

Gradient constraints on the use of Estonian possessive reflexives

Suzanne Lesage Olivier Bonami

Université de Paris, LLF, CNRS

Quantitative Syntax 2019, Paris, August 26

Introduction I

Two ways of referring to the possessor of a noun:

- ▶ adnominal genitive pronouns (1) → NONREFLEXIVE POSSESSIVES,
- ▶ a reflexive forms *oma* (2) → REFLEXIVE POSSESSIVES.

- (1) a. Peeter vii-s mind;
Peeter.NOM lead-PST 1SG.PART minu_i vanema-te juurde.
1SG.GEN parent-PL.GEN at
'Peeter led me to my parents' place.'
- b. Peeter_i vii-s Jaani;
Peeter.NOM lead-PST JaanSG.PART tema_{j/*i} vanema-te juurde.
3SG.GEN parent-PL.GEN at
'Peeter_i led Jaan_j to his_{j/*i} parents' place.'
- (2) a. Ma_i vii-si-n Jaani;
1SG.NOM lead-PST-1SG Jaan.GEN oma_{i/*j} vanema-te juurde.
POSS.REFL parent-PL.GEN at
'I led Jaan to my parents' place.'

Introduction II

- b. Peeter_i vii-s Jaan_j oma_{i/*j} vanema-te juurde.
Peeter.NOM lead-PST Jaan.sg.PART 3SG.GEN parent-PL.GEN at
'Peeter_i led Jaan_j to his_{i/*j} parents' place.'

In canonical constructions,

- ▶ reflexive possessives → subject (2),
- ▶ nonreflexives → local non-subject, as in (1),
→ or be locally unbound.

Introduction, psych verbs I

For psych verbs expressing, both reflexive and nonreflexive possessives → subject or → allative argument.

- (3) a. Mu-lle meeldi-vad kassi-d_i oma_i/nende_i iseloomu
1SG-ALL please-3PL.PRS cat-PL.NOM POSS.REFL/3PL.GEN temper.GEN
pärast.
because
'I like cats because of their temper.'
- b. Lille-de-le_i sobi-b taeva-st alla sadanud
flower-PL-ALL be.suitable-3SG.PRS sky-ELA. adown fallen
vesi oma_i/nende_i pehmuse tõttu väga hästi.
water.NOM POSS.REFL/3PL.GEN fragility because very well
'Rain water is suitable to flowers because of their fragility.'

Introduction, psych verbs II

Two psycholinguistic experiments (Lesage, accepted) :

- ▶ No categorical preference for reflexives being bound by the surface subject (resp. nonreflexives being bound by the allative argument);
- ▶ Binding preferences modulated by word order, with reflexives showing a preference for an initial antecedent irrespective of its grammatical function .

In this talk, results from comprehension experiments confirmed in production?

Data collection and annotation

- ▶ Combination of factors too rare for data to be easily available (no articles, no mandatory overt expression of possession, noncanonical construction found only with a handful of verbs).
- ▶ We relied on resources from the Universal Dependencies community to parse a large web corpus and use it for initial data selection. We trained UDPipe (Straka and Straková, 2017) on the Estonian UD v2.4 treebank, the Universal Dependencies version of the Estonian Dependency Treebank (Muischnek et al., 2014). We then used this to parse the 1.1 billion token Estonian National Corpus (Kallas and Koppel, 2018).

Data selection I

Morphological, POS and dependency annotation to select all sentences containing :

- ▶ a token v of one of eight verbs taking an allative argument: *meeldima* 'please', *sobima* or *kõlbama* 'be suitable for', *meenuma* 'come to one's mind', *võimaldama* 'make possible', *kuuluma* 'belong', *jätkuma* 'be enough', *maitsuma* 'please by its taste';
- ▶ a token p of a reflexive or nonreflexive possessive.
- ▶ The possessive p is the possessor of some noun that has v on its head path.
- ▶ The verb v has an allative dependent.
- ▶ p 's person-number features compatible with the person-number features expressed either on the verb, the subject, or the allative dependent.
- ▶ → 5,593 candidate examples
- ▶ examples sorted by hand to eliminate the numerous false positives → 1,307 sentences.

Data selection II

- ▶ 5 groups according to the grammatical function of possessives' head noun:

Type of relation	Count
Surface subject	415
Allative argument	285
Direct object	86
Other oblique dependent	366
Embedded within a dependent	155

'surface subject' → argument which may trigger agreement on the verb.
Direct objects rare in our corpus, (*meenutama* 'remind' and *võimaldama* 'make possible' takes a direct object).

Annotation

Each example was annotated (information collected from dependency parses + manual work):

- ▶ The type of possessive (reflexive or nonreflexive).
- ▶ The grammatical function of the antecedent (surface subject or allative argument).
- ▶ The grammatical function of the possessed noun.
- ▶ The person and number, and animacy of each argument.
- ▶ The volitional involvement of the subject.
- ▶ The relative order of the two arguments + the possessive and its antecedent.

Results

- ▶ Mixed effects logit models, using the *lme4* and *lmerTest* (Kuznetsova, Brockhoff, and Christensen, 2017) R packages.
- ▶ Dependent variable : the type of possessive
Fixed effect : identity of the verb
- ▶ Two important independent variables : function of the possessed noun and that of the antecedent.

Antecedent	Possessed noun		
	nominative	allative	other oblique
nominative	*	all→subj	other→subj
allative	subj→all	*	other→all

Grammatical functions

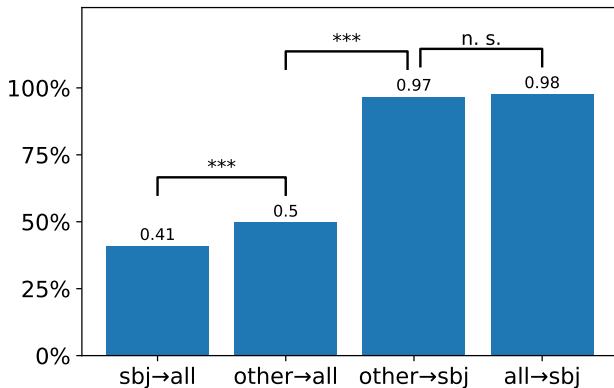


Figure: Proportion of use of the reflexive possessive (vs. nonreflexive possessive) for each combination of grammatical functions of the possessed noun and antecedent.

Grammatical functions

- ▶ Higher preference for reflexive possessives for nominative antecedent than for allative antecedents.
- ▶ Gradient reflex of the well-known observation that relative obliqueness constrains binding (Pollard and Sag, 1992): reflexive dependents of a verb tend to be used when they are bound by a less oblique dependent.

(4) Subject \prec Direct object \prec Oblique argument \prec Adjunct

- ▶ Differences with allative antecedents ($sbj \rightarrow all$ vs. $other \rightarrow all$) are more subtle.
- ▶ In the $sbj \rightarrow all$ condition, we are departing maximally from that situation : not only is the antecedent not a subject, it is also strongly *more oblique* than the possessed noun. \rightarrow nonreflexive \gg reflexive possessive
- ▶ In the $other \rightarrow all$ condition, the antecedent is more oblique than the possessed noun.
Only the strongest expectation that the antecedent be a subject, but not the weaker expectation that it be less oblique than the possessed noun, is violated in the $other \rightarrow all$ condition.

Relative order of antecedent and possessive

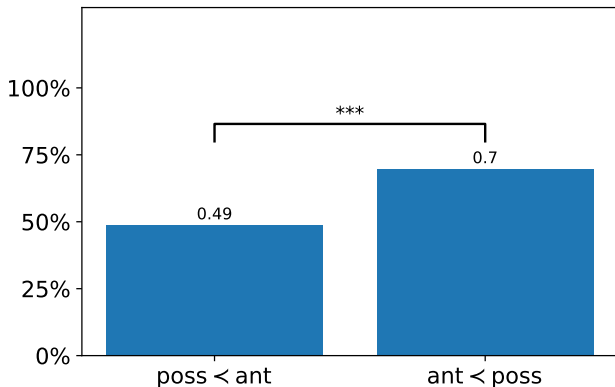


Figure: Proportion of use of the reflexive possessive (vs. nonreflexive possessive) for each word order (poss<ant vs. ant<poss).

Word order I

- ▶ Significant effect of word order : possessives preceding their antecedents favor more a reflexive than possessives that follow their antecedent.
- ▶ Information structure (tight link between word order and information structure in Estonian) realized first in linear order strongly tends to be topical (Lindström, 2005; Tael, 1988).
- ▶ Bickel, 2004 suggests that reflexives can also be topic-oriented rather than subject-oriented: the reflexive tends to take the topic as its antecedent, which will coincide with the subject in most situations, but is likely not to in sentences with an experiencer expressed as an oblique.
- ▶ In Estonian, reflexives are *both* subject-oriented and topic-oriented.
- ▶ Where the antecedent is an allative, obliqueness and topicality pose conflicting constraints on the choice of the possessive form: the obliqueness relation between possessive and antecedent favors a nonreflexive, while the topicality of the antecedent favors a reflexive.
- ▶ Nonreflexives are more common, reflexives are still a relevant option in most cases where the antecedent is the allative argument.

Person of antecedent (co-argument case) I

Optimizing speech for ambiguity avoidance (Grice, 1975).



- (5) a. ... Mulle meeldib **minu** amet.
1SG.ALL pleases 1SG.GEN job.NOM
'I like my job.'

- b. ... Mulle meeldi-b **oma** amet.
1SG.ALL pleases POSS.REFL job.NOM
'I like my job.'
- Diagram: A green arrow connects 'oma' to 'Mulle'. A red arrow connects 'oma' to 'meeldi-b', with three red 'x' marks above it. Another red arrow connects 'oma' to 'Mulle', also with three red 'x' marks above it.

- c. ... Peetrite meeldib **tema** amet.
Peeter.ALL pleases 3SG.GEN job.NOM
'Peeter likes his job.'
- Diagram: A green arrow connects 'tema' to 'Peetrite'. Another green arrow connects 'tema' to 'meeldib'.

- d. ... Peetrite meeldib **oma** amet.
Peeter.ALL pleases POSS.REFL amet.NOM
'Peeter likes his job.'
- Diagram: A green arrow connects 'oma' to 'Peetrite'. A red arrow connects 'oma' to 'meeldib', with three red 'x' marks above it. Another red arrow connects 'oma' to 'Peetrite', also with three red 'x' marks above it.

ambiguous!

Person of antecedent (co-argument case) II

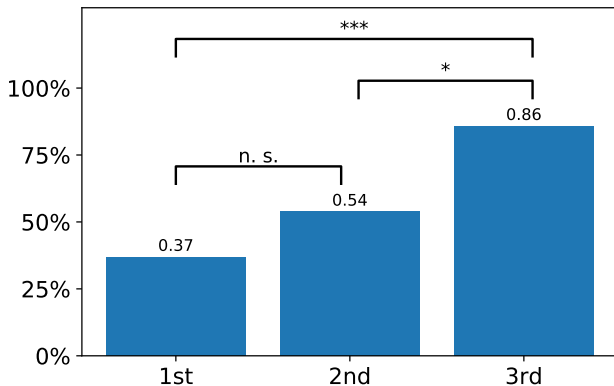


Figure: Proportion of use of the reflexive possessive (vs. nonreflexive possessive) for each antecedent's person.

Conclusion

Our main empirical findings :

- ▶ noncategorical constraints on the choice of possessive,
- ▶ documenting influences of relative obliqueness, word order, and person.

These results are in line with previous observations in comprehension experiments. We explored two separate but complementary lines of explanation for these findings: an interplay of grammatical relations and information structure on the one hand, and an influence of ambiguity avoidance.

References I

- Bickel, Balthasar (2004). "The syntax of experiencers in the Himalayas." In: *Non-nominative Subjects*. Ed. by Peri Bhaskararao and Karumuri V. Subbarao. Vol. 1. Amsterdam: John Benjamins, pp. 77–112 (cit. on p. 14).
- Grice, Paul (1975). "Logic and conversation." In: *The logic of grammar*. Ed. by Donald Davidson and Gilbert H. Harman. Ensino: Dickenson (cit. on p. 15).
- Kallas, Jelena and Kristina. Koppel (2018). "Eesti keele ühendkorpus 2017." Center of Estonian Language Resources (cit. on p. 6).
- Kuznetsova, Alexandra, Per B Brockhoff, and Rune Haubo Bojesen Christensen (2017). "lmerTest package: tests in linear mixed effects models." In: *Journal of Statistical Software* 82.13 (cit. on p. 10).
- Lesage, Suzanne (accepted). "Liage du réfléchi possessif en estonien : une approche expérimentale." In: *Études finno-ougriennes* (cit. on p. 5).
- Lindström, Liina (2005). *Finiitverbi asend lauses: sõnajärg ja seda mõjutavad tegurid suulises eesti keeles*. Vol. 16. Tartu Ülikooli kirjastus (cit. on p. 14).
- Metslang, Helena (2013). "Coding and behaviour of Estonian subjects." In: *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics* 4.2, pp. 217–293.
- Muischnek, Kadri et al. (2014). "Estonian Dependency Treebank and its annotation scheme." In: *Proceedings of the 13th Workshop on Treebanks and Linguistic Theories* (cit. on p. 6).
- Pollard, Carl and Ivan A. Sag (1992). "Anaphors in English and the Scope of Binding Theory." In: *Linguistic Inquiry*, pp. 261–303 (cit. on p. 12).

References II

- Straka, Milan and Jana Straková (2017). “Tokenizing, POS Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe.”
In: *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*.
Vancouver, Canada: Association for Computational Linguistics, pp. 88–99. URL:
<http://www.aclweb.org/anthology/K/K17/K17-3009.pdf> (cit. on p. 6).
- Tael, Kaja (1988). “Infostruktuur ja lauseliigendus.” In: *Keel ja Kirjandus*, pp. 133–143 (cit. on p. 14).

Appendix: model parameters

	Estimate	Std. Error	z value	p-value	
(Intercept)	0.005898	0.582356	0.010	0.991919	
subj→all vs.other→all	-1.181075	0.274058	-4.310	1.64e-05	***
other→subj vs.all→subj	-2.271744	0.426328	-5.329	9.90e-08	***
other→subj vs.all→subj	-0.612193	1.331760	-0.460	0.645741	
person=1	-2.230623	0.296334	-7.527	5.18e-14	***
person=2	-1.308191	0.485689	-2.693	0.007071	**
order=ant_first	0.998497	0.262998	3.797	0.000147	***

Table: Parameters of GLMM modelling the proportion of use of a reflexive possessive.

The role of ambiguity I

In other→all and other→sbj conditions, things are less clear, both conceptually and empirically.

Person of antecedent	Person of other arg.	Poss. type	Local amb.	Global amb.	# of tokens	% of reflexives
-3rd	-3rd	-refl.	no	no	4	33%
-3rd	-3rd	refl.	yes	no	2	
-3rd	3rd	-refl.	no	no	43	51%
-3rd	3rd	refl.	yes	no	45	
3rd	-3rd	-refl.	no	yes	4	95%
3rd	-3rd	refl.	yes	no	79	
3rd	3rd	-refl.	(yes)	yes	51	81%
3rd	3rd	refl.	yes	no	224	

Table: Potential ambiguity of possessives in sentences with two candidate antecedents.

local ambiguity = ambiguity with a clause-local antecedent

global ambiguity = ambiguity with an extra-sentential antecedent

the role of ambiguity

Prediction : Reflexives are

- ▶ rarest where both arguments are nonthird person (using a nonreflexive avoids local ambiguity)
- ▶ most frequent where both arguments are third person (using a reflexive avoids global ambiguity).
- ▶ In the other two situations, it is unclear (in the absence of a hypothesis on the relative costs of local and global ambiguity).

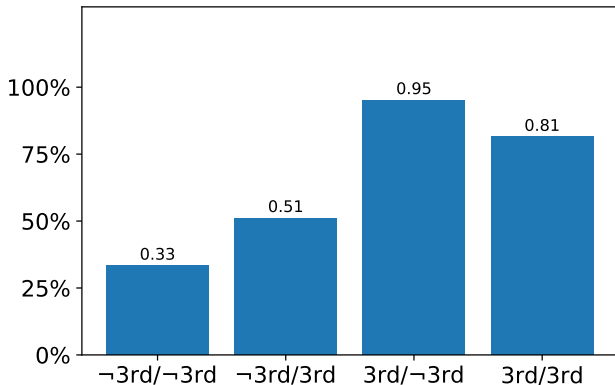


Figure: Proportion of use of the reflexive possessive (vs. nonreflexive possessive) for

phrases pour illustrer le tableau I

- (6) ... Mulle ei sobi sina **minu** religiooni tõttu.
1SG.ALL NEG suit 2SG.NOM 1SG.GEN religion.GEN because
'You are not suitable for me because of my religion.'



- (7) ... Mulle ei sobi sina **oma** religiooni tõttu.
1SG.ALL NEG suit 2SG.NOM POSS.REFL religion.GEN because
'You are not suitable for me because of my religion.'

- (8) ... Mulle ei sobi Peeter **minu** religiooni tõttu.
1SG.ALL NEG suit Peeter.NOM 1SG.GEN religion.GEN because
'Peeter is not suitable for me because of my religion.'



- (9) ... Mulle ei sobi Peeter **oma** religiooni tõttu.
1SG.ALL NEG suit Peeter.NOM POSS.REFL religion.GEN because
'Peeter is not suitable for me because of my religion.'

phrases pour illustrer le tableau II

- (10) ... Maarjale ei sobi sina **tema** religiooni tõttu.
Maarja.ALL NEG suit 2SG.NOM 3SG.GEN religion.GEN because
'You are not suitable for Maarja because of her religion.'
- (11) ... Maarjale ei sobi sina **oma** religiooni tõttu.
Maarja.ALL NEG suit 2SG.NOM POSS.REFL religion.GEN because
'You are not suitable for Maarja because of her religion.'
- (12) ... Jaanile ei sobi Peeter **tema** religiooni tõttu.
Jaan.ALL NEG suit Peeter.NOM 3SG.GEN religion.GEN because
'Peeter is not suitable for Jaan because of his religion.'
- (13) ... Jaanile ei sobi Peeter **oma** religiooni tõttu.
Jaan.ALL NEG suit Peeter.NOM POSS.REFL religion.GEN because
'Peeter is not suitable for Jaan because of his religion.'
-