt sõna, et ta selle tagasi toob* ‘I made him promise to bring it back’; the verb *olema* ‘to be’ expresses abstract possession: *Mul on selle kohta tema sõna* ‘I have his promise concerning this’. Also, the elliptical use is quite common: *Minu sõna selle peale* ‘my word upon it’

## 2.3. *nõu andma* ‘to give advice’

*Nõu* ‘advice’ is a suitable extension of the literal sense<sup>6</sup> of *andma* because in this relation THING (here *nõu*) is usually an entity

<sup>6</sup> A predication that expresses possession has a literal meaning, where GIVER and RECIPIENT are humans and THING is an item that moves

from which RECIPIENT can presumably yield a profit. Prototypically RECIPIENT is BENEFICIENT. The following Estonian compounds stress the importance of giving advice in human relations: *nõuanne* 'advice' (THING), *nõuandja* 'adviser' (GIVER) and *nõuandla* 'counselling centre' (PLACE).

(17) *Kahjuks on toitude ettetoomine lausa katastroofiline, laua-kaaslastele anti nõu toitu üldse mitte tellida, kuna selle laudatoimetamine võtaks kaua aega.* (aja90)

'Unfortunately the food service is disastrous, the fellow diners were recommended not to order any food to the table at all because its delivery would take a long time'

(18) *Euroteaduskonna raames tegutsevad külalislektorid peavad andma nõu ja abistama meie kolleege korraliku õppekava loomisel.* (aja90)

'The visiting lecturers who are working within the framework of the Eurofaculty must give advice and help our colleagues in setting up a decent curriculum'

(19) Finnish: *Annan sinulle neuvoa*

(20) English: *I give you advice*

(21) Polish: *Marysia dala mi rade* (Newman 1996: 143)

(22) Papuan Rumu: *pahe ta-ra*

(23) Swahili: *pa shauri* (Newman 1996: 137)

(24) Bulgarian: *davam savet* (Newman 1996: 137)

(25) French: *donner des conseils*

(26) German: *Rat geben*

The Estonian phrase verb (*väljendverb*) *nõu andma* has been considered to be a German influence. The abundance of similar expressions in various languages indicates that such conceptualization must be substantiated by the structure of cognition as well.

For this expression the schema of giving and taking is not really valid: in Estonian one cannot say *\*nõu võtma* 'to take advice' *Nõu saama* 'to get advice' is fully possible, in addition,

---

in the same time and space with GIVER and RECIPIENT (Tragel 1999: 23).

the expressions *nõuks võtma* ‘to decide’ and *nõu pidama* ‘to discuss’ are used, too.

#### 2.4. *käsku andma* ‘to give a command’, *luba andma* ‘to give permission’

*Käsk* ‘command’ is a suitable argument of *andma* in the role of THING. In such conceptualizations the domain of INSTITUTIONAL RELATIONS is active: GIVER has a higher position than RECIPIENT in the relevant hierarchy, his status enables him to ‘command’ RECIPIENT. *Käsk* ‘command’ may have a signlike character: it can be a conventional sign, for example, a gesture.

(27) *Koer ei ole nagu inimene, et saab sõnadest aru üksipuha mis käskluse sa annad* (kõk)

A dog is not like a human who can understand words no matter what command you give’

(28) *8. augustil andis Norra valitsus sõjaväele kõrgendatud valmisoleku käsu*

‘On August 8 the Norwegian government issued an order to put the troops on a high alert’

(29) *Tema jaoks on kokkulepitud hüpete arv täis, tema annab käsu present kokku rullida, kunsttoss summutada*

‘For him the agreed number of jumps has been achieved, he issues an order to roll up the canvas, to put out the artificial smoke’

(30) Finnish: *Hän antoi käskyn*

(31) English: *He gave an order*

(32) Bulgarian: *davam zapoved* (Newman 1996: 137)

(33) German: *Befehle geben*

(34) Spanish: *dar la ley*

(35) Russian: *dat prikaz*

In these conceptualizations causality is equally important.

The position of THING can be taken by a large number of entities that act similarly to *käsk* ‘command’, for example, *korraldus* ‘order’, *luba* ‘permission’ *juhis* ‘instruction’. It was pointed out in 2.1. that the meaning of the Rumu expression *hei*



*ta-ra* (word give-Purposive) is 'to give an order' (Newman 1996: 141).

(36) *Haridus- ja kultuuriminister andis korralduse tüdruk kooli tagasi võtta.* (aja90)

'The minister of education and culture issued an order to reinstate the girl to her school'

The following examples illustrate giving permission:

(37) *Asi ei ole veel lõplik, loa peab andma Euroopa komisjon.* (aja90)

'The thing is not final as yet, the European Commission must give its permission'

(38) English: *give permission*

(39) Italian: *dare il permesso* (Newman 1996: 183)

(40) Swahili: *pa ruhusa* (Newman 1996: 183)

(41) Malay: *mem-beri izin* (Newman 1996: 183)

GIVER can also be a personified entity (cf. also GIVER as an intermediary):

(42) *Täpsed juhised tegutsemiseks annab jahikorralduskava.* (ESBI)

'The plan of organizing hunting provides detailed guidelines for action'

#### 2.4.1. *infot andma* 'to give information'

There are many parallels but also some rather systematic differences between the usage structure and partial characteristics of the verb *andma* that express verbal communication and the same parameters of the literal meaning of *andma*. In the literal sense THING is transferred from GIVER to RECIPIENT in the process of giving; by the end of the process it is usually in the possession of RECIPIENT and GIVER does not have this THING anymore. In the processes that express verbal communication usually the latter aspect is not included from the source domain: THING will belong to GIVER also after the giving process is over, being at the same time in the abstract possession of RECIPIENT. This process could be called THING sharing (cf. also the expressions *teadmisi, infot jagama* 'to share knowledge,

information') (partly Newman 1996: 138). The source and target domains are similar in that the giving of a real THING brings about some change in the real world (THING changes its owner or possessor), which is an important aspect of interpersonal communication and is often 'registered' linguistically. An act of verbal communication establishes a change as well: RECIPIENT receives something that he presumably did not have before.

- (43) *Soovitav on saadud katseandmed vormistada järgmise lihtsa tabelina, mis annab küllaldaselt esialgset informatsiooni selleks, et otsustada katseandmete usaldatavuse ja kasutatava mõjuri efekti üle.*

'It is recommended to present the experimental data in the form of the following simple table, which provides enough preliminary information to judge on the reliability of the experimental data and the effect of the factor used'

- (44) Finnish: *Asiasta annettiin tietoja*  
 (45) English: *give information to someone*  
 (46) Bulgarian: *davam informatcija* (Newman 1996: 137)  
 (47) Russian: *dat informatciju*  
 (48) German: *Informationen geben*

In Estonian *andma* accepts verbal information as an argument (role of THING) rather freely but not without exception (see e.g. 2.6.).

- (49) *m oota nüüd ma pean sis üles kirjutama ma mõtlen et sa annad mulle kohe mingi faili nime või midagi lihtsamat (kõik)*  
 'Wait a little, I must then write down. I think that you will give me now the name of some file or something simpler'  
 (50) *ee jah aga seal [vist oligi] nimodi=et seal oli mingi: e mingi (.) valeinformatsioon talle antud (kõik)*  
 'eeh yes, but it was [must have been] so that there was: he had been given some misinformation'  
 (51) *Online "tõlkeabi" Anna dokumendi URL*  
 (<http://ee.www.ee>)  
 'Online 'translation assistance' Give the document's URL (<http://ee.www.ee>)  
 (52) *Kas sul on midagi puudu? – Numbri annab "Kontakt"*  
 (telefoniraamatu reklaam ETVs)

'Do you miss anything? – The number will be supplied by "Kontakt" (phone-book commercial on TV)'

In these conceptualizations GIVER is mostly the intermediary of THING and not SOURCE.<sup>7</sup> Sometimes the mediation is elaborated by emphasizing the medium:

(53) K: *mina arvan küll e noh minul ei ole selle vastu midagi kui ma selle saan aga mina olen ainukene kes selle läbi lappab ma annan **informatsiooni** suuliselt edasi*

'K: well, I think that well I have nothing against getting it, but I'll be the only one who thumbs it through, and I will forward the information by word of mouth'

The computer is in the role of GIVER, and it provides (mediates) some information (THING) to the user (RECIPIENT). (One can conclude from this schema that in Estonian the computer is conceptualized as a human participant in the principal schema of human relations.)

(54) *Näedsa nüüt-ta [arvuti] andis mulle jälle mingi veateate. (kõnek)*

'You see now it [the computer] gave me an error message again' (spoken)

(55) *Neid ridu ta [arvuti] mulle siia alla enam ei antki. (kõnek)*

'It [the computer] didn't give me these lines at all under this' (spoken)

## 2.5. *vastust andma* 'to give an answer', 'to be responsible'

*Vastus* 'answer' is also a verbal THING that can be 'moved' with the verb *andma*. Usually *vastust andma* 'to give an answer' differs from the verb *vastama* 'to answer' in that it emphasizes the moment of action (here: answering).

(56) Estonian: *Ta andis vastuse*

(57) Finnish: *Hän antoi vastauksen*

(58) English: *He gave an answer*

(59) German: *Antwort geben*

<sup>7</sup> Participants in the act are described in greater detail in Tragel 1999: 20.

It is unacceptable in Estonian to say *\*küsimumst andma* 'to give a question', one can say *küsimumst esitama* or *küsimumst küsima* 'to ask a question'. However, questions in writing can be given, for example, in the classroom context *Ma annan teile nüüd kontrolltöö küsimused* 'I'll give you now the questions of the test'. It is likely that the literal meaning may act as the motivator here, that is, the questions are usually on sheets that will be handed out. In the classroom context, however, the expression *annan küsimused* 'I'll give the questions' can be applied to the oral forwarding of questions as well.

The expression *millegi eest vastust andma* is used figuratively in Estonian meaning 'to be responsible for sth'. This expression may be used productively also because its phonetic shape resembles the closest literal correspondence of the figurative meaning *vastutama* 'to be responsible' (or perhaps with another figurative equivalent *vastutust kandma* 'to bear responsibility').

(60) *Ettevõtjal tuleb vastust anda oma võlausaldajate, riigil valijate ees*

'Entrepreneurs are held accountable before their creditors, the state before the electorate'

## 2.7. VERBAL THINGS and the domain of social relations

### 2.7.1. *etendust andma* 'to give a performance'

#### and *kontserti andma* 'to give a concert'

For the following conceptualizations, in addition to the domain of verbal communication, the domain of social relations is at least equally active. THING is an event and the public is RECIPIENT, I have called this domain provisionally EVENT FOR THE PUBLIC domain (Tragel 1999). In some cases the verballity of THING in these arguments has a highly general character: a performance can be largely intended for the sense of seeing. Nevertheless, the verb *andma* seems to show some 'preference to verballity': ballet, which we enjoy mostly visually, is not usually given, cf. *\*trupp andis balleti* 'the company gave a ballet' (though e.g. *trupp andis balletietenduse* 'the company gave a ballet performance' is acceptable). However, *ansambel andis kontserdi*

'the band gave a concert' is common although visual perception should have the main role in the processes with a literal meaning – giving or receiving some object is likely to be first and foremost visible.

(61) *Filharmonia Kammerkoor ja Tallinna Kammerorkester sõitsid Ameerika Ühendriikidesse, kus nad annavad dirigent Tõnu Kaljuste juhendamisel kaheksa kontserti.* (aja90)

'The Philharmonic Chamber Choir and the Tallinn Chamber Orchestra flew to the United States, where they will give eight concerts conducted by Tõnu Kaljuste'

(62) English: *give a concert*

(63) Russian: *dat kontsert*

(64) German: *ein Konzert geben*

(65) Bulgarian: *davam kontcert* (Newman 1996: 143)

(66) *Teatrikuul toimus 10 esietendust, anti 420 etendust 168000 külastajale, ..*

'The theatre month witnessed 10 premières, 420 performances were given to 168,000 theatregoers, ..

(67) English: *give a performance*

(68) French: *donner une pièce de théâtre*

However, in Italian the same expression means the staging of a play:

(69) Italian: *dare una commedia* 'to put on a play'

### 2.7.2. *tundi andma* 'to give lessons, to teach'

In the prototypical giving relation RECIPIENT is BENEFICIENT, GIVER has something (THING) by which RECIPIENT will get richer (see also 2.5.). In the following conceptualizations THING is knowledge or skill 'in the possession' of GIVER, which is moved in the giving process. This meaning of passing on knowledge or a skill is likely to have extended into the expression of the event of giving knowledge or a skill. Similarly to the meaning of the previous group of expressions this arrangement, too, is related to the EVENT FOR THE PUBLIC domain.

Similarly to the domain of verbal communication in general, in this group, too, **THING** is, so to say, a shared value – it is true that it moves in the giving process, but **GIVER** will keep it, too.

(70) *Paides valiti ta 1919. aastal Poeg- ja Tütarlaste-gümnaasiumi õpetajaks, ühtlasi hakkas ta tunde andma ka teistes koolides.*

'He was elected as a teacher of the Coeducational Gymnasium in Paide in 1919, at the same time he began to give lessons at other schools as well'

(71) *Ta annab nüüd meie vanas koolis matemaatikat*

'He is teaching math at our old school now'

(72) Finnish: *Hän antaa pianotunteja*

'He gives piano lessons'

(73) English: *give lessons*

(74) Bulgarian: *davam urotci* (Newman 1996: 142)

(75) Russian: *dat uroki*

RECIPIENT of the conceptualization *tundi andma* 'to give lessons' is not a prototypical PATIENT of the giving relation: he is expected to be active in accordance with the agreed rules. On the level of speculation one might ask why is it impossible to say in Estonian *\*loengut andma* 'to give a lecture' In English this conceptualization is very common:

(76) English: *to give a lecture*

The same is true about *kõne* 'speech' In Estonian we say *loengut pidama* 'to hold a lecture' and *kõnet pidama* 'to hold a speech' but, for example, in English and Bulgarian you can 'give' them as well:

(77) English: *to give a speech*

(78) Bulgarian: *davam retch* (Newman 1996: 142)

The verb *andma* is used in Estonian (at least at university) for the designation of at least one teaching format:

(79) *Ta annab eripedagoogikas ka seminare*

'He teaches seminars in special education as well' (from a conversation overheard at the Department of Estonian)

Continuing the speculation one might ask if a seminar is a teaching format that presumes more activeness from RECIPIENT than from a person who attends a lecture?

There is no German influence on Estonian here because the German expression is *eine Stunde halten* 'to hold a lesson' (the verb *halten* 'to hold' expresses also the processes of delivering lectures and speeches: *einen Vortrag halten* 'to make a presentation', *eine Rede halten* 'to make a speech').

All the previous uses of *andma* are also related to the causality of the concept 'to give'. Relations can be established also with the domain of CREATION (cf. Lakoff and Johnson 1980: 69–76). In the following examples it is difficult to decide which domain is primary in the positioning of predication:

(80) .. *et riik on sulle nagu ühe hariduse juba andnud (kõk)*  
 '.. that the state has already given you an education'

(81) .. *kahe päeva jooksul saaksime ühele grupile anda sis (.) selle teooria mis mida peaks üks müüja teadma (kõk)*  
 '.. during two days we could then give one group the theory that a salesperson should know'

### 2.7.3. *tervitust, õnnitlust üle/edasi andma* 'to convey a greeting, congratulation'

The verb *andma* can also express the mediation of greetings and congratulations. The Estonian expression *tervitust üle andma* 'to convey greetings' has a constraint in the properties of GIVER: GIVER is usually the mediator of THING and not the author (cf. also 2.5.). *Õnnitlust üle andma* 'to convey congratulations' may mean either mediating somebody's congratulations (*Ma annan sulle üle õnnitluse ka oma ema poolt* 'I'll also convey my mother's congratulations to you') or the moment of extending the author's own congratulations (*Ma andsin üle oma südamliku õnnitluse* 'I extended my own cordial congratulations'). Estonian uses phrasal verbs in these expressions. The adverbial components *üle* and *edasi* of phrasal verbs, the primary meaning of which denotes place and movement, emphasize the course of progress in predication.

- (82) *Olin Urmas Oti “Carte Blanche”-raamatu presentatsioonil, kus Urmas ütles, et oli Maiaga kohtunud, ja andis mulle ter-vitused edasi.* (aja90)

‘I attended the presentation of the book “Carte Blanche” by Urmas Ott, where Urmas said that he had met Maia and conveyed her greetings to me’

- (83) English: *Give my best wishes/ regards to Mary*

The conceptualization of English GIVER acts as a mediator as well.

- (84) *Majandusminister andis üle ka peaministri õnnitluse*

‘The minister of economy conveyed the prime minister’s congratulations as well’

Bulgarian has the verb prefix *pre-*, which has developed from the preposition *prez* ‘through’, and it is a sign of the mediator’s role in this conceptualization (example 85). The same pattern can be found in Russian (86).

- (85) Bulgarian: *pre-davam pozdravi* (Newman 1996: 140)

- (86) Russian: *predat pozdravlenija*

#### 2.7.4. *õnnistust andma* ‘to give a blessing’

It is characteristic of the literal meaning of *andma* that the participants in the process undergo some change in the course of the process (see also 2.5.). In the case of a figurative meaning this circumstance may motivate the use of performatives, where a participant’s utterance changes the status or condition of the participants (Newman 1996: 141). This use may have been influenced by the schema of divine giving in the religious context, cf. *Kui jumal annab ..* ‘if the Lord gives .. or *jumalast antud* ‘given by the Lord’ At least in Estonian the meaning of the expression *õnnistust andma* ‘to give a blessing’ seems to develop toward ‘wishing well’

The utterance retains a performative connotation, however.

- (87) Estonian: *Ta andis tütrele oma õnnistuse*

‘He gave a blessing to his daughter’

- (88) Finnish: *antaa siunauksensa*



- (89) English: *to give a blessing*  
 (90) Italian: *dare la benedizione*  
 (91) Bulgarian: *davam blagoslavija* (Newman 1996: 137)  
 (92) German: *seinen Segen geben*  
 (93) Russian: *dat blagoslovenije*

In fact, the Javanese expression that has the same origin has a generalized meaning 'to wish sb well'

- (94) Javanese: *awéh slamet* 'give a greeting/blessing; to wish someone well' (Newman 1996: 137)

The Estonian expression *õnnistust andma* 'to give a blessing' has an extended use that expresses 'promise' Nevertheless, the use of the following example sentence could be related to the so-called schema of divine giving as well:

- (95) *Teisedki Euroopa riigid ei kõhelnud toetamast, niipea kui Lenin oma õnnistuse andis.* (aja90)  
 'Nor did the other European countries hesitate to support as soon as Lenin gave his blessing'

## 2.8. Symbol-like verbal THINGS

One part of THINGS, which have a verbal character, would rather have a symbolic meaning in the giving process. The transfer of a symbol-like object is expressed, for example, by *nime andma* 'to give a name (example 97), which in turn extends to a figurative use (in the meaning 'to name sth somehow, example 98).

- (96) *Pojale anti nimeks Jaan*  
 'The son was given the name Jaan'  
 (97) *Rahvas kutsus seda Valuojaks ja oja omakorda oli nime andnud Wana talule.* (ILU)  
 'People called it Valuoja and oja 'stream' had in turn given its name to Wana farmstead'  
 (98) *Sellele loole ei oska kohe nimegi anda*  
 'One cannot even give a name to this story'

It is likely that in the following examples the meaning of giving a verbal THING may motivate the giving of THINGS that are similar to linguistic symbols:

(99) *Põhimõte oli järgmine: esikoht mingil alal annab ühe punkti, teine koht kaks punkti jne.* (aja90)

'The principle was as follows: the first place in an event gives one point, the second place two points, etc'

(100) *Kõik osalejad andsid enda meelest suurimale uudisele kümme punkti, järgmisele üheksa ja nii edasi.* (aja90)

All the participants gave ten points to what they considered the top news, eight to the next one, and so on'

### 3. Conclusions

Comparison is a characteristic feature of our perceptual processes. A new object is compared with what has already been experienced and recorded in the system and is classified according to similarity or suitability. The usage of the verb *andma* in the domain of verbal communication indicates similarity between the perception of transfer of a concrete item and a verbal unit.

Similarly to many other languages, in Estonian, too, the verb *andma* is undergoing grammaticalization (for grammaticalization in the broader sense see Heine *et al* 1991). The causal meaning of *andma* must serve as a reason for this (see 2.1. and 2.4.). It is equally important that *andma* belongs to those linguistic means by which the basic and daily interpersonal relations are expressed (e.g. giving all kinds of information, see 2.5.). Parallels between the participants, their characteristics, and the event structure of the primary meaning and the extended meaning play an important role in the motivation of meaning extensions. However, systematic differences are very common as well. For example, in the literary meaning of the *andma* process THING is passed from GIVER to RECIPIENT. At the end of the process it is in the possession of RECIPIENT, whereas GIVER does not have this THING anymore. In the domain of verbal communication the latter aspect is usually not included from the source domain. THING will remain in the possession of GIVER also after the giving process is over, and at the same time it will be in the abstract possession of RECIPIENT as well (e.g. in the groups *infot andma* 'to give information' and *nõu andma* 'to give advice').

Among the frequent uses of the domain of verbal communication are productive those schemas that affect language use. For example, if it is possible to say *sõna võtma* 'to take the floor' then it is also possible to say *sõna andma* 'to give the floor' and *sõna saama* 'to get the floor'. These schemas operate, for example, in the expressions *sõna andma* and *sõna saama*. This use can be related to the fact that a certain degree of activeness is expected from RECIPIENT (in accordance with the rules of the schema).

This schema does not operate perfectly, however, in *nõu andma* 'to give advice' *käsku andma* 'to give a command' *luba andma* 'to give permission' *infot andma* 'to give information' and *õnnistust andma* 'to give a blessing'. For example, in Estonian one can say *nõu andma* 'to give advice' and *nõu saama* 'to get advice' but one cannot say *\*nõu võtma* 'to take advice' (*nõuks võtma* 'to decide' is possible, however, see 2.3.). In these schemas no activeness or agent-likeness is expected from RECIPIENT. Various differences are revealed when languages are compared (see e.g. 2.1. and 2.7.1.).

It is doubtful whether such a schema operates also in *tervitusi üle/edasi andma* 'to convey regards'. It is possible to say *tervitusi kätte saama* 'to receive the regards' but not *\*tervitusi võtma* 'to take regards'.

The schema is not valid at all in *etendust, kontserti andma* 'to give a performance, a concert' which belongs to the domain of social relations. Thus, both the expressions *\*kontserti/etendust võtma* 'to take a concert/performance' (from the point of view of the event's recipient) and *\*kontserti/etendust saama* 'to receive a concert/performance' are unacceptable.

A similar situation can be observed in *tundi andma* 'to give lessons'. However, some fossilized expressions point to the possibility that under certain conditions and in certain meanings this schema is still possible. Thus, it is possible to say in Estonian *võtta* or *saada eesti keele tunde* 'to take or receive Estonian lessons', but one cannot say *\*võtta* or *\*saada eesti keele tundi* 'to take or receive an Estonian lesson' (With the exception of the meaning where *tundi võtma* 'to take lessons' is used in the meaning 'to give lessons'). As in this expression THING must be

in the plural, we can consider it a fossilized expression, and thus the schema is unproductive. In addition, one can say *sain oma õppetunni* 'I learned my lesson' and *andsin õppetunni* 'I taught a lesson' (but not *\*võtsin oma õppetunni* 'I took my lesson').

Interrelations between language and thought constitute the central problem of cognitive linguistics. It is discussed whether the linguistic differences can cause differences in thought, which would in turn give rise to behavioural differences. On the basis of the present article one might ask, for example, whether the behaviour of Estonian and English teachers when giving their lessons would be different from a German teacher who 'holds a lesson' (*eine Stunde halten*). Or is an English-language lecturer, who gives a lecture, closer to the giving process also in real life; is he different from the Estonian-language and German-language lecturers who 'hold lectures' Or why is it so that not all of the linguistic units expressing verbal units are suitable as arguments of *andma* (cf. e.g. Estonian *\*kõnet andma* 'to give a speech', *\*laulu andma* 'to give a song' (see 2.7.2.)?

The meaning of *andma* has not become generalized to a degree that one could unconditionally state that it has become grammaticalized. However, its meaning is general enough so that instead of counting the usage possibilities it would be expedient to speak about constraints (although their number is much smaller). Consequently, there is some ground for the claim that regardless of the universal nature of cognitive structure each language has something intrinsic to the world-picture of its native speakers, and it is valid on the lexical level as well.

## ABBREVIATIONS

- kõk: Spoken Estonian Corpus of the University of Tartu  
 kõnek: other source references: examples of spoken language collected by the author  
 aja90: Corpus of Written Estonian, 1990s journalism  
 ILU: Corpus of Written Estonian, fiction

## REFERENCES

- Casad, Eugene 1997:** Lots of ways to GIVE in Cora. – John Newman (ed.), *The Linguistics of Giving*. Typological Studies in Language 36. John Benjamins Publishing Company. Amsterdam/Philadelphia.
- Heine et al 1991 = Heine, Bernd; Claudi, Ulrike; Hünemeyer, Friederike 1991:** *Grammaticalization. A Conceptual Framework*. The University of Chicago Press. Chicago.
- Klaas, Birute 1996:** Similarities in Case Marking of Syntactic Relations in Estonian and Lithuanian. – Mati Ereht (ed.), *Estonian: Typological Studies I*. (Tartu Ülikooli eesti keele õppetooli toimetised 4. Publications of the department of Estonian of the University of Tartu 4.) Tartu. 37–67.
- Lakoff, George; Johnson, Mark 1980:** *Metaphors we live by*. The University of Chicago Press. Chicago.
- Langacker, Ronald W. 1987:** *Foundations of cognitive grammar, vol. I. Theoretical Prerequisites*. Stanford University Press. Stanford, California.
- Newman, John 1996:** *Give. A Cognitive Linguistic Study*. Cognitive linguistics research 7. Mouton de Gruyter. Berlin.
- Newman, John (ed.) 1997:** *The Linguistics of Giving*. Typological Studies in Language 36. John Benjamins Publishing Company. Amsterdam/Philadelphia.
- Norvig, Peter; Lakoff, George 1987:** *Taking: A Study in Lexical Network Theory*. – Aske et al (eds.), *Proceedings of the 13th Annual Meeting of The Berkeley Linguistics Society*. Berkeley, California. 195–206.
- Rudzka-Ostyn, Brygida 1988:** *Semantics extensions into the domain of verbal communication*. – Brygida Rudzka-Ostyn (ed.), *Topics in Cognitive Linguistics. Current Issues in Linguistic Theory* 50. Benjamins. Amsterdam/Philadelphia. 507–554.
- Tragel, Ilona 1999:** *Verbi andma kognitiivlingvistiline analüüs*. Tartu Ülikool. (MA thesis, ms)

## VERBI ANDMA KASUTUS VERBAALSE SUHTLUSE VALDKONNAS

Ilona Tragel

1. Andmise mõistel on inimsuhetes keskne koht. Kogemused andmis-suhetest on meie tunnetuses organiseerunud süsteemi, mis võimaldab meil muuhulgas keele abil kirjeldada ja ära tunda andmise aktide kirjeldusi. Sellesse süsteemi kuulub ka see, et tajume sarnasust konkreetsete esemete andmise protsessi ja abstraktsete üksuste "andmise" protsessi vahel.

Tüüpjuhul väljendab andmissuhet verb, näiteks eesti keeles *andma*, inglise keeles *give*, vene keeles *dat*, saksa keeles *geben*, prantsuse keeles *donner*, hispaania keeles *dar*, itaalia keeles *dar* jne. See verb väljendab seda, et mingi olem (*entity*) siirdub ühe olemi valdusest teise olemi valdusse. On keeli, kus sellist suhet väljendavast verbist on kujunenud grammatiline element. Näiteks eve keele andmisverb *ná* on teatud kontekstis otstarbeadpositsiooni funktsioonis: *me-wo do 'vévíé ná dodókpó lá* 'Ma töotan kõvasti eksamiks/eksami jaoks' (Heine *et al* 1991: 1–2).

Artiklis analüüsin verbi *andma* ja selle teiste keelte vastete kasutust VERBAALSE SUHTLUSE valdkonnas, näiteks: .. *ja politsei andis allkirju seal. Ich werde ihm das Wort geben*. Selleks olen analüüsinud ligi 400 verbi *andma* sisaldavat lauset, mille süntaktiline objekt või muu suhte pidepunktiks määrateldav argument sätestab verbaalse suhtluse valdkonda. Teiste keelte kohta olen materjali kogunud sõnaraamatutest ja emakeelseid kõnelejaid küsitledes, lisaks kasutan näitelauseid kahest raamatust: Newman 1996 ja Newman (ed.) 1997. Näidete keelte valik on juhuslik. *Andma* vastete objekti e pidepunkti tähistava lekseemi vastavuse määramisel lähtun peamiselt prototüüpsusest. Vastavate väljenditühmade juures käsitlen ka selliseid eesti keele näiteid, mis selle mõisterühma juurde sobivad ning mille kasutus võib olla motiveeritud käsitletavast põhiväljendist.

### 2. Verbaalsed ASJAD

Artikli põhiosas käsitlen predikatsioone (kognitiivse grammatika tähenduses), milles:

- 1) verbaalse ASJA "saatmise" protsessi väljendatakse verbiga *andma*;
- 2) verbaalne ASI on kuuldav ja/või nähtav;
- 3) ANDJA ja SAAJA on tavaliselt elus, enamasti inimesed.

### 2.1. sõna andma [kõnejärge loovutama] (näited 4–7 artiklist)

Prototüüpne verbaalne üksus on sõna. Selle tähelepaneku piltlikuks tõestuseks võib tuua eri keelte väljendid *sõna andmise* kohta, milles see sõnaühend tähendab 'kõnevooru, -järge (üle) andma'. Verbiga *andma* väljendatud andmisprotsessis on SAAJA PATSIENT, kelle aktiivsus protsessis on tavaliselt (kuid mitte eranditult) minimaalne. Aktiivse SAAJaga skeemis on perspektiiv andmisele vastupidine, SAAJA on VÕTJA ja AGENT ning protsessi väljendatakse tavaliselt verbiga *võtma*, nt *Järgmine kord ma võtan ise sõna*. Eesti keeles on kasutusel isegi liitsõnad *sõnavõtt* ja *sõnavõtja*.

Verbil *andma* ja selle vastetel teistes keeltes on suhteliselt universaalselt ka kausaalne tähendus (Newman 1996: 172): ANDJA põhjustab selle, et SAAJA saab ASJA. Selles väljendirühmas käsitletud näidetes põhjustab ANDJA (prototüüpses situatsioonis isik, kellel on kontroll toimuva üle) kõnejärje ülemineku SAAJA kätte.

### 2.2. sõna andma [lubadust andma] (näited 9–16)

Sõnaühend *sõna andma* tähendab paljudes keeltes ka lubaduse andmist: *Annan sulle sõna, et artikkel on homseks valmis*. Ka selles skeemis on eesti keeles täidetud vastandpool: võime öelda *Ma võtsin talt sõna, et ta selle tagasi toob*. Üsna tavaline on elliptiline kasutus: *Minu sõna selle peale*.

### 2.3. nõu andma [soovitama, juhendama] (näited 17–26)

*Nõu* sobib *andma* literaalse tähenduse laienduseks seetõttu, et ASI (siin *nõu*) on sellises suhtes tavaliselt entiteet, millest SAAJA eeldatavasti kasu saab. SAAJA ongi prototüüpselt BENEFITSIENT, nt *Euroteaduskonna raames tegutsevad külalislektorid peavad andma nõu ja abistama meie kolleege korraliku õppekava loomisel*. (aja90)

Eesti keeles on verbi *andma* tüve sisaldavate liitsõnadena viimistletud järgmised mõisted: *nõuanne* (ASI), *nõuandja* (ANDJA) ja *nõuandla* (KOHT).

Väljendverbi *nõu andma* on eesti keeles peetud saksamõjuliseks. Samalaadsete väljendite rohkus eri keeltes (vt näited artiklis) osutab, et selline mõistetus on tunnetuse struktuuriga põhjendatud.

Andmise-võtmise skeem selle väljendi puhul täielikult ei tööta: eesti keeles ei öelda *\*nõu võtma*. *Nõu saada* on täiesti võimalik, lisaks on kasutusel väljendid *nõuks võtma* ja *nõu pidama*.

#### 2.4. *käsku andma* [käskima, käsutama], *luba andma* [lubama]

(näited 27–42)

*Andma* argumentiks ASJA rolli sobib ka *käsk*. Selliste mõistestuste puhul on ühtlasi aktiivne nn INSTITUTIONAALSETE SUHETE valdkond: ANDJA on selle suhte seisukohalt olulises hierarhias SAAJAs kõrgemal positsioonil, tal on staatus, mis lubab tal SAAJAt “käsutada”: 8. *augustil andis Norra valitsus sõjaväele kõrgendatud valmisoleku käsu*. ASJA positsiooni sobivad ka paljud *käsuga* sarnaselt toimivad olemid, näiteks *korraldus* (*Haridus- ja kultuuriminister andis korralduse tüdruk kooli tagasi võtta*), *luba* (*Asi ei ole veel lõplik, loa peab andma Euroopa komisjon*), *juhis* jt. ANDJA võib olla ka personifitseeritud entiteet: *Täpsed juhised tegutsemiseks annab jahikorralduskava*.

#### 2.5. *infot andma* [informeerima, teavitama] (näited 43–55)

Verbi *andma* verbaalsuhtlust väljendavate kasutuste ning *andma* litaalse tähenduse kasutuse struktuuris ja osaliste omadustes on palju paralleele, kuid on ka suhteliselt süstemaatilisi erinevusi. Litaalses tähenduses *andma* protsessis siirdub ASI ANDJAlt SAAJAlle, protsessi lõpuks on see tavaliselt SAAJA valduses ja ANDJAl seda ASJA enam pole. Verbaalset suhtlust väljendavates protsessides seda viimast aspekti tavaliselt lähtevaldkonnast kaasatud ei ole: ASI jääb ANDJAlle alles ka peale andmisprotsessi, olles samal ajal ka SAAJA abstraktses valduses, nt *Soovitav on saadud katseandmed vormistada lihtsa tabelina, mis annab küllaldaselt esialgset informatsiooni*. Sellist protsessi võib nimetada ASJA jagamiseks (vrd ka väljendid *teadmisi, infot jagama*).

#### 2.6. *vastust andma* [vastama], [vastutama] (näited 56–60)

Ka *vastus* on verbaalne ASI, mida *andma*-verbiga saab “liigutada”: *Eksamil annate vastuseid kordamööda*.

Väljendil *millegi ees(t) vastust andma* on eesti keeles piltlik kasutus tähenduses ‘vastutama, vastutust kandma’: *Ettevõtjal tuleb vastust anda oma võlausaldajate, riigil valijate ees*. Selle väljendi produktiivne kasutus võib olla tingitud sellestki, et see sarnaneb häälikujult piltliku tähenduse lähima litaalse vastega *vastutama* (ehk ka teise piltliku vastega *vastutust kandma*).



## 2.7. VERBAALSED ASJAD ja sotsiaalsete suhete valdkond

### 2.7.1. *etendust andma* [etendama]

ja *kontserti andma* [kontserdiga esinema] (näited 61–69)

Nende mõistestuste puhul on verbaalse suhtlemise valdkonna kõrval vähemalt võrdselt aktiivne sotsiaalsete suhete valdkond, ASI on sündmus, SAAJA publik: *Teatrikuul toimus 10 esietendust, anti 420 etendust 168000 küllastajale*. ASJA verbaalsus on niisuguste argumentide puhul mõnel juhul väga üldistatud: *etendus* võib ju olla suures osas ainult visuaalsele meelele. Mingi “verbaalsuse eelistus” paistab *andma*-verbil siiski olevat: *balletti*, mida naudime eelkõige siiski vaadates, tavaliselt ei *anta*, vrd \**trupp andis balleti* (kuigi sobib nt *trupp andis balletietenduse*), aga *ansambel andis kontserdi* on tavaline, kuigi literaalses tähenduses protsesside puhul peaks põhiline roll olema just visuaalsel tajul: mingi eseme andmine või saamine on ilmselt eelkõige nähtav.

### 2.7.2. *tundi andma* [õpetama] (näited 70–81)

Selle rühma mõistestustes on ASI ANDJA “valduses olev” teadmine või oskus, mis andmisprotsessis siirdub: *Ta annab nüüd meie vanas koolis matemaatikat*. Sellisest teadmiste ja/või oskuste edasiandmise tähendusest on *andma* kasutus tõenäoliselt laienenud teadmiste ja/või oskuste andmise sündmuse väljendamiseks: *Ta hakkas tunde andma ka teistes koolides*. Sarnaselt eelmise väljendirühma tähendusega on seegi paigutus seotud SÜNDMUSRUUMiga.

Saksa keele mõju eesti keelele ei ole siin võimalik, sest saksa keeles tundi ei *anta*: *eine Stunde halten* (verbiga *halten* ‘pidama’ väljendatakse ka loengu ja kõne siirdamise protsesse: *einen Vortrag halten, eine Rede halten*).

### 2.7.3. *tervitust, õnnitlust üle / edasi andma* [tervitama], [õnnitlema]

(näited 82–86)

Verbiga *andma* on võimalik väljendada ka *tervituse* ja *õnnitluse* vahendamist. Eesti keeles on väljendil *tervitust üle andma* piirang ANDJA omadustes: ANDJA on tavaliselt ASJA vahendaja, mitte autor (vrd ka 2.5.). *Õnnitlust üle andma* võib tähendada nii kellegi õnnitluse vahendamist (*Ma annan sulle üle õnnitluse ka oma ema poolt*) kui autori enda õnnitluse üleandmise hetke (*Ma andsin üle oma südamliku õnnitluse*). Nendes väljendites on eesti keeles kasutusel ühendverbid. Ühendverbi adverbilised komponendid *üle* ja *edasi*, mis põhitähenduselt on kohta ja

liikumist märkivad mõisted, rõhutavad predikatsioonis protsessi kulgemise teed.

#### 2.7.4. *õnnistust andma* [õnnistama] (näited 87–95)

*andma* literaalset tähendust iseloomustab see, et protsessi osalistes toimub protsessi käigus mingi muutus. Piltliku tähenduse puhul võib see asjaolu motiveerida performatiivide kasutamist (*Ta andis tütrele oma õnnistuse*), milles üks osaline oma lausungiga muudab osaliste staatust või seisundit (Newman 1996: 141).

### 3. Järeldused

Meie tajumehhanismis toimivad võrdlusprotsessid. Uut objekti võrreldakse juba kogetu ja süsteemi talletatuga ning “lahterdatakse” sarnasuse või muu sobivuse alusel. Verbi *andma* kasutus verbaalse suhtluse valdkonnas osutab konkreetse eseme ja verbaalse üksuse siirdumise tajumise sarnasust.

Nagu paljudes teistes keeltes, toimub ka eesti keeles *andma*-verbi grammatikaliseerumisprotsess (grammatikaliseerumine on siin laiemas tähenduses, vt Heine *et al* 1991). Üks selle protsessi põhjustest on kindlasti *andma* kausaalsustähendus. Sama oluline on see, et *andma* kuulub nende keele vahendite hulka, millega väljendatakse põhilisi ja igapäevaseid inimestevahelisi suhteid (näiteks informatsiooni andmine).

Tähenduslaienduste motivatsioonis on olulisel kohal põhitähenduse ja laienuvad tähenduse osaliste, nende omaduste, sündmuse struktuuri jms vahelised paralleelid, kuid väga sageli on nende vahel ka suhteliselt süstemaatilisi erinevusi. Näiteks literaalses tähenduses *andma* protsessis siirdub ASI ANDJAlt SAAJAlle ja on protsessi lõpuks SAAJA valduses, kusjuures ANDJAl seda ASJA enam pole, verbaalse suhtluse valdkonnas seda viimast aspekti tavaliselt lähtevaldkonnast kaasatud ei ole: ASI on ANDJA valduses ka peale andmisprotsessi ja on samal ajal ka SAAJA abstraktses valduses (näiteks rühmades *infot andma, nõu andma* jt).

Verbaalse suhtluse valdkonna sagedastes kasutustes toimivad produktiivselt skeemid, mis mõjutavad keekekasutust, näiteks kui *sõna* saab *võtta*, siis saab seda ka *anda* või *saada* jne. Need skeemid töötavad näiteks väljendirühmades *sõna andma* [kõnejärge loovutama] ja *sõna andma* [lubadust andma]. Selline kasutus võib olla seotud sellega, et SAAJAlt oodatakse teatavat (skeemis töötavate reeglite kohast) aktiivsust.

Rühmades *nõu andma* [soovitama, juhendama]; *käsku andma, luba andma, infot andma, vastust andma, õnnistust andma* ei tööta see skeem

aga täiuslikult, nt saab eesti keeles *anda nõu* ja *saada nõu*, kuid mitte *\*võtta nõu* (*nõuks võtta* on võimalik). Neis skeemides ei oodata SAA-JAlti aktiivsust, agendilisust. Mitmesuguseid erinevusi ilmneb ka keelte võrdlemisel.

Küsitav on selline skeem ka rühmas *tervitusi üle/edasi andma*: *tervitused* on võimalik *kätte saada*, kuid neid ei saa *\*võtta*.

Skeem ei tööta üldse sotsiaalsete suhete valdkonda kuuluvas rühmas *etendust, kontserti andma*: nii ei saa näiteks *\*kontserti/etendust võtta* (sündmuse vastuvõtja seisukohalt) ega ka *\*kontserti/etendust saada*.

Samalaadne on olukord ka rühmas *tundi andma*, ent mõningad kivistunud väljendid viitavad võimalusele, et teatud tingimustel ja teatud tähendustes on täidetud skeem siiski võimalik. Nii on eesti keeles võimalik *võtta* või *saada eesti keele tunde*, ent mitte *\*võtta* või *\*saada eesti keele tundi*. (Välja arvatud tähenduses, kus *tundi võtma* kasutatakse tähenduses 'tundi andma'.) Kuna selles väljendis peab ASI olema tingimata mitmuses, võime seda pidada kivistunud väljendiks, seega skeem ei ole produktiivne. Lisaks on võimalik ka: *sain oma õppetunni, andsin õppetunni* (ent mitte: *\*võtsin oma õppetunni*).

Kognitiivses lingvistikas peetakse keskseks küsimust keele ja mõtlemise seostest ja arutletakse selle üle, kas keelelised erinevused võivad põhjustada mõtlemise erinevusi, mis omakorda põhjustaks käitumise erinevusi. Käesoleva artikli materjali taustal võib näiteks küsida, kas eesti õpetaja ja inglise õpetaja käitumine *tundi andes* on teistsugune kui saksa õpetaja oma, kes *tundi peab* (*eine Stunde halten*) või kas ingliskeelse loengupidaja, kes 'loengut annab' (*give lecture*), on andmisprotsessile lähedasem ka reaalses tegevuses, erinevalt eesti- ja saksa-keelsest lektorist, kes *loengut peavad*? Või miks siiski kõik verbaalseid üksusi väljendavad keeleüksused ei sobi *andma* argumendiks (vrd nt eesti keeles *\*kõnet andma* 'kõnet pidama' *\*laulu andma* 'laulma')?

*Andma* tähendus ei ole eesti keeles nii üldistunud, et võiks tingimusteta väita, et see on grammatikaliseerunud, kuid tähendus on piisavalt üldine, et kasutusvõimaluste loetlemise asemel tundub otstarbekas rääkida hoopis piirangutest (kuna neid on tunduvalt vähem). Järelikult on põhjust väita, et olenemata tunnetuse struktuuride universaalsusest on igal keelel siiski midagi omast, mis ehk iseloomustab selle keele emakeelsete kõnelejate maailmapilti, ja seda ka sõnavara tasandil.

# ESTONIAN AND ENGLISH: A LEXICOGRAPHER'S POINT OF VIEW I

Enn Veldi

*University of Tartu*

*Dictionaries are highly complex things which are barely understood by their authors, not understood at all by the publishers, advertising agents, and booksellers and understood least of all by the buyers (Hausmann 1989: 100)*

## 1. Introduction

Estonian–English and English–Estonian lexicography has been a neglected topic in Estonian linguistics. There are of course the existing dictionaries and their reviews, and there is also some material that describes how the lexicographers did it (Silvet 1976; see also Vihma 1999). However, to the best of my knowledge, two of our outstanding lexicographers – Johannes Silvet (1895–1979) and Paul F Saagpakk (1910–1996) – did not carry out any theory-oriented lexicographic research. Actually, one can find in the literature even the statement that “no actual unified theory of bilingual lexicography has ever been published, although certain essential partial aspects have been discussed” (Kromann *et al* 1989: 2715). The lexical structures of English and Estonian have been approached mainly from the point of view of translation studies (Lehtsalu 1964; Lehtsalu, Liiv 1972). The aim of the present article is to open discussion in matters that concern Estonian–English bilingual lexicography. The first part of the article will focus on some problems of the macrostructure (inventory) of general-purpose English–Estonian and Estonian–English dictionaries.

### 1.1. Some typological features of Estonian and English

For bilingual lexicography each language pair is unique and requires an individual approach. Ladislav Zgusta has written that

“anisomorphism can be manifested by any component of the lexical meaning, in any degree and dimension” (Zgusta 1971: 296).

As for its grammatical structure, Estonian is remarkably more synthetic than English and contains a large number of inflected forms – many of its grammatical relations are expressed ‘within the word’ English, on the other hand is highly analytic, and most of its grammar lies ‘outside the word’ Therefore, it is common that English requires two or sometimes even three words to express an idea that can be expressed in Estonian by means of a single word (e.g. *õpetajale* – *to a teacher*, *süüakse* – *is/are (being) eaten*). On the other hand, Estonian has a number of analytic features, too – a characteristic feature being the use of postpositions preceded by nouns in the genitive case (e.g. *minu kõrval* – *next to me*, *maja taga* – *behind the house*).

It is also clear that both Estonian and English reveal a high degree of idiomaticity, which requires careful lexicographic treatment. It is characteristic of the Estonian language to form idiomatic compounds: *hädavares* ‘lame duck’ *naljahammas* – *joker, punster*; *pullivend* – *prankster*, *päevavaras* – *loafer, idler*. The typical Estonian greeting *kuidas käsi käib* – *how is it going* is also highly idiomatic, or think of the expressions *mine metsa* or *mine kuu peale* – *go to hell* (lit. ‘go to the forest’ and ‘go to the moon’). Collocational problems can be illustrated by three examples from athletics: *ketast heitma* – *to throw the discus*, *kuuli tõukama* – *to put the shot*, and *oda viskama* – *to hurl/throw the javelin*.

## 1.2. Sources of lexical similarity between Estonian and English

From the point of view of etymology and culture, Estonian has been subject to extensive ‘Europeanization’ Throughout its history Estonian has absorbed large numbers of words from Germanic languages (as well as Slavic and Baltic languages).

Germanic loanwords (i.e. words borrowed before the beginning of the Low German impact in the 12<sup>th</sup> century), for example, have given rise to the following similarities between Estonian and English: *aer* – *oar*, *jõul* – *Yule(tide)*, *katel* – *kettle*, *kuld* – *gold*, *kuningas* – *king*, *küünal* – *candle*, *laen* – *loan*,

*lammas – lamb, lauk – leek, leib (gen leiva) – loaf, lima – slime, ling – sling, lukk – lock, nael – nail, nõel – needle, paat – boat, põld – field, rand – strand, rooste – rust, rõngas – ring, sadul – saddle, sama – same, sättima – to set, tina – tin, vaal – whale, õlu – ale.*

However, it is the Low German element that is most extensive in Estonian. According to Huno Rätsep, the number of Low German stems in Estonian is between 771 and 850 (Rätsep 1983: 546). It is usually pointed out that the Estonian vocabulary that describes medieval city culture is mostly of Low German origin. What has not been pointed out, however, is that this factor gives rise to numerous lexical similarities between Estonian and English as well. The pedagogical implications of these similarities for the teaching of English have never been realized in Estonia – on the one hand, as a mnemonic aid and, on the other, as a source of possible *faux amis*. To the best of my knowledge, no comprehensive list has ever been attempted from the point of view of English. Below are some examples, where the similarities between Estonian and English have arisen because the Estonian word was borrowed from Low German (for Low German loanwords in Estonian see Liin 1968 and Ariste 1981).

*ankur – anchor, haak – hook, haamer – hammer, hakkima – to hack, hinged – hinges, iisop – hyssop, kaal – (Am) scale, (Br) scales; kamber – chamber, kamm – comb, kelder – cellar, kokk – cook, kook – cake; (Am) cookie; koolmeister – schoolmaster, korter – quarter, kraapima – to scrape, kratsima – to scratch, kruvi – screw, käärid – shears, kühvel – shovel, lamp – lamp, liilia – lily, mölder – miller, münt – mint, naaber – neighbour, näärid – new year, pakkima – to pack, pink – bench, pannkook – pancake, plaat – plate, pleekima – to bleach, pool (gen pooli) – spool, pruulima – to brew, pump – pump, raam – frame, redel – ladder, reede – Friday, rent – rent, ruum – room, röövel – rover; robber; salpeeter – saltpeter, seep – soap, taaler – dollar, tangid – tongs, tool – stool, torm – storm, tosin – dozen, triip – stripe, tärklis – starch, tüürima – to steer, vaagima – to weigh, viil – file, villima – to fill.*

A further contributing factor to such similarities is that English has absorbed large numbers of Low Dutch loanwords. The

number of these words in English may amount to 2,500 (Baugh and Cable 1978: 188).

In addition, there are of course numerous English loanwords in Estonian. Aino Jõgi's dissertation discusses the impact of English on Estonian up to the beginning of the 1970s (Jõgi 1971). However, the increased impact of English on Estonian in recent years has yet to be established. It is clear that many new anglicisms have been added, e.g. *AIDS – AIDS*, *miksima – to mix*, *reivima – to rave*, *ruulima – to rule*, *šoppama – to shop* (for an interesting study of anglicisms in various European languages see Görlach 1994 and 1998).

To sum up, considering the large variety of European influences on Estonian (some of which were not discussed above, e.g. international terminology), English and Estonian share more similarities than most people are aware of.

## **2. Typology of dictionaries**

Henri Béjoint (1994: 32) points out that, in principle, it is possible to make a distinction between a classification of dictionaries and a typology of dictionaries. Librarians classify dictionaries, and linguists work out typologies. From the librarian's point of view, dictionaries can be classified into a large number of categories. There are general-purpose dictionaries; monolingual, bilingual, and multilingual dictionaries; dictionaries of synonyms and antonyms (thesauruses), pictorial and photo dictionaries, terminological dictionaries, dialect dictionaries, slang dictionaries, dictionaries of false friends, dictionaries for native speakers and foreign learners, bilingualized learners' dictionaries, dictionaries for children, students and adults, phrase books for tourists, etc.

Linguists have tried to apply more theoretical approaches. Yakov Malkiel worked out perhaps the best-known linguistic typology of dictionaries (Malkiel 1962, for the discussion see also Béjoint 1994: 33–34). Malkiel's typology is based on three criteria: 1) range, 2) perspective, and 3) presentation. Range is a criterion that characterizes the number of entries, the number of languages, and the percentage of linguistic information (i.e. the ratio of purely linguistic to encyclopedic information). It is clear that 'range' (content) is a problematic feature because the density

of a dictionary can be measured in two ways: as 1) density of the macrostructure and 2) density of the microstructure. Range is also the criterion that measures the types of lexical items included, such as words and idioms, archaic words, dialect words, and slang words, etc.

The second criterion – ‘perspective’ – divides dictionaries into diachronic and synchronic ones. One has to admit that this criterion is somewhat problematic because most dictionaries of contemporary language are not purely synchronic and tend to have a mixed character. In addition to the current vocabulary, they usually include also a proportion of archaic and outdated words. Perspective also shows how the material is arranged in a dictionary – it can be alphabetic (or reverse alphabetic) or conceptual. Malkiel’s third criterion – ‘presentation’ – measures the precision of metalanguage (e.g. the language of definitions), it assesses the nature of examples (citations, invented examples), the presence of illustrations as well as such features as phonetic description and usage labels. Malkiel also makes a distinction between descriptive and prescriptive dictionaries, which is important for Estonian-language lexicography. Erelt *et al* (1997: 569–574) classified Estonian-language dictionaries according to 1) the headword list and 2) description. The macrostructure reflects the type of vocabulary (e.g. general language, dialects, slang, terminology, neologisms, abbreviations, etc.) that is included as well as its arrangement. The latter can be alphabetic or non-alphabetic (i.e. notional, according to frequency, etc.). The description may be provided in the same language or in another language. It may include explanations, grammatical information, lexical relations with other words, etymology, word frequencies, etc. The information provided may have a descriptive or prescriptive character. According to Erelt *et al* the type of a dictionary can be established by combining these features.

After discussing the existing typologies Béjoint is of the opinion that “it is impossible to classify dictionaries in a way that would be both orderly and applicable to all societies. Dictionaries come in more varieties than can ever be classified in a simple taxonomy” (Béjoint 1994: 37). I believe that he is right. One could point out, for example, that the 1990s gave rise to a new criterion for classification – ‘format’. If we use ‘format’ as a



criterion, we can speak of 1) print dictionaries, 2) CD-ROM dictionaries, and 3) web-based dictionaries.

The Russian linguist Lev V Shcherba, whose work was until recently relatively unknown in the West, made a useful theoretical innovation that has greatly affected the recent discussion of bilingual dictionaries (see Shcherba 1995). His typology of dictionaries is based on the fact that dictionaries are needed for two basic functions – decoding (comprehension) and encoding (production, i.e. writing or speaking). Dictionaries for decoding are passive dictionaries and dictionaries for encoding are active ones. Thus, four dictionaries are needed for each language pair:

- 1) L2→L1 (for decoding)
- 2) L2→L1 (for encoding)
- 3) L1→L2 (for encoding)
- 4) L1→L2 (for decoding)

If L1 stands for Estonian and L2 is English, then dictionaries of types 1) L2→L1 (for decoding) and 3) L1→L2 (for encoding) are needed for Estonians and types 2) L2→L1 (for encoding) and 4) L1→L2 (for decoding) for non-Estonians. Besides, dictionaries can be monodirectional or bidirectional. A bidirectional dictionary includes two functions, for example, a bidirectional Estonian–English dictionary can be used both by Estonians and non-Estonians.

### **2.1. Estonian–English and English–Estonian dictionaries**

The list of general-purpose Estonian–English and English–Estonian dictionaries is not long. We will focus our attention on three general-purpose dictionaries: the Estonian–English dictionaries by Paul Saagpakk (1992) and Johannes Silvet (1980) and the two-volume *English–Estonian Dictionary* by Johannes Silvet (1989). Saagpakk's dictionary is a bidirectional dictionary. It was intended as a decoding dictionary for the English-speaking world and as an encoding dictionary for native speakers of Estonian. Silvet's dictionaries were compiled for Estonians – the *English–Estonian Dictionary* is for decoding English texts and the *Estonian–English Dictionary* is for producing texts in English. The *English–Estonian Dictionary of Idioms* by Urve Hanko and Gustav Liiv (Hanko and Liiv 1998) is the only comprehensive

dictionary of its kind; it is a monodirectional dictionary for comprehension. There is no Estonian–English dictionary of idioms as yet, and bilingual English–Estonian and Estonian–English synonym dictionaries are non-existent. Specialized dictionaries will remain beyond the scope of the present article.

### 3. How many words are there in Estonian and English?

This problem is of great interest for Estonian lexicographers who work with English. Is Estonian an equal partner to English, which is known as a language with a huge vocabulary? When reading the figures about the richness of English vocabulary, one is tempted to assume that English must have many more words than Estonian. How does that affect dictionary making? Although it is impossible to count the possible size of English vocabulary, it is also quite certain that English has over one million words (McArthur 1998: 641). There are also bolder estimates that put the figure at four million words. It is of course true there is no dictionary that would include all these words. Nevertheless, some impressive unabridged dictionaries have been published (e.g. the second edition of the *Oxford English Dictionary* (1989) contains over 500,000 words and 2.5 million quotations).

How about the Estonian language? The entry *sõnavara* ‘vocabulary’ in the Estonian Encyclopedia informs us that there are at least 500,000 lexical items in Estonian (EE 9: 65). Huno Rätsep has estimated that the number of words that may have been used in written Estonian could well reach 600,000–700,000 items (Rätsep 1976: 144). Although these figures are rather impressive, the number of Estonian words that have been registered in a monolingual dictionary of Estonian, and thus immediately accessible for a lexicographer, is relatively small. The hitherto published fascicles of the *Explanatory Dictionary of Written Estonian* (henceforth, EKSS) include about 118,904 entries (Margit Langemets and Heiki-Jaan Kaalep, personal communication). The letter S has come out already, and it is unlikely that the total figure will reach well over 150,000 items.

It is also interesting to note that in the English-speaking world dictionaries are divided into unabridged (over 250,000 words), semi-unabridged (130,000–250,000 words) and abridged

dictionaries (less than 130,000 words) (see Béjoint 1994: 32–33). I'm afraid that the same scale cannot be applied to Estonian, where the upper limit seems to be 150,000 at present.

It is a well-known fact that the Estonian language is in need of constant modernization, especially in the area of specialized terminology. Actually, good progress has been made. According to Eevi Ross (Ross 1997: 453), the period between 1961 and 1990 saw the publication of over 120 dictionaries that contained about 350,000 terms. Most of these dictionaries included Russian as one of the languages. At least 74 specialized dictionaries were published in 1991–1995 (Erelt 1997: 306).

Unfortunately, the terminology of various fields is scattered in a large number of specialized dictionaries. Therefore, the bilingual lexicographer is faced with two challenges when looking for translation equivalents for a large English–Estonian dictionary. First, there can be thousands of instances, where no Estonian-language equivalent is available, and, second, the lexicographer cannot fully rely on the existing monolingual dictionaries. He or she has to look for translation equivalents in numerous specialized dictionaries and primary sources. This circumstance makes progress rather slow.

In order to find out whether my ideas have any foundation, two 100-item-runs of entries were culled from English and Estonian general-purpose monolingual dictionaries from letter J. The entry-lists were then compared with the runs of entries in the existing general-purpose English–Estonian and Estonian–English dictionaries.

#### **4.1. The macrostructures of NODE, COD 10, and SILVET 3**

A 100-item run of entries was culled from the beginning of the letter J of the *New Oxford Dictionary of English* (1998) (henceforth, NODE). NODE is a new single-volume monolingual dictionary of English that contains about 350,000 words. This sequence was then compared to the entry-list of the tenth edition of the *Concise Oxford Dictionary* (1999) (henceforth, COD 10). COD 10 is based on NODE, and its scope is 240,000 words. The entries were counted in the following way – included were headwords, different parts of speech (e.g. nouns, verbs, adjectives,

etc.), derivatives, and phrases. Excluded were different senses and sub-senses, as well as compounds that were listed by their second components. It appeared that NODE included about ten percent of proper names, which were not provided in COD 10. Therefore I decided to delete the proper names from the entry-list. The choice of the *Concise Oxford Dictionary* was not accidental. Namely, Johannes Silvet pointed out in the preface to the third edition of his dictionary that *COD* had served as a model for his dictionary in a number of ways. He mentions such features as spelling, grammatical classification and terminology (Silvet 1989: 5). It is interesting to note that *COD* had served as a model also for the famous *Oxford Advanced Learners' Dictionary* by A.S. Hornby (Cowie 1999: 47).

I then compared the entry-list to the macrostructure of the third edition of the two-volume *English–Estonian Dictionary* by Johannes Silvet (1989) (henceforth, SILVET 3). When SILVET 3 was published ten years ago, it was in a reviewer's opinion outdated by the time of its publication (Rähesoo 1990). Therefore, it was also of great interest to see whether the gaps that would be revealed could be dealt with successfully or not.

NODE	COD 10	SILVET 3
ja	–	–
ja well	–	–
ja well no fine	–	–
jab (verb)	+	+
jab (noun)	+	+
jabber (verb)	+	+
jabber (noun)	+	+
Jabberwocky	+	–
jabiru	+	–
jab jab	+	–
jaborandi	+	–
jabot	+	+
jacal	+	–
jacamar	+	–
jacana	+	–

<b>NODE</b>	<b>COD 10</b>	<b>SILVET 3</b>
jacaranda	+	-
jacinth	+	+
jack(1) (noun)	+	+
before one can say Jack Robinson	-	+
every man jack	+	+
I'm all right, Jack	-	-
jack of all trades (and master of none)	+	+
on one's jack (or Jack Jones)	-	-
jack someone around	+	-
jack in (or into)	+	-
jack something in	+	-
jack off	+	-
jack up	+	-
jack something up	+	+
jack (2) (noun)	+	-
jack (3) (adjective)	+	-
jackal	+	+
jackanapes	+	+
Jack and Jill party	+	-
jack arch	-	-
jackaroo (noun)	+	+
jackaroo (verb)	+	-
jackass	+	+
jack bean	-	-
jackboot	+	+
jackbooted	+	-
Jack-by-the-hedge	+	-
jack chain	-	-
Jack cheese	-	-
jackdaw	+	+
jackeen	+	-
jackeroo	-	-
jacket (noun)	+	+
jacket (verb)	+	+

NODE	COD 10	SILVET 3
jacket potato	+	+
jackfish	+	-
Jack Frost	+	+
jackfruit	+	-
jackhammer (noun)	+	-
jackhammer (verb)	+	-
jackie hangman	-	-
jack-in-office	-	+
jack-in-the-box	+	+
Jack-in-the-pulpit	+	+
jackknife (noun)	+	+
jackknife (verb)	+	-
jackknife clam	+	-
jackknife fish	+	-
Jackleg	-	-
jack light	-	-
jack mackerel	+	-
jack-o'-lantern	+	+
jack pine	+	-
jack plane	+	+
jack plug	+	-
jackpot	+	+
hit the jackpot	+	+
jackrabbit	+	+
Jack Russell (terrier)	+	-
jack screw	+	+
jack shaft	-	-
jackshit	-	-
jacksie	+	-
jack snipe	-	+
jack socket	-	-
Jacksonian	-	-
jackstaff	+	+
jackstay	-	-

NODE	COD 10	SILVET 3
jackstone	+	-
jackstraw	+	+
jacksy	+	-
Jack Tar	+	+
Jack the Lad	+	-
jack-up (rig)	-	-
Jacky	-	-
Jacky lizard	-	-
Jacky Winter	-	-
Jacobean (adjective)	+	+
Jacobean (noun)	+	-
Jacobethan	-	-
Jacobian (adjective)	-	-
Jacobian (noun)	-	-
Jacobin	+	+
Jacobinic	+	-
Jacobinical	+	-

Of the counted entries COD 10 had only one (*-jacketed*), which is not in NODE. It is interesting to note that SILVET 3 included 10 items, which were not in NODE and COD 10: *Jack and Gill* (or *Jill*), *Jack Johnson*, *Jack Ketch*, *Jack Pudding*, *Jack Straw*, *Jack-a-dandy*, *to dust smb.'s jacket*, *jack-ladder*, *jack-towel*, and *jacobin cap*.

The count on the basis of this 100-item list reveals that COD 10 includes about 72% of the entries to be found in NODE. This percentage is rather close to 69% (240,000 words in COD 10 versus 350,000 in NODE), taking into account that roughly 10% entries in NODE are proper nouns. SILVET 3 includes 37% of the entries to be found in NODE and 51% of the entries to be found in COD 10. It would be perhaps too much to modernize SILVET 3 by 40–50%. However, a 30% modernization programme would be perhaps desirable.

Does the Estonian language have the required translation equivalents or not? I examined some of the existing gaps with the purpose of finding suitable translation equivalents. At this I tried

to find some empirical evidence and indicate the source where the equivalent was found. However, in most cases the primary source was not a monolingual dictionary but an encyclopedia or some other reference book. It appeared that it is a time-consuming job to establish suitable equivalents from primary sources. In a number of cases I failed to come up with suitable equivalents. Below are the results:

ja (Safr) jah

**Jabberwocky** (coined by Lewis Carroll, not found)

**jabiru** (*Jabiru mycteria*) *zool* jabiiru (Kesk- ja Lõuna-Ameerika toonekurglane) (LE6: 76)

**jab jab** (West Indian) kurat (er karnevalimaskeraadil)

**jaborandi** (not found, suggestion: 1) jaborandilehtedest valmistatud droog, 2) (*Pilocarpus jaborandi*) *bot* jaborandi

**jacal** (not found, suggestion: rookatusEGA plonnhütt)

**jacamar** (*Galbulidae*) *zool* läiklindlane (Lõuna-Ameerika metsalindude sugukond) (EE6: 10)

**jacana, jaçana, lily-trotter** (*Jacanidae*) *zool* lootoselindlane, jassaanalane (EE13: 142–143)

**jacaranda** (*Jacaranda*) *bot* jakaranda (ENE4: 58)

**jack (3)** (Aus/NZ) tülpinud, tüdinud

**jack off** (vulgar slang) pihku taguma

**Jack and Jill party** (Nam) (not found, suggestion: eelpulm)

**jack arch** (not found)

**jack bean** (*Canavalia ensiformis*) *bot* mõõkjäs kanavaalia (ETA)

**Jack-by-the-hedge, hedge garlic** (*Alliaria petiolata*) (*bot* not found)

**jackchain** (not found)

**Jack cheese** (Monterey Jack) (not found, suggestion: Ameerika juustusort (< Monterey maakond Californias)

**jackeen** (Irish, chiefly derog) vurle, linnauntsakas; pätt er dublinlase kohta

**jackeroo** (var of jackaroo)

**jackfish** (chiefly North Am) *zool* havi, purikas, haug; havinolk

**jackfruit** (tropical Asian tree) (*Artocarpus heterophyllus*) *bot* erilehine leivapuu (EE5: 478); erilehise leivapuu vili

**jackhammer** (chiefly NAmer) pneumaatiline vasar, pneumo-vasar, suruõhuvasar (ÕS 1999)



- jackie hangman** (SAfr) (*Lanius collaris*) zool punaselg-õgija (EE10: 586)
- jackknife clam, razor clam** (Am) (*Solenidae*) zool kaunkarplane (LE2: 109)
- jackknife fish** (*Equetus lanceolatus*) (zool not found)
- jackleg** (US informal) soss-sepp, oskamatu töömees; suli
- jack light** (N.Amer) latern õiseks kalapüügiks, tulusepüügilatern
- jack mackerel** (*Trachurus symmetricus*) (zool not found)
- jack pine** (*Pinus banksiana*) bot hall mänd (ETA)
- jack plug** ühe tikuga pistik
- Jack Russell terrier** (zool not found)
- jack shaft** (not found)
- jackshit** (US vulgar slang) kama kõik, sitta kah, kama kaks, üks ta kama kõik
- jacksie** (Brit. informal) pepu, tagument, taguots
- jack socket** pistikupesa
- Jacksonian** (med not found)
- jackstay mer jaktaak** (Olev 1981)
- Jack the Lad** (informal) ennast täis tüüp, ennast täis sell; häbematu nolk
- jack-up rig** (not found, suggestion: avamere naftapuurtorn, mis toetub merepõhjale)
- Jacky** (Aus, offensive) (*halvust*) pärismaalane, aborigeen
- Jacky lizard** (*Amphibolus muricatus*) (zool not found)
- Jacky Winter, Australian brown flycatcher** (*Microeca leucophaea*) (zool not found)
- Jacobethan** (Jacobean + Elizabethan) (not found)
- Jacobian mat** jakobiaan (determinant) (ENE4: 60)

#### 4.2. The macrostructure

##### of EKSS and ÖS 1999 vs. SAAGPAKK and SILVET

It was also of great interest to compare a 100-entry sample (also from the letter J) in two monolingual dictionaries with the macrostructure of two Estonian–English dictionaries. The *Explanatory Dictionary of Written Estonian* (EKSS) is the first monolingual purely descriptive dictionary of Estonian. It has been published in fascicles since 1988, and by the time of writing this article its publication has reached the entry *sulforühm*. The brand-

new *Dictionary of Estonian* (henceforth, ÕS 1999) is a standard-descriptive dictionary (the term comes from Zgusta 1989: 70) with an emphasis on vocabulary modernization. In addition to providing a lexicographic description of contemporary Estonian (actual usage), it also lists coinages that were invented by the authors. The new coinages, however, lack any empirical foundation. Therefore, the authors assume that the native speakers of Estonian are willing to learn and use these words. The prescriptive nature of ÕS 1999 is also manifested in the inclusion of 'negative linguistic material' – it includes 1,180 'undesirable' words in braces that are to be avoided (Margit Langemets, personal communication). They are followed by recommended forms.

The *Estonian–English Dictionary* (1982, 1992) by Paul Saagpakk (henceforth, SAAGPAKK) is the largest Estonian–English Dictionary. Its scope is 150,000 words (Vihma 1999: 71). This fact is of great interest because, as was pointed out earlier, the hitherto published fascicles of the EKSS contain about 118,904 lexical items. Zgusta has written that the number of entries in a bilingual dictionary is “nearly always smaller than that in a comparable monolingual dictionary” (Zgusta 1971: 304). In the case of SAAGPAKK this is not the case. Saagpakk's dictionary is an attempt to provide a complete list of Estonian words together with their English equivalents. It is a bidirectional dictionary that serves a double purpose – as an encoding dictionary for speakers of Estonian and as a decoding dictionary for speakers of English.

It is also of great interest to compare the macrostructure of SAAGPAKK with the *Estonian–English Dictionary* by Johannes Silvet (henceforth, SILVET). The latter dictionary was intended for speakers of Estonian. Silvet's dictionary contains about 30,000 words (Raun 1989: 2394).

<b>EKSS</b>	<b>ÕS 1999</b>	<b>SAAGPAKK</b>	<b>SILVET</b>
jaaguar	+	+	+
jaagup	+	-	-
jaagupipäev	+	+	-
ja-ah	-	-	-
jaah	+	-	-
jaa-jaa	+	-	-
jaa-jah	-	-	-
jaala	+	+	+
jaam	+	+	+
jaamaesine	+	+	+
jaamahoone	-	+	+
jaamakomandant	-	-	-
jaamakorraldaja	+	+	+
jaamapuhvet	-	-	-
jaamarestoran	-	-	-
jaamatöötaja	-	-	-
jaamavahe	+	+	+
jaamaülem	+	+	+
jaan	+	+	-
jaanalind	+	+	+
jaanalinnulik	-	-	-
jaanalinumuna	+	-	-
jaanalinnupoliitika	+	+	+
jaanalinnusulg	+	+	-
jaanihein	-	-	-
jaanijuur (van)	+	+	-
jaanik	+	+	+
jaanikakar	+	+	+
jaanikann	+	+	-
jaanikask	+	-	-
jaanikaun	+	+	+
jaanikesa	-	-	-
jaanikomme	-	-	-

EKSS	ÕS 1999	SAAGPAKK	SILVET
jaanikuine	-	-	-
jaanikuline	+	-	-
jaanikuu	+	+	+
jaanikõder	+	-	-
jaanilaat	-	-	-
jaanilaul	+	-	-
jaanilaupäev	+	+	-
jaanileib	-	+	+
jaanileivapuu	+	+	-
jaanilill	+	+	+
jaanimardikas	+	+	+
jaanipuna	-	+	+
jaanipäev	+	+	+
jaanirukis	-	+	-
jaanituleline	+	+	-
jaanituli	+	+	+
jaaniuss	+	+	+
jaaniõhtu	+	+	+
jaaniöö	+	+	-
jaanuar	+	+	+
jaanuarikuu	+	+	+
jaanuaripakane	+	-	-
jaapani	+	+	+
jaapanikeelne	+	-	-
jaapanlane	+	+	+
jaapanlanna	+	+	-
jaapanlik	-	-	-
jaarmark	+	+	-
jaa-sõna	+	+	+
jaatama	+	+	+
jaatav	+	+	+
jaatavalt	+	+	-
jaatus	+	+	+
jaatussõna (lgv)	-	+	-

<b>EKSS</b>	<b>ÕS 1999</b>	<b>SAAGPAKK</b>	<b>SILVET</b>
jaava	–	+	–
jaavalane	–	+	–
jabur	+	+	+
jaburdama	+	+	+
jaburduma	+	–	–
jaburdus	+	+	–
jaburus	+	+	+
jada	+	+	+
jadalülitus (el)	+	+	–
jadamisi	+	+	–
jadeiit (geol)	+	+	–
jae	+	+	+
jaehind	+	+	+
jaekaubandus	+	+	+
jaekaubandusettevõte	–	–	–
jaekaubandusvõrk	–	–	–
jaekäive	–	–	–
jaemüük	+	+	+
jaevõrk	–	–	–
jagaja	+	+	+
jagama	+	+	+
jagamatu	+	–	+
jagamatult	+	–	–
jagamine	+	+	+
jagamismärk	+	+	+
jagatav	+	+	+
jagatis	+	+	+
jagelema	+	+	+
jagelus	+	+	+
jagin	+	+	–
jagu	+	+	+
jaguma	+	+	+
jagumatu	+	+	+

The word-list of EKSS calls for some comments. As a purely descriptive dictionary its macrostructure is unrestricted. Thus it attempts to include everything that has been recorded. This factor affected to some extent the composition of the initial word-list and the subsequent result of counting. First, EKSS provides various spellings of the same word as separate entries (*ja-ah, jaah*). Second, the list provides as separate entries *jaanikaun, jaanikõder, jaanileib* 'carob bean, locust bean', whereas according to ÕS 1999 the preferred term is *jaanikaun*. Another difference from ÕS 1999 is that EKSS claims that the preferred botanical term is *jaanileivapuu*, whereas ÕS 1999 and the *Estonian Encyclopedia* prefer the term *jaanikaunapuu* 'carob tree, locust tree' (ÕS 1999; EE 4: 33). The latter term is not included in EKSS. Third, EKSS includes some multi-component compounds (*jaekaubandusettevõte, jaekaubandusvõrk*) that can be regarded as unnecessary inclusions. Fourth, the words *jagamatus* 'indivisibility' and *jaanipidu* 'Midsummer Eve bonfire' are missing (EKSS records the latter compound only by its second component in the entry *pidu*).

The comparison with ÕS 1999 revealed that EKSS had 24 words that were missing in ÕS 1999. It should also be pointed out that ÕS 1999 lists *jaamahoone* under *hoone*. In contrast, ÕS 1999 includes 13 words that were not recorded in EKSS. The additions in ÕS 1999 were as follows:

- |                      |                           |
|----------------------|---------------------------|
| 1) jaamatagune       | 8) <i>jackpot</i>         |
| 2) jaanikaunapuu BOT | 9) <i>jadaõpe</i> PED     |
| 3) jaanipidu         | 10) <i>jaehinnakiri</i>   |
| 4) jaanuarilaps      | 11) <i>jaekaup</i>        |
| 5) jaanuarinumber    | 12) <i>jagamatus</i>      |
| 6) jaapanipärane     | 13) <i>jagamispea</i> TEH |
| 7) jaatuslause KEEL  |                           |

In my opinion, most additions are justified; the only compounds that reveal a low degree of 'cruciality' are *jaanuarilaps* 'January child' and *jaanuarinumber* 'January issue'

However, the macrostructure of SAAGPAKK revealed very interesting differences both from EKSS and ÕS 1999. While 30 words that are included in EKSS were missing, SAAGPAKK records 46 words that can not be found in EKSS. Six of them are

proper nouns, and they will not be taken into consideration in the overall count:

- |                           |                          |
|---------------------------|--------------------------|
| 1) jaagerpuri             | 24) jaanuari             |
| 2) jaa-hääl               | 25) Jaapan (pn)          |
| 3) Jaak (pn)              | 26) jaapanitar           |
| 4) Jaakobus (pn)          | 27) jaarl                |
| 5) jaamaline              | 28) jaatamine            |
| 6) jaamaooteruum          | 29) Jaava (pn)           |
| 7) jaamapakihoiuruum      | 30) Jaavet (pn)          |
| 8) jaamapiletikassa       | 31) jaa-ütleja           |
| 9) jaamapiirkond          | 32) jabin                |
| 10) jaamaplats            | 33) jadakäik             |
| 11) jaamaraamatukapp      | 34) jadaühendus (elektr) |
| 12) jaanalinnufarm        | 35) jaekaupmees          |
| 13) jaanalinnukasvandus   | 36) jagamata             |
| 14) Jaani (kihelkond, pn) | 37) jagamisplaan         |
| 15) jaanileet             | 38) jagamissirkel        |
| 16) Jaanilinn (pn)        | 39) jagamus              |
| 17) jaanipidu             | 40) Jagannatha           |
| 18) jaanipäevaaegne       | 41) jaga(ta)matu         |
| 19) jaanipäevane          | 42) jaga(ta)mattus*      |
| 20) jaanirohi             | 43) jaga(ta)matus        |
| 21) jaanitee              | 44) jagatavus            |
| 22) jaanivagel            | 45) jagav                |
| 23) jaaniöine             | 46) jagelemine           |

Is the inclusion of these 40 words justified? My impression is these words have relatively little value for contemporary Estonian. Perhaps they could be helpful in the comprehension pre-war (i.e. before World War 2) Estonian texts. It would be an interesting idea to check their empirical foundation. The list is rather heterogeneous. One group of words includes multi-component compounds: *jaamaooteruum*, *jaamapakihoiuruum*, *jaamapiletikassa*, *jaamaraamatukapp*, *jaanalinnukasvandus*. Then there are words that I perceive unnatural or outdated as a native speaker of Estonian born in 1955: *jaa-hääl*, *jaamaline*, *jaamaplats*, *jaamapiirkond*, *jaapanitar*. The word *jaga(ta)mattus\** is provided with an asterisk, which means that it must be a neologism. Judging by this 100-item sample, SAAGPAKK is overburdened with

unnatural words that are not used in contemporary Estonian. Although it contains 150,000 words, almost one-third of its entries could be of relatively little value to the present user.

In contrast, the *Estonian-English Dictionary* (1920) by Johannes Silvet has a well-balanced macrostructure. It included 50% of the word-list of EKSS, but it did not include any unnecessary words either. Its macrostructure is firmly based on monolingual dictionaries.

## 5. Concluding remarks

The present study was limited to two 100-item runs of entries. Therefore, the results obtained enable us to make only rough estimates. It is clear that English-Estonian dictionaries are in need of modernization. However, there are two problems that make the task more complicated. Where translation equivalents are missing, new words need to be coined. This is a job that a lexicographer should generally avoid. The lexicographer's primary task is to deal with the existing words and meanings and not to coin words extensively. Therefore, cooperation between bilingual lexicographers and Estonian philologists is needed. The second problem is that a term may exist in Estonian, but it is not easily available to the lexicographer (i.e. the existing monolingual dictionaries are inadequate). Therefore, the lexicographer has to put in considerable efforts to pin down the existing term in primary sources, which is a time-consuming job.

For this reason, at present it is easier to compile Estonian-English dictionaries. In the light of the recent changes they need considerable modernization, too. As the monolingual descriptions of English are more adequate, the task is easier. Here the challenge is to find suitable equivalents for roughly 50,000 Estonian words that have never before been provided with equivalents in an Estonian-English dictionary.



## REFERENCES

- Ariste, Paul 1981:** Keelekontaktid. Tallinn.
- Baugh, Albert C., Thomas Cable 1978:** A history of the English language. Third edition. London.
- Béjoint, Henri 1994:** Tradition and innovation in Modern English dictionaries. Oxford.
- COD 10 = The Concise Oxford Dictionary.** Tenth edition. Edited by Judy Pearsall. Oxford UP 1999.
- Cowie, A. P. 1999:** English Dictionaries for Foreign Learners: A History. Oxford.
- EE = Eesti entsüklopeedia, vol. 5–10.** Tallinn. 1990–1998.
- EKSS = Eesti kirjakeele seletussõnaraamat: A–sulforühm.** Tallinn. 1988–1999.
- ENE = Eesti nõukogude entsüklopeedia, vol. 1–4.** Tallinn. 1985–1989.
- EPE = Eesti põllumajandusentsüklopeedia, vol. 1.** Tallinn, 1998.
- Erelt, Mati, Tiiu Erelt, Kristiina Ross 1997:** Eesti keele käsiraamat. Tallinn.
- Erelt, Tiiu 1997:** Eesti oskussõnastikud 1991–1995 I. – Keel ja Kirjandus, no. 5. 305–314.
- ETA = Eestikeelsete taimenimetuste andmebaas:** [www.ut.ee/BGBA/ld](http://www.ut.ee/BGBA/ld).
- Görlach, Manfred 1994:** A Usage Dictionary of Anglicisms in Selected European Languages. – International Journal of Lexicography, vol. 7, no. 3. 223–246.
- Görlach, Manfred 1998:** The Usage Dictionary of Anglicisms in Selected European Languages: a report on progress, problems and prospects. – Links & Letters 5. 209–221.
- Hanko, Urve, Gustav Liiv 1998:** English–Estonian Dictionary of Idioms. Tallinn.
- Hausmann, Franz Josef 1989:** Dictionary criminality. – Wörterbücher. Dictionaries. Dictionnaires. An International Encyclopedia of Lexicography. NY, Berlin. 97–101.
- Jõgi, Aino 1971:** Inglise päritolu sõnad eesti keeles. Tartu. (Diss. kand. fil. MS).
- Kromann, Hans-Peder et al 1989:** Principles of Bilingual Dictionaries. – Wörterbücher. Dictionaries. Dictionnaires. An International Encyclopedia of Lexicography. NY Berlin. 2711–2728.
- LE = Loomade elu, vol. 1–7.** Tallinn. 1981–1987.
- Lehtsalu, Urve 1964:** Ilukirjanduse inglise keelest eesti keelde tõlkimise probleeme (J. Galsworthy teoste eestinduste põhjal). Tartu. (Diss. kand. fil. MS).

- Lehtsalu, Urve, Gustav Liiv 1972:** Ilukirjanduse tõlkimisest inglise keelest eesti keelde. Tartu.
- Liin, Helgi 1968:** Alamsaksa laensõnadest 16. ja 17 sajandi eesti kirjakeeles. Emakeele Seltsi aastaraamat 1967. Tallinn. 47–69.
- Malkiel, Yakov 1962:** A Typological Classification of Dictionaries on the Basis of Distinctive Features. – Fred W Householder and Sol Saporta (eds.), Problems in Lexicography. Bloomington. 3–24.
- McArthur, Tom 1998:** Concise Oxford companion to the English language. Oxford.
- NODE = The New Oxford Dictionary of English.** Edited by Judy Pearsall. Oxford UP 1998.
- Olev, Kulno 1981:** English–Estonian–Russian Maritime Dictionary, Tallinn.
- Raun, Alo 1989:** Lexicography of the Uralic Languages IV: Estonian and Livonian. – Wörterbücher, Dictionaries, Dictionnaires. An International Encyclopedia of Lexicography. NY Berlin. 2392–2935.
- Ross, Eevi 1997:** Eesti leksikograafia XX sajandi teisel poolel: arengujooni ja kokkuvõtteid. – Keel ja Kirjandus, no. 7 449–458.
- Rähesoo, Jaak 1990:** Arvustuse asemel. – Keel ja Kirjandus, no. 6. 374–375.
- Rätsep, Huno 1976:** Kui palju on eesti keeles sõnu? – Keel, mida me uurime. Tallinn. 141–144.
- Rätsep, Huno 1983:** Eesti kirjakeele tüvevara päritolu. – Keel ja Kirjandus, no. 10. 539–548.
- Saagpakk, Paul 1992:** Estonian–English Dictionary. Tallinn.
- Shcherba, L.V. 1995:** Towards a General Theory of Lexicography. – International Journal of Lexicography, vol. 8, no. 1. 314–350.
- Silvet, Johannes 1976:** Kuidas valmis inglise–eesti sõnaraamat. – Keel, mida me uurime. Tallinn. 1976. 145–149.
- Silvet, Johannes 1980:** Estonian–English Dictionary. Second edition. Tallinn.
- Silvet, Johannes 1989:** An English–Estonian Dictionary. Third edition. Tallinn.
- Vihma, Helgi 1999:** Aaviku Seltsi esimene auliige Paul Saagpakk. – Artikleid ja arhivaale II. Tallinn. 71–82.
- Zgusta, Ladislav 1971:** Manual of Lexicography. The Hague.
- Zgusta, Ladislav 1989:** The Role of Dictionaries in the Genesis and Development of the Standard. – Wörterbücher. Dictionaries. Dictionnaires. An International Encyclopedia of Lexicography. NY, Berlin. 70–79.
- ÕS 1999 = Tiiu Erelt (ed.) 1999:** Eesti keele sõnaraamat. ÕS 1999. Tallinn.

## INGLISE JA EESTI KEEL LESIKOGRAAFI PILGU LÄBI. I

Enn Veldi

Eesti–inglise ja inglise–eesti leksikograafiat on suhteliselt vähe uuritud. Vastav erialakirjandus piirdub retsensioonide ja ülevaadetega ilmunud sõnastike kohta. Johannes Silvet (1895–1979) ja Paul F Saagpakk (1910–1996) jutustasid küll sellest, kuidas nad oma sõnastikke koostasid, kuid ei avaldanud teoreetilisi käsitlusi. Seni on eesti ja inglise keele leksikaalseid struktuure uuritud peamiselt tõlketeooria seisukohalt (Urve Hanko ja Gustav Liiv). Et iga keelepaar on unikaalne, siis võib inglise ja eesti keele kõrvutatav uurimine pakkuda suurt huvi. Käesolev artikkel heidab pilgu inglise–eesti ja eesti–inglise sõnaraamatute makrostruktuurile.

Artikkel algab tõdemusega, et eesti ja inglise keelel on küllaltki palju leksikaalseid sarnasusi. Aja jooksul on eesti keel päris ulatuslikult 'indoeuroopastunud' mille tulemusena on sarnasuste hulk pidevalt kasvanud.

Artikli järgnev osa käsitleb lühidalt sõnaraamatute tüpoloogiat. Henri Béjoint on seisukohal, et sõnastike liigitus ja tüpologia ei ole kattuvad mõisted. Raamatukoguhoidjad tegelevad sõnastike liigitamisega, aga lingvistide poolt pakutavad tüpoloogiad on teoreetilist laadi. Siiski tõdeb Béjoint, et sõnaraamatute näol on tegemist niivõrd mitmepalgelise nähtusega, et kõiki kultuure ja ühiskondi rahuldavat sõnaraamatute liigitust pole võimalik välja töötada. Võime öelda, et 1990 aastatel tekkis uus kriteerium sõnastike liigituseks – infokandja vorm. Traditsiooniliste pabersõnastike kõrvale on lisandunud CD-ROM ja veebipõhised sõnastikud. Taas võib täheldada suurt huvi aktiivsete ja passiivsete sõnastikega seonduva problemaatika vastu Lev V Štšerba järgi vajab iga keelepaar nelja sõnastikku: 1) L2→L1 sõnastikku dekodeerimiseks, 2) L2→L1 sõnastikku kodeerimiseks, 3) L1→L2 sõnastikku kodeerimiseks ja 4) L1→L2 sõnastikku dekodeerimiseks.

Eesti–inglise keelepaari puhul pakub suurt huvi olemasoleva sõnavara kvantitatiivne maht. Kirjanduse andmeil on inglise keeles mitte vähem kui üks miljon sõna. Julgemad ennustajad on pakkunud sõnade arvuks isegi nelja miljonit. Ühede kaante vahel pole kõiki sõnu siiski võimalik leida. *Oxford English Dictionary* teine väljaanne (1989) sisaldab umbes 500 000 sõna ja 2,5 miljonit tsitaati. Eesti keele sõnavara suuruseks on pakutud 600 000 – 700 000 sõna. See number on üsna muljetavaldav, kuid kahjuks pole kõik need sõnad leksikograafide

kergesti kättesaadavad. *Eesti kirjakeele seletussõnaraamatu* seni ilmunud vihikud (kuni *s*-täheni) sisaldavad 118 904 märksõna. Võib arvata, et sõnastiku kogumaht jääb siiski väiksemaks kui 150 000 märksõna. Samuti tuleb märkida, et eesti keele puhul on paljude valdkondade oskussõnavara leksikograafide raskesti kättesaadav. Seepärast ei saa kakskeelsete sõnastike koostaja tugineda ainult ükskeelsetele kirjeldestele, vaid peab palju aega kulutama vastete otsimisele algallikaist. Siia lisandub veel see probleem, et eestikeelne vaste võib sootuks puududa.

Et leida kinnitust oma mõtetele, võrdlesin 100 sõnakirjest koosnevat lõiku (märksõnad ja kuni *Jacobinical*) kolmes sõnaraamatus: *The New Oxford Dictionary of English* (1998), *The Concise Oxford Dictionary* (1999) ja Johannes Silveti *Inglise–eesti sõnaraamat* (1989). Selgus, et Silveti sõnaraamat sisaldas 37% esimesena mainitud sõnaraamatu kirjetest ning 51% teisena mainitud sõnaraamatu kirjetest. Kuigi kasutatud võrdlusviis annab üsna ligikaudseid tulemusi, näitab võrdlus ilmekalt, et Silveti *Inglise–eesti sõnaraamat* vajab ulatuslikku moderniseerimist.

Teiseks võrdlesin 100 kirjest koosnevat lõiku (märksõnad *jaaguar* kuni *jagumatu*) neljas sõnaraamatus, mis lähtuvad eesti keelest: *Eesti kirjakeele seletussõnaraamat*, ÕS 1999, Paul Saagpaku *Eesti–inglise sõnaraamat* (1992) ja Johannes Silveti *Eesti–inglise sõnaraamat* (1980). Siin ilmsesid samuti huvitavad asjaolud. Selgus, et Saagpaku sõnaraamatus, mille maht on 150 000 sõna, puudus 30 sõna, mis on EKSS-is. Samas oli Saagpaku sõnaraamatus 40 sõna (koos pärisnimedega 46), mis puudusid EKSS-is, nt *jaa-hääl*, *jaamaline*, *jaamaooteruum*, *jaamapakihoiuruum*, *jaamapiirkond*, *jaamapiletikassa*, *jaamaplats*, *jaamaraamatukapp*, *jaanalinnufarm*, *jaanalinnukasvandus*, *jaapanitar*, *jaa-ütleja*. Vaadeldud lõigu alusel võib teha esialgse järelduse, et tänapäeva eesti keele seisukohalt võib liiasus Saagpaku sõnaraamatus ulatuda kuni ühe kolmandikuni sõnaraamatu mahust. Kindlasti vajab Saagpaku sõnaraamatu empiiriline põhi edasist uurimist.

# OTSE, SIRGE AND ÕIGE: A DOMAIN OF METAPHORIC EXTENSION IN ESTONIAN

Haldur Õim

University of Tartu

## 0. Introduction

The original meanings of the Estonian words *sirge* 'straight' *õige* 'straight' and *otse* 'straight' are related to the spatial relations – they characterize the shape or motion of objects in space. In contemporary language these meanings partly overlap, but only partly. At any rate, they cannot always be used in the same situations. Their most general opposite is *kõver* 'curved' although things and motions can be curved or bent in a number of ways. The opposite of *kõver* – *sirge* – explains best the meaning under discussion. *Sirge* combines, for instance, with wire, board, tree, and motion. And if something is *sirge*, then it must be without any curves or bends, at least within some agreed range or distance. *Õige* is practically not used in this sense anymore. However, the existence of this sense is proved by the expressions (*traati*) *õigeks painutama* 'to bend the wire straight' and *õgvendama* 'to straighten' (with the corresponding dialectal adjective *õgev* 'straight'). The use of the word *otse(ne)* in this sense is least of all – and most vaguely – associated with straightness. The word is practically not used in this sense with regard to things. Perhaps only in such expressions as *seiske otse* in the meaning 'stand straight'. However, when we say *pilt on/ei ole otse* 'the picture is/is not straight' and *sättis pildi otseks* 'he straightened the picture', there is no relation to straightness in the above sense. On the other hand, the sense 'straight' is included in the content of the uses that characterize the word *otse* and its combinations: *otsetee* 'a straight (path)' *ta läks otse üle põllu/läbi metsa* 'he went straight across the field/through the forest' and *tuli otsejoores minu juurde* 'he came straight up to me'. Here, too, there are differences in the shades of meaning in comparison with *sirge*.

Therefore, it is likely that they cannot be substituted freely in these uses, e.g. *otsetee (läbi metsa)* 'a straight (path) (through the forest)' need not be a physically visible path as an object, and it need not be straight in the sense 'without curves or bends'. *Otse* would rather mean here motion towards the target without **unnecessary** curves or bends, thus not necessarily absolutely straight but as straight as possible. This constitutes the connection between the discussed meanings of *otse* and *sirge*. And this may explain why the Finnish *suora* and *suoraan* are appropriate and acceptable for the expression of both meanings, as are the Russian *prjamo* and *prjamoj*.

In addition, one can point out that the word *otse* has such uses in the expression of spatial relations or relationships that the words *sirge* and *õige* do not have. For example, when we say *Otse üle järve oli (paistis) mingi valge ehitis* 'Straight across the lake there was a white building' then what does it actually mean? It has no direct association with the meaning of *sirge* because one cannot, in fact, see or look in an arched or zigzag manner. It would rather imply, first, the presence of a viewer, and, second, that if the viewer, who is standing crosswise with the shoreline, looks across the lake without turning his eyes or head (one's eyes and the forehead are closely related), then he will see the mentioned object. *Otse* 'straight' is etymologically associated with the word *ots(a)* 'end' and the Finnish etymological dictionary provides as the primary Finnish etymology of *otsa* 'Stirn, oberer Teil der Stirnseite', i.e. the forehead (SSA2: 275).

However, the main purpose of this article is not to differentiate the original meanings of the above-mentioned words and to compare them with the translational equivalents in other languages. Instead, I would like to focus on the system that governs the metaphoric extensions of their meanings. In fact, all these words have multiple uses in Modern Estonian that came into existence by means of metaphoric extension in relation to their original meanings. And most importantly, similar transfers in the same direction have taken place in other languages as well, although – and it will make the problem an interesting one – not exactly in the same way.

Below are some preliminary examples from Estonian: *sirge* ~ *sirgjooneline* 'straight ~ straight-line' can be used not only with regard to objects or motion but also with one's character or interaction. Similarly, the word *otse* has the metaphoric derivatives *otsekohe*, *otsemaid* 'straight away' not to speak of the meanings of *otse* "immediate" (*otsene laim* 'downright slander' and *otsene tähendus* 'direct meaning'), "downright, outright" (*otse niimoodi ma ei mõelnud* 'I didn't mean it so actually'). The words *õige* 'correct' and *õigesti* 'correctly' have a figurative meaning that contrasts with the meaning "wrong, incorrect" and has undoubtedly become the basic meaning in Modern Estonian (*õige asend* 'correct position' *õige jutt* 'a truthful story' and *õige mõte* 'a correct idea'). The same applies to the word *õigus* 'a right' (e.g. *õigus teha midagi* 'the right to do sth').

The present article was prompted by an article by Alan Cienki "STRAIGHT: An Image Schema and its Metaphorical Extensions" which was published in the journal "Cognitive Linguistics" (Cienki 1998). Cienki claims the existence of a universal image schema, which he designates with the English word STRAIGHT, and certain (rather) universal extension rules from this image schema as the source domain to the more abstract target domains. Cienki's examples come in addition to English from Russian and to a lesser extent from other languages (Polish, Hungarian, and Japanese). I took an interest in what the data from Estonian and our closest cognate language Finnish would enable us to conclude by comparison with Cienki's claims, especially with regard to the role of the STRAIGHTness concept in the system of semantic change and transfer.

## 1. Background

The problem under discussion is, in a broader sense, a typical problem in cognitive semantics: metaphoric extensions as a source of semantic change, image schemas as the main basis for metaphoric extensions (that need not always lead to semantic change). I am mainly interested in one aspect of this problem – the essence of semantic change, the possibly valid regularities, the various manifestations of this regularity in different languages –

and only then in the claims of cognitive semantics concerning the role of image schemas and other theoretical constructions in it.

Academic interest in the regularities of semantic change has a history of at least a hundred years. When M. Bréal introduced the term *semantics* (in French *sémantique*) at the end of the 19th century, he had in mind a branch of linguistics that would study the regularities of semantic change. As is known, at that time the scientific study of language meant the study of its history, be it speech sounds, morphological forms, syntactic constructions, or meanings.

The first half of the previous century witnessed considerable interest in semantic change. Linguists hoped to find regularities that would be analogous to the regularities that are valid in phonetic change. Emphasis was still on possible causes of semantic change. For example, A. Meillet emphasized the role of social factors: a change in the function of things and the content of human relations in society will bring along changes in the meanings of the words that denote them although the signifier will remain the same (Meillet 1906). In contrast, H. Sperber in his book *Einführung in die Bedeutungslehre* (Sperber 1923) stresses the role of psychological (emotional) factors. In addition to the role of taboos, euphemisms, irony, and pejoration, he points out that metaphors often fulfil the same function. For example, when the French soldiers called the submachine gun as a typewriter (*une machine à coudre*), their purpose was to soften the negative force of the object under discussion.

However, such a 'macroanalysis' of the causes of semantic change yielded little about the language-internal mechanisms of change. The best-known work of the so-called old period was "Meaning and Change of Meaning" by G. Stern (Stern 1931). Structuralism, which prevailed after the 1930s, focused on the synchronic description of language, and diachronic semantics was largely neglected.

In the meantime semantics became a fully-fledged independent branch of linguistics, but its focus had a purely synchronic character. Serious diachronic semantics has been revived only during the past few decades, which was largely thanks to the cognitive, in a broader sense functional, approach to



language, which had an interdisciplinary character. The latter claim could be supported by a highly characteristic and valid quotation from a paper that formally belongs to cognitive anthropology: "Over the past twenty years, the rise of functional approaches to linguistics, the birth of cognitive semantics, the expansion of the Greenbergian-style approach to linguistic universals, and an increased interest in the mechanisms and pathways of grammaticalization have all, in their own way, led to an increased theoretical interest in diachronic semantics. These trends have led to a view that, instead of being wild, chaotic and unpredictable, semantic change is in many cases regular and is often predictable, and that natural tendencies of semantic change arise out of the universality of certain cognitive processes and/or out of universal traits to be found in the production and interpretation of fully contextualized natural discourse" (Wilkins 1993: 3). Recent years have seen a number of important papers in diachronic semantics, and most of them are associated with the cognitive (functional) approach to language (Geeraerts 1997; Sweetser 1990; Traugott 1985; Lehrer 1986; for a more detailed overview see e.g. Wilkins 1993).

For the sake of accuracy, one has to point out that interest in semantic change has taken two paths. First, interest in those processes that are known under the term 'grammaticalization': the meaning of a lexical unit changes so that the resulting unit is not anymore part of the lexical system of a language but belongs to its grammatical system. Second, interest in those processes that take place within the framework of the lexical system, first and foremost polysemy, genesis of new lexical meanings that are based on the existing meanings (the semantic system of a language can be renewed in a number of other ways as well, e.g. by adopting loanwords for new meanings).

The present article will focus only on the second type of semantic change, i.e. the genesis of new lexical meanings on the basis of the existing meanings, whereas the new meanings come into existence by means of a mechanism that enables us to make a generalization that can be observed in more than one language.

## 2. Image schema STRAIGHT and the Estonian *otse*, *sirge*, and *õige*

Typically, an image schema is defined in cognitive semantics as “a recurring, dynamic pattern of our perceptual interactions and motor programs that gives coherence and structure to our experience” (Johnson 1987: XIV). Incidentally, coherence and structure serve as a basis for the extension of specific properties of an image schema into abstract domains. Lists of basic image schemas have been put forward (e.g. CONTAINER, PATH, FULL-EMPTY, NEAR-FAR; see e.g. Cienki 1998: 108). Cienki has suggested that STRAIGHT could also fit as a relatively universal image schema. Being originally an organizer of knowledge about the shape and motion of objects, it serves as a basis for metaphoric extensions to such domains as time, events, discourse, thought, control, social norms, morality, truth, and law. As has been said, his data come first and foremost from English and Russian.

There is no doubt that the Estonian words *otse(ne)*, *sirge*, and *õige* belong to this image schema. Below I will discuss their main metaphoric extensions as compared to Cienki’s treatment.

### 2.1. *Otse/otsene/otseselt*

First and foremost, the Estonian *otse* ‘straight’ is of special interest (as contrasted with *sirge* and *õige*). Therefore, we will take a closer look at its primary meanings.

1) It is difficult to say what was the original meaning of *otse* – the one that gave rise to metaphoric extensions into other domains. However, it is quite likely that this meaning may be manifested in such words and expressions as *vaatama otse* ‘look straight’ (cf. also the command *otse vaat* ‘eyes front!’ and the above-mentioned sentence *ots üle järve paistis/oli ..* ‘right across the lake there was ..’). Thus, this meaning belongs to the visual sphere that is confirmed by the mentioned circumstance that etymologically the Estonian *otse* ‘straight’ is associated with *ots* ‘end’ (see Raun 1982: 112). Looking straight assumes that not only the eyes but also the head is turned ‘straight’ (one’s eyes cannot look straight

if the head is not turned in the same direction). The meaning of *ots(a)* is associated with *otsaesine* ‘forehead, lit. end-front’ or with the facial part of the entire head, cf. *otsa vaatama/vahtima* ‘to look at sb, lit. to look at the end’

2) The closest extension into a non-visual (but non-abstract) sphere is manifested in such Estonian uses as *pööra (auto) rattad otse* ‘turn the (car) wheels straight’ where the analogy between the eyes and looking is transparent. The same use that proceeds from the image schema STRAIGHT can be observed in English, Finnish, and Russian:

*to be oriented straight on*

*(auton) pyörät ovat suoraan* ‘the (car) wheels are straight’

*poverni koljosa prjamo* ‘turn the wheels straight’

3) The semantic relation between looking and motion has been under discussion for some time already and is, in fact, beyond doubt. For example, in Estonian *vaatas läbi akna välja* ‘looked out the window’ and *hüppas läbi akna välja* ‘jumped out of the window’ have a similar syntactic and semantic structure. It is true, of course, that, as immediate parallels to looking, only such motions are suitable that assume a straight-line trajectory (as *hüppama* ‘to jump’), which cannot be changed at some point of motion (e.g. walking or running may have trajectories with any degree of complexity).

Therefore, one could bring out the motion-related meaning(s) as the next meaning of *otse* and its derivatives. Some examples of this kind were provided above. Actually, there are two (sub)meanings. The one that is more closely related to the previous meanings is manifested in such uses as *ta läks otse edasi* ‘he went straight ahead’ and *auto sõitis otse* ‘the car moved straight’ which implies simply motion without straying from the previous direction.

The other and much commoner meaning is associated with motion towards a target (that can be either directly visible or not). To move straight in this sense means to move towards the target straight without any curves or bends on the way:

*ta läks otse läbi metsa* ‘he went straight through the forest’

*talvine otsetee läks üle järve*

'the winter straight road was across the lake'

*ta tuli otsejoones/otseteed meie juurde*

'he came straight up to us'

*kui ei ole võimalik minna otse, kas siis poleks võimalik minna ringiga?*

'if it is impossible to go straight, then wouldn't it be possible to take a circuitous route?'

As has been pointed out, *otse* need not mean here totally straight motion. First, it is important that there is a destination towards which one is moving (simple straight-line motion does not assume it), and, second, in this context *otse* means that one will move towards this destination **as straight as possible** (along a straight path), which does not exclude physical detours in case of obstructions (e.g. when walking straight through a forest one must move around the trees and bushes). Thus, in this case *liikuma otse* 'to move straight' means "to move towards the destination as straight as possible without any curves or bends on the way for which there is no objective need" A peculiar proof of the relation between this meaning of *otse* and its primary (visual) meaning is the circumstance that *silmnäolt* 'straight, lit. eye-faced' is a synonym for *otse*: *ta läks silmnäolt Jüri juurde ja ..* 'he went straight up to Jüri and .. The only possible justification for the peculiar structure of this word seems to be the need to emphasize the role of one's eyes and the face in this motion – one would move so that one's eyes and face are directed straight at the target.

As has been said, this semantic extension has been actualized in English, Russian, and Finnish. However, the lexical realizations once again reveal some differences. The English *to move along a straight path* or *to go directly to X*, the Russian *idti prjamo (tsherez pole)* 'to go straight across the field', and the Finnish *mennä suoraan (pellon üli/läpi metsän)* 'to go straight (across the field/through the forest)' are not associated with looking/seeing. They emphasize the straightness of the path, and the literal translation of the corresponding words into Estonian would be *sirge(lt)*. Therefore, Cienki includes straight motion into the source domain of the metaphors under discussion: "While some of the metaphors characterize the target domain in terms of a straight object, many involve processes, and therefore the source domain

often involves straight motion (along a path)” (Cienki 1998: 116). The enumerated languages, including Finnish, do not make the distinction that can be found in Estonian between *otse* and its derivatives and *sirge* and its derivatives. In none of the meanings under discussion *sirge(lt)* would be an acceptable synonym for *otse* in Estonian. In contrast, at times *õige* or its dialectal or colloquial equivalent *õkva* would be acceptable, e.g. *pööra (auto) rattad õigeks* ‘turn the (car) wheels straight’ and *läks õkva üle põllu* ‘went straight across the field’

By the way, it is worth pointing out that the English *straight* derives etymologically from the same source as the verb *to stretch* ‘to make stiff or hard as in straining a cord’ (Skeat 1993: 475). A taut string is straight, of course. Thus, this source domain is unrelated to looking.

4) It is not my purpose to analyze all the meanings of *otse* but only those that are associated with metaphoric extensions into more abstract domains. However, before examining the extensions of *otse* (as well as *sirge* and *õige*) and comparing them to Cienki’s data, I would like to discuss another usage type (meaning) of *otse*, which cannot be traced back so clearly to the visual source domain, but which seems to act as a mediating source domain for some abstract extensions. It is the meaning that can be illustrated by the following expressions:

*need saidad on otse ahjust*

‘these buns are straight from the oven’

*X oli otse tema ees/taga*

‘X was right in front of him/X was right behind him’

*otse teeraja kõrval* ‘right next to the path’

*tabati otse teolt* ‘was caught red-handed’

*otseühendus* ‘direct link’ *otseülekanne* ‘live broadcast’

Here the use of *otse* is not based on the straightness of the relative position (distance) or motion. Nor can it be observed from the ‘frontal view’ The critical semantic component would rather be ‘absence of the connecting link’ An acceptable synonym is mostly *vahetult* ‘immediately’ and *sirge(lt)* and *õige(sti)* are unacceptable. The dialect word *õkva* is again acceptable.

However, this meaning is yet clearly related to the previously discussed ones. For example, in case of motion an unnecessary deviation from the straight line, a zigzag, can be regarded as a connecting link that makes the motion non-straight.

And this meaning is expressed, for example, in Finnish and Russian by means of the same expressions as in case of the previous meanings:

Finnish: *suoraan meidän edessä* 'right ahead of us' *suora yhteys* 'direct link' *suora lähetys* 'live broadcast'

Russian: *prjamo pod tvoim nosom* 'right under your nose', *prjamaja svjaz* 'direct link' *prjamaja peredacha* 'live broadcast'

## 2.2. Metaphoric extensions into abstract domains

Next I am going to examine those extensions into the abstract domains, which were discussed by Cienki with regard to the image schema STRAIGHT from the point of view of English and Russian. As we will see, most of them, though not all, are valid for Estonian as well. However, even in those cases, where the extension is valid, Estonian has its own peculiarities.

### 2.2.1. Communication, character, and thought

#### 1) Communication and thought

It is a well-known fact that motion as a source domain gives rise to metaphoric extensions to the domain of communication. It is natural because in communication some information – some knowledge but immediately the text – moves from one person to another person. The problem is which aspects of physical motion are transferred in such extensions and what kind of meaning they will acquire in the target domain – in the communication schema.

1a) Cienki (1998: 116) uses the following wording for the main principle of the aspect that is of interest to us: TO SPEAK IN A MAXIMALLY INFORMATIVE WAY IS TO TRANSFER WORDS ALONG A STRAIGHT PATH:

*tell it to me straight*

*to say something straight from the heart [with one's true emotions]  
she told him flat out that she was not interested*

The same is, in principle, valid for Russian (Cienki: 130–131):

*prostite za prjamoje slovo*  
'excuse me for my directness, lit. for the direct word'  
*prjamoi otvet* 'a straight answer'  
*govorit' naprjamik*  
'to be to the point, to say something point blank, lit. to speak straight'

When we examine Estonian and Finnish, we will see that this claim is, in principle, valid for these languages as well.

Estonian:

*ütles otse välja* 'said point-blank'  
*otsesõnu öeldes* 'to put it bluntly'  
*otsekohene jutt* 'straightforward talk'  
*sirgjooneline arvamusavaldus* 'straightforward opinion'

It is noticeable that here, too, *otse* prevails as the carrier of the corresponding feature. However, *sirge* is an acceptable synonym, too (in the combinations *sirgjooneline*, *sirgjooneliselt*). It is likely that the difference that we pointed out in the treatment of the original physical motion (straight motion to somewhere need not mean (absolutely) straight motion) may not be relevant in communication. In case of physical motion the obstacles on the way may prove to be important, which makes straight motion non-straight. At least in face-to-face communication (that the previous expressions refer to) they are ignored.

Virtually all the Estonian expressions can be translated into Finnish by means of *suora*, its derivatives, or compounds:

*otse, otsekoheiselt* = *suoraan* (*puhua suoraan, suorapuheinen, suorasana-*  
*nainen*)

As said before, Finnish does not make a lexical difference between *otse* and *sirge*. Therefore, for example, *sirgjooneline (jutt)* 'a straightforward talk' is also *suora (puhe)*.

It is worth noting that in Estonian, too, a straightforward talk (as well as the corresponding Finnish expressions with the word *suora*) refers to the most informative kind of communication. However, when we refresh our memory with the maxims that are known to us from pragmatics, then such a conversation may not meet the demands of far from all the maxims, not at least the maxims of the politeness principle (e.g. the maxim of tact; see Leech 1983: Ch. 8). The latter requirement would often be met if the information were presented indirectly, in the form of hints, so that there would be a detour in the source domain, i.e. the destination will be reached by taking some sideroads. Therefore, one cannot say that STRAIGHT communication represents a communicative norm or **typical** communication as a whole.

Second, it is noticeable that Estonian does not use such expressions for the characterization of communication that would directly correspond to the English *flat*, which is associated first and foremost with negative communicative forms (*a flat denial*), see also the above-mentioned example). Although Cienki (111) finds that the extension is natural because in case of a horizontal object *flat* is an analogue to a *straight* (line) – and we could claim the same for Estonian (and in certain restricted cases we do use it: *silus kortsus paberi sirgeks* ‘smoothed out the crumpled paper’) – I have not found any cases where the Estonian equivalents *tasane*, *sile*, and *lame* of the English *flat* would be extended to the communication domain in this meaning. It is true that one can say *lame nali* ‘a stale joke’ in Estonian, but it seems unrelated to the image schema STRAIGHT. The same can be said about Finnish.

**1b)** However, the Estonian *otse(ne)* occurs in the characterization of communication in other uses as well, not only as an expression of informativeness that is an extension of the straightforwardness of motion. One can say:

*ütles otse näkku*

‘said point-blank, lit. said straight into sb’s face’

*valetas otse suu sisse*

‘lied flatly, lit. lied straight into sb’s mouth’

Here the source domain is the meaning that emphasizes not the straightforwardness (informativeness) of the conversation but



rather the 'absence of connecting links' (in the above-mentioned sense (4)). As it is a characteristic use of *otse* in the source domain, then it is not surprising that *sirge* and *õige* (with the exception of the dialectal *õkva*) do not occur in this meaning. The uses need not have only a negative meaning although the examples could suggest it (though they must be more frequent). According to the author's linguistic instinct, it would also be acceptable to say

*otsene tõde* 'plain truth'

It would be the truth that was told in a straightforward manner.

**1c)** One aspect of communication is the authenticity or truthfulness of an utterance. It is possible in Estonian to express this aspect by means of the adjective *õige* 'right, correct' and the noun *õigus* 'right' which has been derived from the former. The primary meaning of *õige*, as was pointed out earlier, belongs to the image schema STRAIGHT, being synonymous with the meaning of *sirge* and applied to objects rather than motion (a wire can be bent straight). In Modern Standard Estonian *õige* is seldom used in this meaning, but instead the meaning 'truthful' is mainly expressed by *õige* or its derivatives:

*See jutt on õige* 'This story is true'

*Sul oli õigus* (*kui sa ütlesid, et ..*)

'You were right (when you said that ..)'

*See teade osutus õigeks* 'This news proved to be accurate'

It is natural – and it is characteristic of not only Estonian – that *õige* and its derivatives in this sense can be used not only in the context of communication (with regard to the forwarded information) but also with regard to other objects of thought if there is a need to assess their correspondence to reality (opinions, thoughts, decisions, and plans):

*Sa arvasid õigesti, et ..*

'You were right when you thought that ..'

*See on õige mõte* 'That's a good idea'

*See oli õige otsus* 'That was the right decision'

The same is basically true of the Finnish equivalent *oikea* to the Estonian *õige*, which carries the same meaning:

*Olit oikeassa (kun sanoit, että ..)*

'You were right (when you said that ..)'

*Se oli oikea päätös* 'That was the right decision'

Thus, in Estonian and Finnish *õige* has become a basic word that expresses TRUTH(FULNESS). Again there is a difference from English, where Cienki (120) postulates the principle TRUTH IS PLAIN on the basis of such expressions as

*the plain (and simple) truth*

*the unadulterated (unvarnished, naked) truth*

Similarly to the previously discussed *flat (flat denial)*, the Estonian equivalents to the word *plain* that belong to the image schema STRAIGHT (the Estonian equivalents again being *lame* 'flat' and *ühetasane* 'even') cannot be used as metaphoric equivalents to truthfulness. According to Estonian (and Finnish), THE TRUTH IS A RIGHT (= STRAIGHT) OBJECT.

The difference is, of course, not a principal one because there is a perceptual relation between a straight line and an even surface. However, the difference is of interest from the point of view of lexical typology.

For the sake of completeness one should point out that the above-mentioned principle was not intended to cover the entire domain of TRUTHFULNESS. In Estonian, too, one might say  *alasti tõde* 'naked truth' and *karm tõde* 'bitter truth', but these expressions do not proceed from the image schema STRAIGHT. It is true that in the negative sense one could say in Estonian, for example, *vildakas arusaam* 'misconception, lit. slanting understanding' A surface can be slanting but not a one-dimensional object. However, one might suspect that this expression may be a translation loan rather than a metaphoric extension that originated from the semantic system of Estonian (the positive equivalent is absent).

## 2) Character

It is quite understandable that the way how a person communicates with other people largely depends on his or her character. Therefore, it is not surprising that the expressions that are applied to someone's character and communication are the same. In that case the use of the same expressions indicates that a person would **typically** communicate as described by the corresponding expression, that this manner of communication depends on one's character. Of the expressions that are relevant in this context, there are two that are characteristic of Estonian:

*sirgjooneline (inimene, iseloom)*

'a straightforward (person, character)'

*otsekohene (inimene, iseloom)*

'a straightforward (person, character)'

The dictionaries list also *otsemeelne* 'lit. straight-minded' and *otsesõnaline* 'lit. straight-worded' (e.g. Õim 1991: 348), but they are rarely used. Nonetheless they evidently follow the extension pattern that is characteristic of Estonian.

### 2.2.2. Time and events

Cienki (122) points out that "one conceptualization of time in European and American cultures involves the metaphor TIME IS A MOVING OBJECT .. The default assumption is that this motion is STRAIGHT" When applied to time and the temporal sequence of events in this context, an **immediate** sequence of these events is implied: IMMEDIATE ACTION IS MOTION ALONG A STRAIGHT PATH (Cienki 1998: 123):

*to discover something straight off*

*a straightaway denial*

*straight ahead of the news is a story about ..*

It is the meaning for which Estonian uses *otse* (with its derivatives) in the sense 'immediately, without any connecting links'. It is not important whether each of the previous examples can be translated into Estonian by means of *otse*. What is important is that for Estonian, too, the same general principle is

valid. The expressions *otsekohe* and *otsemaid* 'straightaway' represent, in fact, the immediate temporal sequence of events. The expression *otsekohe* is especially interesting in that *kohe* alone means largely the same, but *otsekohe* is more resolute and emphasizes the exclusion of intermediate actions or events:

*mine ja räägi temaga kohe* 'go speak to him right away'  
*mine ja räägi otsekohe temaga* (the same)

The second sentence requires that there should be no other events between the moment of communication and speaking; the first sentence does not require it so categorically. In this case *kohe* 'at once' can be supplemented either implicitly or explicitly by intermediate events 'right after X ..'; *otsekohe* 'right away' excludes this possibility.

The role of the Finnish equivalent *suora* to the Estonian *otse* is even more apparent. Dictionaries translate *otsekohe* into Finnish as *suuraa päätä, oikopäätä*, where *oiko* is obviously related to *oikea*.

### 2.2.3. Order and control

According to Cienki, the source domain STRAIGHT gives rise to the target domains of order and control – something is in order, has been tidied up, 'is as is proper' (Cienki 1998: 124ff):

*The room is straight again*  
*to straighten out what someone else did wrong*

In Estonian there is no such clear lexical relation between *otse/sirge/õige* and, for example, putting a room in order. However, the Estonian verb *õiendama* (< *õige*) has on a more general level the meaning 'to straighten out':

*Ma õiendan selle asja kohe homme ära*  
 'I will straighten this matter out tomorrow'

It is also possible to use as synonyms of the verb *korraldama* 'to arrange' (derived from the noun *kord* 'order') – though only in certain styles – *joonde ajama, jutti ajama* 'to straighten out', where *joon* and *jutt* 'line' are semantically related to the adjective *sirge* 'straight'. To end with, Estonian has the idiomatic

expressions (*midagi sirgeks rääkima* 'to smooth away a problem, lit. to talk sth straight' and *teeme sotid sirgeks* 'let's straighten it out, lit. let's straighten out the accounts' where *sirge* is represented even lexically.

Thus, the relation between straightness and order, as pointed out by Cienki, is not unknown in Estonian although it is impossible to translate each English expression that contains the corresponding lexical unit into Estonian (by using the dictionary equivalents) and the other way round. However, this is just the reason why this phenomenon is of great interest.

In the case of Finnish the verb *suorittaa* 'to perform' attracts attention because it has an obvious association with the adjective *suora*. One could compare it to the Estonian verb *õiendama*, but the former has a much more general and stylistically neutral meaning in Finnish than the verb *õiendama* in the lexical system of Modern Estonian.

### 3. Summary

In the course of time everything changes in language – in each language. The meanings of words are subject to change, too. When semantics under its present name came into existence about a hundred years ago, it was the branch of linguistics that was to study the regularities of semantic change.

After a few decades linguists reached an understanding that in comparison with, for example, phonetic change that had been studied extensively by that time in the case of meaning the situation is hopelessly vague. Apart from some general types of change, as broadening or narrowing of meaning, linguists could not establish any such regularities that would be comparable to the phonetic laws that were valid in the case of the phonetic side of words.

A possible reason was undoubtedly the fact that very little was known about the essence of meaning and its organization in language. Contemporary semantics – although it has branched off into many theories that are in part very different – can say much more about these aspects of language.

One of the first topics that cognitive semantics focused on was metaphORIZATION. The chief tenet was that metaphors

constitute a natural part of language use, and that new metaphors are (mostly) created according to certain general regularities. The most important regularity is that metaphors are typically based on the so-called image schemas, the structures by means of which humans organize their knowledge, experience – and also meanings. Image schemas constitute the framework in which humans extend their concrete experience, beginning with sensuous perception of the world, into more abstract domains. The main mechanism of these modifications is metaphoric or metonymic extensions.

Only one more step was needed to realize that metaphorization underlies (many) semantic changes (which had been known for a long time) and that the same image schemas will enable us to clarify the types and regularities of semantic change.

The present article was prompted by an idea that was put forward by Cienki (1998). The image schema STRAIGHT, which originally organizes our sensuous experience with regard to things and motion, can also explain the extension of the corresponding meanings into more abstract domains. The present article examined the Estonian words *otse*, *sirge*, and *õige*. Their primary, 'perceptual' meanings belong to the image schema STRAIGHT. Their more abstract meanings were then compared with the ones that Cienki suggested for English and Russian. It appears that in Estonian the primary meanings were extended into the same basic domains that were highlighted by Cienki: discourse, thought, truth, time and events, and order.

It also appeared that it was not important where the meanings of the corresponding words originated that express the straightness of things or motion, as the English *straight* or the Estonian *otse*. The English *straight* is related to the verb *to stretch* (e.g. a cord), and the Estonian *otse* is related to looking. However, once these words have acquired the meaning that expresses the straightness of things or motion, and, thus, they belong to the image schema STRAIGHT, then similar semantic changes will occur.

As the mutual influence of the languages under discussion is excluded, one should draw the conclusion that the described semantic changes are based on more general regularities that are

independent of the specific language, which evidently belong to the sphere of human cognition and thought.

As is known, one of the tenets of cognitive semantics is that the organization of human cognition and thought constitutes the basis of the semantic system of language. Thus, the regularity that is valid in the described semantic changes – directly derived from the analysis of linguistic data – can be regarded as proof of the validity of this abstract claim.

## REFERENCES

- Cienki, A. 1998:** STRAIGHT: An image schema and its metaphorical extensions. – *Cognitive Linguistics*, vol. 9, no. 2. 107–149.
- Geeraerts, D. 1997:** *Diachronic Prototype Semantics. A Contribution to Historical Lexicology.* Clarendon Press. Oxford.
- Johnson, M. 1987:** *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason.* University of Chicago Press. Chicago.
- Leech, G. N. 1983:** *Principles of Pragmatics.* Longman. London and New York.
- Lehrer, A. 1985:** The influence of semantics fields on semantic change. – J. Fisiak (ed.), *Historical Semantics and Historical Word-formation.* Mouton de Gruyter. Berlin. 283–296.
- Meillet, A. 1906:** Comment les mots changent de sens. – *Année sociologique*, 9. 1–38
- Õim, A. 1991:** *Sünonüümisõnastik.* Tallinn.
- Raun, A. 1982:** *Eesti keele etümolooiline teatmik.* Maarjamaa. Rooma-Toronto.
- Skeat, W.W. 1993:** *The Concise Dictionary of English Etymology.* Wordsworth Reference. Norhaven.
- Sperber, H. 1923:** *Einführung in die Bedeutungslehre.* Schroeder. Bonn.
- SSA2 1995:** Suomen sanojen alkuperä, 2. (Suomalaisen Kirjallisuuden Seuran Toimituksia 556; Kotimaisten Kielten Tutkimuskeskuksen Julkaisuja 62.) Suomalaisen Kirjallisuuden Seura; Kotimaisten Kielten Tutkimuskeskus. Helsinki.
- Stern, G. 1931:** *Meaning and Change of Meaning.* Elanders Boktryckeri Aktiebolag. Gothenberg.

- Sweetser, E. 1990:** From Etymology to Pragmatics. The Mind-as-Body Metaphor in Semantic Change. Cambridge University Press. Cambridge.
- Traugott, E. 1985:** On regularity in semantic change. – Journal of Literary Semantics, vol. 14. 155–173.
- Wilkins, D. 1993:** From Part to Person: Natural Tendencies of Semantic Change and the Search for Cognates. Cognitive Anthropology Research Group at the Max Planck Institute of Psycholinguistics. Working Paper No 23. Nijmegen.

*Otse, sirge ja õige:*

## ÜHEST METAFORSE ÜLEKANDE ALAST EESTI KEELES

Haldur Õim

Aja jooksul muutub keeles – igas keeles – kõik. Muutuvad ka sõnade tähendused. Kui semantika oma praeguse nime all umbes 100 aastat tagasi sündis, siis nimelt kui keeleteaduslik distsipliin, mis pidi uurima tähenduste muutumise seaduspärasusi.

Mõnekümne aastaga jõuti aga arusaamisele, et võrreldes näiteks häälikumuutustega, mida selleks ajaks piisavalt hästi tunti, on tähenduste puhul olukord lootusetult segasem. Peale üksikute üldiste muutumise tüüpide, nagu tähenduste laienemine või kitsenemine, ei osatud siin mingeid selliseid regulaarsusi välja tuua, mis olnuksid võrreldavad sõnade häälikulise külje puhul kehtivate häälikumuutuste seadustega..

Üks põhjusi oli kahtlemata see, et tähenduste endi olemust ja nende organisatsiooni keeles tunti väga vähe. Tänapäeva semantika – ehkki see jaguneb paljudeks osalt vägagi erinevateks teooriateks – oskab keele nimetatud aspektide kohta palju rohkem öelda.

Kognitiivses semantikas oli üheks esimeseks teemaks, millesse süveneti, metaforisatsioon. Põhiteesiks sai, et metafoorid moodustavad keeikasutuse loomuliku osa ja uute metafooride loomine toimub enamasti teatud üldiste seaduspärasuste järgi, millest olulisim on seaduspärasus, et tüüpiliselt on metafooride aluseks teatud kognitiivsed mudelid või skeemid, struktuurid, mille abil inimesed organiseerivad oma teadmisi, kogemusi – ja ka tähendusi. Need skeemid on raamiks, milles inimesed kannavad üle oma konkreetseid kogemusi, alates maailma meelelisest tajust, abstraktsematele aladele; ning põhimehhanismiks neis teisendustes on metafoorsed või metonüümilised



tähendusülekanded. Siit oli vaid samm ideeni, et paljude tähendusmuutuste taga on metaforisatsioon (seda teati tegelikult ammu) ning et samade kognitiivsete skeemide najal võib seega püüda selgitada tähenduste muutumise tüüpe ja seaduspärasusi.

Käesoleva artikli aluseks on A. Cienki (1998) poolt esitatud idee, et kognitiivne skeem (Cienki kasutab terminit *image schema*) STRAIGHT, mis algselt organiseerib meie meelelist kogemust asjade ja liikumiste "sirguse" kohta, võimaldab seletada ka vastavate tähenduste ülekandmist abstraktsematele aladele. Artiklis on vaadeldud eesti keele sõnu *otse*, *sirge* ja *õige*, mille algsed, "pertseptuaalsed" tähendused väljendavad asjade ja liikumiste ruumilist sirgust, ning võrreldud nende abstraktsemaid tähendusi nendega, mida Cienki on kästlenud eelkõige inglise ja vene keele osas. Täiendavalt on kõrvutuseks lisatud näiteid soome keelest kui sugulaskeelest (mida inglise ega vene keel teatavasti ei ole). Selgub, et eesti keeles võib konstateerida algühenduste ülekandumist samadele põhilistele sihtaladele, mille on esile toonud Cienki, kuid eesti keel pakub ka huvitavaid eripärasusi.

Seejuures võib aga kohe ühe olulise järeldusena esile tuua asjaolu, et ei ole oluline, kust vastavate sõnade tähendused, mis väljendavad asjade või liikumise sirgust, nagu inglise *straight* või eesti *otse*, algsemalt pärit on: kui inglise *straight* on seotud verbiga *to stretch* 'pingutama (nt nõõri)' siis eesti *otse* on ilmselt seotud vaatamisega. Ent kui need sõnad on omandanud tähenduse, mis väljendab asjade või liikumise sirgust, ja seega kuuluvad kognitiivse skeemi STRAIGHT alla, hakkavad toimuma ühesugused tähendusmuutused.

Kuna ilmselgelt on välistatud vaadeldavate keelte vastastikune mõju antud valdkonnas, tuleb järeldada, et kirjeldatud tähendusmuutuste aluseks on üldisemad, konkreetsest keelest sõltumatud seaduspärasused, mis kuuluvad inimese tunnetuse ja mõtlemise sfääri.

Kognitiivse semantika üks põhiteese on teatavasti, et inimliku tunnetuse ja mõtlemise organisatsioon on keele semantilise süsteemi alus. Kirjeldatud tähendusmuutustes kehtivat seaduspärasust – mis on tuletatud otseselt keelematerjali analüüsist – võib seega vaadelda kui üht tõendit selle abstraktse teesi paikapidavuse kohta.

Eesti keele sõnade *otse*, *sirge* ja *õige* (ja nende tuletiste) algsed, ruumilist sirgust väljendavad tähendused kattuvad tänapäeva keeles ainult osaliselt. Kõigi nende vastandiks selles tähenduses sobib *kõver*. Kõige selgemalt väljendab kõnealust tähendust – kõvera vastandit – *sirge*. Sirged võivad olla asjad (traat, puu, laud), tee, liikumine. *Õige* on tänapäeva kirjakeeles selle tähenduse praktiliselt kaotanud, aga et see tähendus tal olemas on, näitavad väljendid nagu (*traati*) *õigeks*

*painutama*, verb *õgvendama* jms. Sõna *otse(ne)* kõnealune kasutus on kõige keerulisemalt seotud sirgusega. Asjade kohta seda mainitud tähenduses praktiliselt ei kasutata. Küll võime öelda nt *vaatas otse minu poole*, *läks otse metsa*. Ent ka siin on *sirgega* võrreldes erinevused. Nt kui keegi läheb otse läbi metsa, siis ei tähenda see, et ta läheb täiesti sirgelt. *Otse* puhul lisandub tüüpiliselt eeldus, et liigutakse teatud sihtmärgi poole (lihtne sirge liikumine seda ei eelda), ja *otse* tähendab, et liigutakse sinna **üleliigsete** kõverusteta (nt ümber puude või põõsaste ringimine ei muuda läbi metsa liikumist veel mitteotseks); niisiis mitte tingimata päris sirgelt, aga nii sirgelt kui võimalik. Selles on *otse* ja *sirge* vaadeldavate tähenduste seos. Ja ilmselt seletab see suurel määral asjaolu, et soome keeles kasutatakse mõlema tähenduse väljendamiseks *suora/suoraan*, nagu ka vene keeles *prjamo* ja *prjamoi*.

Ülekannetest abstraktsetele (s.o. mitteruumilistele) aladele on artiklis vaadeldud järgmisi.

**1. Suhtlus.** Seda ülekannet illustreerivad eesti keeles väljendid nagu *ütles otse*, *otsesõnu öeldes*, *otsekohene/sirgjooneline jutt*. Samad tähendusülekanded on olemas soome, inglise ja vene keeles. Tähenduse “ruumiline” taust avaldub eesti keeles eriti selgesti väljendites nagu *ütles otse näkku*, *valetas otse suu sisse*.

**2. Iseloom:** *sirgjooneline*, *otsekohene*, *sirgemeelne* (inimene, iseloom). Vahetuks lähteks on ilmselt eelmine tähendus.

**3. Suhtlemisega seostub** räägitu (ja selle lähteks olevate arvamuste, otstuste jms) **tõelevastavus**. Eesti keeles kasutatakse (vaadeldavatest sõnadest) selle tähenduse väljendamiseks vaid *õiget* ja selle tuletisi: *õige jutt*, *õige otsus*, *õige mõte*, *arvasid õigesti*. Sedasama võib öelda soome keele kohta, kus kõnealust tähendust väljendavad *oikea/oikein*. Ilmneb aga huvitav erinevus inglise keelest, kus Cienki järgi kehtib pigem üldine tees “truth is plain” vrd *plain (and simple) truth*. Tõde on inglise keele järgi niisiis pigem tasapinnaline kui ühedimensionaalne objekt. Ühetasase pinna ja sirge (joone) vahel on küll ilmne pertseptuaalne seos (ja ka eesti keeles võib öelda *silus kortsus paberi sirgeks*), kuid vaadeldavale tähendusosalale see eristus/analoogia ei ole eesti keeles kandunud.

**4. Aeg, sündmused.** Selles sihtalas tähendab sirgus ajas järjestatud sündmuste vahetut järgnevust, ja seda väljendatakse eesti keeles eelkõige *otse* mõnede tuletiste abil: *otsekohe*, *otsemaid*. “Sirgus” on siin justkui tagatud vahelülide puudumisega. Pole juhuslik, et *otse* (aga mitte *sirge* või *õige*) algse ruumilise tähenduse puhul leiame analoogilise

kasutusvariandi: nt *need (laual olevad) saiad on otse ahjust tähendab, et teel ahjust lauale pole saiad mujal olnud.*

**5. Kord ja kontroll.** Cienki järgi on inglise keeles üks *straight* tüüpilisi tähendusi kord/korrasolek/korrastamine: *the room is straight again* 'tuba on taas korras' *to straighten out what someone else did wrong* 'midagi korda ajama/tegema, mis keegi teine valesti tegi' Eesti keeles ei ole leksikaalset seost *otse/sirge/õige* ja nt asjade korrasoleku vahel. Küll aga leiame nt verbi (*ära*) *õiendama*, mis üldisemal "asjaajamise" tasandil kannab tähendust "korda ajama, korrastama" Sii võib lisada väljendid *joonde ajama, jutti ajama*, aga ka (*midagi*) *sirgeks rääkima* jms. Niisiis puht semantiliselt ei ole ka see ülekanne eesti keelele võõras.

Sama võib väita soome keele kohta, kus väärrib tähelepanu nt verb *suorittaa*, millel on selge seos omadussõnaga *suora*. Seda võib kõrvutada eesti keele verbiga *õiendama*, ent esimesel on soome keele leksikaalses süsteemis semantiliselt hoopis üldisem ja kaalukam koht.



ISSN 1406-0183  
ISBN 9985-4-0144-1