

NODALIDA 2007 Tutorial

Typological Singularities of Estonian and Their Role in Computer Analysis

Kadri Muischnek, Kadri Vider,
Haldur Õim,
University of Tartu

Estonian language is typologically...

- Finno-Ugric language
- Mainly agglutinative language
- Mainly SVO language

Readings:

- Mati Erelt (ed.), Estonian Language. – Linguistica Uralica. Supplementary Series, vol. 1. Tallinn: Estonian Academy Publishers, 2003. 413 pp.
- Alo Raun, Andrus Saareste 1965. Introduction to Estonian Linguistics. Ural-Altäische Bibliothek XII. Wiesbaden.
- Wikipedia

Topics covered in this presentation

Phonology

- quantity system

Morphology

- word-classes

- inflection

- compounds

Syntax

- coding of syntactic functions

- word order

Alphabet and orthography

- Latin alphabet with special letters:
š ž ō [ɣ] ä [æ] ö ü
- Essentially phonemic orthography
- Missing graphemes for palatalisation
- Missing graphemes for distinguishing
2nd and 3rd degrees of phoneme length
(so-called syllabic quantity, Q2 and Q3)

Phonology (1): Phonemes

- Vowels (9): a e i o u õ ä ö ü
19 segmental diphthongs
- Consonants (19): k p t t' s s' z z' m n n'
ŋ l l' r f v j h

Phonology (2): Gradation of phonemes

- 1) Qualitative changes
arg:ara (*timid*), tuba:toa (*room*), mesi:mee (*honey*); nd:nn, rs:rr; b:v, d:j g:j
- 2) Quantitative changes – **syllabic quantity**

		Syllable length	Syllable weight
Syllabic quantity	Q1	short	light
	Q2	long	
	Q3		heavy

Morphophonology: the case system (1)

Grammatical cases – rich in stem variation

- **NOM**

(without suffix)

uba lipp hüpe pöial idee

- **GEN**

(stem vowel,
∅ or quantity
changes)

oa lipu hüppe pöidla idee

∅ -u ∅ -a ∅

Q2 Q3 Q3

- **PART**

(suffix –t(d),
∅ or quantity
changes)

uba lippu hüpet pöialt ideed

∅ ∅ -t -t -d

Q3 Q2 Q2

'bean' 'flag' 'jump' 'thumb' 'idea'

Morphology: word classes

1. Words that inflect for mood, time and person - verbs
2. Words that inflect for case (14) – nominals – nouns, adjectives, numerals, pronouns
3. Words that have no inflectional forms – adverbs, adpositions, conjuncts, interjections

Morphology: the case system (2)

System of **adverbial cases** (Erelt et al 2003:33), concrete suffixes, have 'more meaning'

	directional	static	separative
Interior	ILL: -sse <i>into</i>	INE: -s <i>in</i>	ELA: -st <i>from</i>
Exterior	ALL: -le <i>onto</i>	ADE: -l <i>on</i>	ABL: -lt <i>off</i>
Limited	TERM: -ni <i>up to...</i>		
Existential	TRAN: -ks <i>into the state of being</i>	ESS: -na <i>as ...</i>	
Instrumental		COM: -ga <i>with</i>	ABE: -ta <i>without</i>

Morphological ambiguity: grammatical homonymy

auto-Ø

car-NOM

auto-Ø

car-GEN

auto-t

car-PART

Isa-Ø

father-NOM

isa-Ø

father-GEN

isa-Ø

father-PRT

Cmp: Finnish: **isä: isän: isää**

Grammatical homonymy: consequences

Maailm-a *juhti-va-d* *majandus-riigi-d*
world-**GEN** lead-PRS.PTCP-PL economic-state-PL.NOM

'world's leading economic nations'

Maailm-a *juhti-va-d* *majandus-riigi-d*
world-**PART** lead-PRS.PTCP-PL economic-state-PL.NOM

'economic nations leading the world'

Word classes: adpositions

puu *läbi*
tree-GEN through
'because of the tree'

läbi *puu*
through tree-GEN
'through the tree'

Word classes: adpositions

Tera ***keerle-b*** ***puu*** ***üumber***
Blade-NOM revolve-3.SG tree-GEN around(ADPOS)

kindla-t ***joon-t*** ***mööda.***
certain-PART line-PART along(ADPOS)

'The blade is revolving around the tree along a certain line.'

Tera ***keerle-b*** ***üumber*** ***puu***
Blade-NOM revolve-3.SG around(ADPOS) tree-GEN

mööda ***kindla-t*** ***joon-t.***
along(ADPOS) certain-PART line-PART

'The blade is revolving around the tree along a certain line.'

Adverbial particle vs adposition

<i>Ta</i>	<i>ela-s</i>	<i>üle</i>	<i>vapustus-e.</i>
s/he	live-PST.3SG	over(PARTICLE)	shock-GEN
[SUBJ]	[+FMV]	VPART]	[OBJ]

's/he experienced a shock'

<i>Ta</i>	<i>ela-s</i>	<i>üle</i>	<i>tänav-a.</i>
s/he	live-PST.3SG	over(ADPOS)	street-GEN
[SUBJ]	[+FMV]	[[PREP	<NOUN]=ADVL Loc]

's/he lived across the street'

Compounding

Sup-i+söömine

soup-GEN+eating

puhta-ks+pühkimine

clean-TRANSL+sweeping

selle lamba liha

this-GEN sheep-GEN meat-NOM

'the meat of this sheep'

see lamba+liha

this-NOM sheep-GEN+meat-NOM

'this mutton'

Compound splitting

kuusetukad

1) *kuuse/tukad*

kuusk-GEN | clump+PL.NOM

‘fir clumps’

2) *kuu/setukad*

moon-GEN | crock-PL.NOM

‘moon crocks’

Verb expressions

verb + adverb particle = particle verb

Kukku-s ***alla*** ‘fell off’

fall-PST.3.SG down

Kukku-s ***kokku*** ‘collapsed’

fall-PST.3.SG together

Verb expressions (2)

verb + noun/adjective phrase =
verbal idiom (expression verb)

Saa-b *aru* ‘understands’

get-3.SG sense-PART

Verb expressions in a clause

Ta sa-i luuletuse-st aru.
s/he get-PST.3SG poem-**ELA** sense-**PART**
[SUBJ] [+FMV [**ADV L Obj**] **V PART**]
's/he understood the poem'

Ta sa-i poe-st piima.
s/he get-PST.3SG shop-**ELA** milk-**PART**
[SUBJ] [+FMV] [**ADV L Locfrom**] [**OBJ**]
's/he got some milk from the shop'

Syntax (1)

Coding of syntactic functions: mostly by morphological devices

IDEAL:

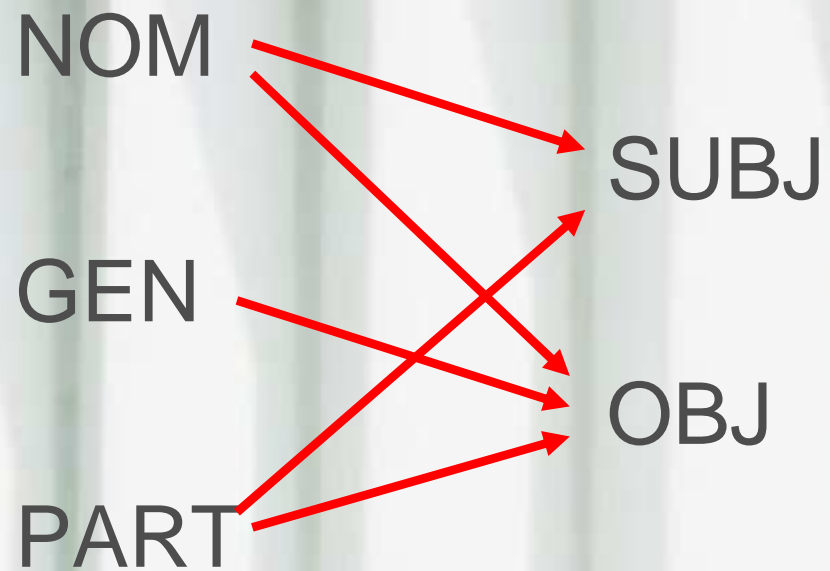
NOM → SUBJ

GEN → ATR

PART → OBJ

Syntax (2)

Coding of syntactic functions: reality



Syntax: word order

Basic: SVO

Laps *söüb* *saia*
child-NOM eat-3SG bun-PART
[SUBJ] [+FMV] [OBJ]
'The child is eating the bun'

But

Laps saia söüb.
Saia söüb laps.
Söüb saia laps.

Saia laps söüb.
Söüb laps saia.

NP in PL.NOM: subject or object ?

Eesti rahu-valvaja-d vaheta-si-d Bosnia-s välja

Estonian peace-keeper-PL replace-PST.3PL Bosnia-INE out(PARTICLE)

[**NOUN** > **SUBJ**] [+FMV [ADVL Loc] <ADVL]

lätlase-d.

Latvian-PL

[**OBJ**]

'The Estonian peace-keepers replaced Latvians in Bosnia'

[**NOUN** > **OBJ**] [+FMV [ADVL Loc] <ADVL]

[**SUBJ**]

'Latvians replaced Estonian peace-keepers in Bosnia'

N in SG.GEN: object or premodifier?

Ta aseta-s mantl-i tool-i seljatoe-le.

S/he place-PST.3SG coat-**GEN** chair-**GEN** back-ALL
[SUBJ] [+FMV] [OBJ] [NOUN> ADVL Locto]

's/he placed the coat on the back of the chair'

[SUBJ] [+FMV] [NOUN> OBJ] [ADVL Locto]

's/he placed the chair of the coat on the back'

N in adverbial cases: adverbial or premodifier?

<i>Ta</i>	<i>võtt-i-s</i>	<i>prae-ahju-st</i>	<i>pann-i</i>
s/he	take-PST.3SG	owen- ELA	pan- GEN
[SUBJ]	[+FMV]	[ADVL Locfrom]	[OBJ]

's/he took the pan from the owen'

<i>Ta</i>	<i>võtt-i-s</i>	<i>vase-st</i>	<i>pot-i.</i>
s/he	take-PST.3SG	copper- ELA	pot- GEN
[SUBJ]	[+FMV]	[NOUN>	OBJ]

's/he took the copper pot'

Object cases and aspect

Mees ehita-s suvila-t

Man-NOM built summer-house-**PART**

‘The man built a summer-house/did some summer-house-building.’

Mees ehita-s suvila

Man-NOM built summer-house-**GEN**

‘The man built a summer-house.’

Object cases and aspect (2)

Jaan sõ-i sup-i ära.

Jaan-NOM eat-PST.3.SG soup-**GEN** up

‘Jaan ate up the soup’

Finnish: Jaan söi keiton.

Jaan luge-s raamat-u läbi.

Jaan-NOM read-PST.3.SG book-**GEN** through

‘Jaan read through the book’

Impersonal and/or passive

- no subjective action passive
- Impersonal voice marked on the verb:

siin ela-ta-kse

here live-IPS

'smn lives here'

loe-ta-kse

read-IPS

'smn reads a book'

raamatu-t

book-PART

Impersonal vs passive

	<i>Impersonal</i>	<i>Personal passive</i>
Present	<i>Puud raiuta-kse maha.</i> 'One is felling the trees'	<i>Puud on maha raiutud.</i> 'The trees are felled.'
Preterite	<i>Puud raiut-i maha.</i> 'One was felling the trees'	<i>Puud olid maha raiutud.</i> 'Trees were felled.'
Pres. perf.	<i>Puud on maha raiutud.</i> 'One has been felling the trees'	<i>Puud on maha raiutud olnud.</i> 'trees have been felled.'
Past perf.	<i>Puud oli maha raiutud</i>	<i>Puud olid maha raiutud olnud.</i>

Impersonal vs passive (2)

Puud on maha raiutud.

Tree-PL is-3.SG down fell-PST.PCTP

OBJ [Verbal chain]

'One has been felling the trees'

SUBJ V [PRD]

'The trees are felled.'

Estonian vs Standard Average European: the possessor

Mu-l on koer.
I-ADE be/have-3.SG dog-NOM
[**ADVL Poss**] [+FMV] [**SUBJ**]
'I have a dog'

Ma oma-n koer-a.
I-NOM own-1.SG dog-PART
[**SUBJ**] [+FMV] [**OBJ**]
'I have a dog'

Estonian vs Standard Average European: the experienter

Mu-lle

I-ALL

[**ADVL Exper**]

'I like dogs.'

meeldi-va-d koera-d.

like-3.PL

dog-PL

[+FMV] [**SUBJ**]

Ma armasta-n koer-i.

I-NOM

love-1.SG

dog-PL.PART

[**SUBJ**] [+FMV] [**OBJ**]

'I love dogs.'

References

- Mati Ereht (ed.), Estonian Language. – Linguistica Uralica. Supplementary Series, vol. 1. Tallinn: Estonian Academy Publishers, 2003. 413 pp.
- K. Müürisep. Eesti keele arvutigrammatika: süntaks. Dissertationes Mathematicae Universitatis Tartuensis 22. Tartu 2000.
- Roosmaa, T., Koit, M., Muischnek, K., Müürisep, K., Puolakainen, T, Uibo, H. 2001. Eesti keele formaalne grammatika. Tartu Ülikool, Arvutiteaduse Instituut, 158 pp.
- Tiina Puolakainen. Eesti keele arvutigrammatika: morfoloogiline ühestamine. Dissertationes mathematicae Universitatis Tartuensis, 27. Tartu, 2001
- http://en.wikipedia.org/wiki/Estonian_language