The powerful mix of animacy and initial position: 
A psycholinguistic investigation of Estonian case and subjecthood

Elsi Kaiser*, Merilin Miljan†§, Virve-Anneli Vihman§‡

*Univ. of Southern California, †Univ. of Tallinn, ‡Univ. of Tartu, §Univ. of Manchester
elsi.kaiser@gmail.com, merilinmiljan@gmail.com, virve.vihman@gmail.com

Forty Years after Keenan 1976, 7–9 September 2016
Ghent University, Ghent, Belgium

When defining grammatical relations, we often prioritize morphological and syntactic properties, such as case-marking, agreement, linear position (see, e.g., Keenan 1976, Keenan and Comrie 1977) over meaning-related criteria (e.g. animacy).

- Subject in nominative-accusative languages is typically in nominative case.
- Pre-verbal/clause-initial noun phrase is the subject (e.g. English)

Estonian, a nominative-accusative language, presents an interesting example:

- three different morphological cases – NOM, PART, GEN – are used for marking the core arguments, subject and object:
  - Subject is marked by two of these: NOM and PART (section 1)
  - Object is marked by NOM, PART and GEN (section 1.1)
- flexible word order, i.e. pre-verbal position does not define subjecthood (section 1.2)

⇒ all three core cases are syntactically ambiguous, while the constituent order is free

Questions: How do the (prototypical) morphological and syntactic properties of subjects – case marking and linear position – cue for the grammatical role of subject in Estonian?

- What information influences the interpretation of the subject relation in Estonian?
- Is case in Estonian a stronger subjecthood feature than word order, because in terms of frequency, the position of the subject varies?

This talk: We report on two psycholinguistic experiments: participants were shown case-marked nouns and asked to use them in a sentence. In the first experiment, the sentence had to begin with the prompt noun; the second experiment placed no constraints on linear order.

Experiment 1: What grammatical relations are assigned to the core cases on nouns in sentence-initial position during incremental processing, before any other linguistic information is encountered? How strong a cue is initial position (which is statistically associated with S) for subject interpretation, in a situation where case-marking does not force a subject interpretation?

Experiment 2: What syntactic information is associated with case marking on nouns when sentence position is flexible? If a noun is not restricted to clause-initial position, how strongly are NOM, PART and GEN associated with particular grammatical roles?
AIM: To better understand what factors influence the likelihood of a noun being interpreted as the subject in a flexible word-order language.

1. Subject marking in Estonian

- Nominative (NOM, ex.1) and partitive case (PART, ex.2) mark subjects in Estonian. This is an instance of differential case marking on the subject with intransitive (activity) verbs.
- Partitive case can mark subjects on plural count nouns (2a) and on mass nouns (2b); singular count nouns are unacceptable under partitive marking (2c).
- The nominative in (1) gives rise to a perfective reading and the partitive in (2) to an imperfective or partially complete/indefinite quantity reading.

(1)  
a. **Nominative singular subject (count noun)**
    Külaline saabus.
guest.NOM.SG arrive.PST.3SG  
‘(The) guest (has) arrived.’

b. **Nominative plural subject (count noun)**
    Külalised saabusid.
guest.NOM.PL arrive.PST.3PL  
‘(The) guests (have) arrived.’

(2)  
a. **Partitive plural subject (count noun)**
    Külalisi saabus.
guest.PART.PL arrive.PST.3SG  
‘(Some) guests arrived.’ / ‘Guests were arriving.’

b. **Partitive singular subject (mass noun)**
    Uut sõjatehnikat saabus täna Tapale.
new.PART.SG war.equipment.PART.SG arrive.PST.3SG today Tapa.ADE  
‘(Some) new war equipment arrived today to Tapa.’

c. **Partitive singular subject (* count noun)**
    * Külalist saabus.
guest.PART.SG arrive.PST.3SG  
intended reading: ‘(Part of) a guest arrived.’

1.1 The “subject” cases also mark objects

- Nominative and partitive both also mark objects.
- Partitive marks both singular (3a) and plural objects (3b).
- Partitive alternates with genitive on singular objects (4a), and with nominative on plural objects (4b). That is, object case-marking crucially involves number: singular objects in unmarked, active voice constructions take genitive case, whereas nominative case marks plural, affected objects (4c). This is sometimes analysed as accusative.¹

¹ There is no morphological evidence for a morphologically distinct accusative case in Estonian, and its existence remains a matter of debate (see, e.g. Hiiyetam 2003, Caha 2009, Miljan and Cann 2013, Norris 2015).
• The alternation between partitive and genitive/nominative signifies a difference in boundedness, either quantity of the affected object or grammatical aspect (perfectivity).
• (Objects in the scope of negation must be in partitive case and do not show the genitive/partitive alternation.)

(3)  
a. Partitive singular object  
\[\text{Poiss} \quad \text{sõi} \quad \text{kooki}.\]  
\(\text{boy.NOM.SG} \quad \text{eat.PST.3SG} \quad \text{cake.PART.SG}\)  
‘The boy was eating (a) cake.’ / ‘The boy ate (some) cake.’

b. Partitive plural object  
\[\text{Juhan} \quad \text{kirjutas} \quad \text{luuletusi}.\]  
\(\text{Juhan.NOM} \quad \text{write.PST.3SG} \quad \text{poem.PART.PL}\)  
‘Juhan wrote poems.’ / ‘Juhan was writing (the) poems.’

(4)  
a. Genitive singular object  
\[\text{Poiss} \quad \text{sõi} \quad \text{koogi}.\]  
\(\text{boy.NOM.SG} \quad \text{eat.PST.3SG} \quad \text{cake.GEN.SG}\)  
‘The boy ate (the) cake.’

b. Nominative on plural object  
\[\text{Juhan} \quad \text{kirjutas} \quad \text{luuletused}.\]  
\(\text{Juhan.NOM} \quad \text{write.PST.3SG} \quad \text{poem.NOM.PL}\)  
‘Juhan wrote (the) poems.’

c. \text{Poiss kirjutas luuletuse / luuletused.}  
\(\text{boy.NOM.SG write.PST.3SG poem.GEN.SG / poem.NOM.PL}\)  
‘(The) boy wrote (a) poem / poems.’

• Note that case-marking in Estonian can result in transitive sentences with both the subject and object in nominative case, as demonstrated by (4b).
• Nominative also marks the object in imperatives (5) and impersonals (6).

(5)  
Nominative object in imperative  
\[\text{Vii} \quad \text{koer} / \text{koerad} \quad \text{jalutama!}\]  
\(\text{take.IMP.SG} \quad \text{dog.NOM.SG/NOM.PL} \quad \text{walk.INFINITIVE}\)  
‘Take (the) dog/dogs for a walk!’

(6)  
Nominative object in impersonal  
\[\text{Koer} / \text{koerad} \quad \text{viidi} \quad \text{jalutama}.\]  
\(\text{dog.NOM.SG/NOM.PL} \quad \text{take.PST.IMPERS.} \quad \text{walk.INFINITIVE}\)  
‘(The) dog/dogs was/were taken for a walk.’

---

**Nominative case in Estonian is not restricted to subjects**

**Subjects** can be nominative or partitive  
**Objects** can be nominative, partitive or genitive
1.2 Free constituent order

• Estonian has flexible word order, i.e. the object and other grammatical roles may freely occur in sentence-initial position, as in (6) and (7).
• In a transitive sentence, all six possible configurations of S, V and O are grammatical in appropriate discourse-pragmatic contexts.

(7) Sõbra kutsus ta kinno.
friend.GEN.SG invite.3SG.PST 3SG.NOM cinema.SG.ILL
‘S/he invited a friend to the cinema.’

• Subjects may also occur sentence-finally, as in (8):

(8) Talveks tulevad tuppa hiired.
winter.SG.TRA come.3PL.PRS room.SG.ILL mouse.NOM.PL
‘For winter, mice come into the room.’

• Subjects are more frequent sentence-initially than other grammatical relations. For example, Lindström (2004) reports that in spoken Estonian, 41% of all sentences have SV(X) order.

1.3 Summary

• Nominative and partitive mark both subjects and objects; the details depend on number and syntactic construction.
• Genitive marks singular objects, among many other functions (e.g. possessor, adverbials, complements of postpositions).
• Ambiguous morphological and syntactic coding properties for subjects in Estonian: (i) syntactically ambiguous case marking, and (ii) flexible constituent order.

2. Psycholinguistic experiments\(^2\)

• The ambiguous coding properties described above open up interesting questions, especially when combined with the psycholinguistic finding that sentences are interpreted incrementally (e.g. Tanenhaus et al. 1995):
  - Readers/hearers begin building an interpretation immediately, even when the sentence is still unfolding and the input is ambiguous.
  - Readers/hearers rapidly use information such as case-marking to guide the possible interpretations they consider/activate during incremental processing.

1. What kinds of expectations are triggered in Estonian speakers’ minds about grammatical roles when encountering case-marked nouns?
2. How relevant are semantic criteria for assigning grammatical roles to case-marked nouns?

\(^2\) Earlier versions of some of this data have been presented elsewhere, see Kaiser, Miljan and Vihman (2015, under review).
2.1 Subjecthood and animacy

Psycholinguistic work on other languages:
- Speakers tend to mention animate nouns before inanimate nouns (e.g. Bock 1986, Ferreira 1994, see also Prat-Sala & Branigan 2000).
- The interaction between morphological case and animacy in determining grammatical role assignment is not absolute even in a single language: it varies according to the case information available in a particular construction (Bornkessel-Schlesewsky & Schlesewsky 2008):
  - In German constructions with NOM and ACC arguments, animacy has no effect.
  - In German constructions with NOM and DAT arguments (either could be subject or object), animate entities are more preferred as subject than inanimate ones
    \[ \Rightarrow \text{a favored linearization: animate-before-inanimate} \]
- Animacy has not been investigated as a relevant factor in grammatical role assignment in Estonian; neither has animacy been invoked in any description of case assignment rules in Estonian

2.2 Subjecthood and initial position

Psycholinguistic work on other languages:
- ‘Subject-first’ bias: Comprehenders tend to interpret sentence-initial nouns as subjects, regardless of animacy (e.g. Demiral et al. 2008). Universal (?) language comprehension strategy: Subject preference (e.g. Bornkessel & Schlesewsky 2006):
  
  Human language comprehension system tends to assign minimal (grammatical) relation – subject (S/A) – to initial base-form noun phrases or unmarked nominative nouns which are syntactically ambiguous (Bornkessel & Schlesewsky 2006:790).

We conducted 2 sentence-completion experiments:
- Experiments 1 and 2 used the same nouns as prompts.
- 1 & 2 differed in terms of the constraints imposed on clause position of the prompt noun.

**Experiment 1:**
The nouns were given as sentence-initial (e.g. Sõbra....) and participants were explicitly instructed to begin their sentences with the prompt noun.

**Aim:** To mimic the incrementality of real-time language processing in a minimally-constraining context, in order to see how case-marking, animacy and number influence comprehenders’ choices about what grammatical role to assign the sentence-initial nouns.
3. Experiment 1: Interpretation of case-marking on sentence-initial nouns

Main aims of experiment

- #1 To investigate what grammatical role is associated with the noun marked with NOM, PART and GEN in sentence-initial position.

- #2 What factors influence the interpretation of case-marked nouns?
  - Animacy
  - Number (sing/plural)
  - Mass/count [focus here on count only]

- Number is relevant because the case-marking system of Estonian shows sensitivity to number: certain syntactic functions are not available for both singular and plural nouns (section 1).

Why sentence-initial nouns?

- No verb information or clause structure information is present, i.e. that does not limit/guide interpretation of case.
- We designed our experiment so that the clause structure or information from the predicate would not determine case interpretations or grammatical role. All the overt cases used in the experiment could in principle be used in a range of syntactic functions.

Experiment 2:

Participants were instructed to write a sentence using the prompt noun *anywhere in the sentence*.

Aim: To look at the interpretation of case-marking in a flexible context that can tell us about how people interpret case-marking without word order constraints. We again manipulated case marking, animacy and number.
3.1 Method, design, participants

- **Method:** Sentence completion
- Production task: Write a complete sentence beginning with the given noun.

(10) (a) *Hiirt...* (mouse-PART.SG)
(b) *Porgand...* (carrot-NOM.SG)
(c) *Raamatute...* (book-GEN.PL)
(d) *Rebase...* (fox-GEN.SG)

- This method taps into grammatical-role expectations triggered in participants’ minds by information from case-marked nouns (when no information is available about verb/clause structure)
- 42 native Estonian participants, web-based experiment.

- **Design/materials**
  - Manipulated: (i) animacy (animate/inanimate), (ii) number (singular/plural), (iii) count/mass, (iv) case-marking (NOM, PART, GEN)

(11) example: *hiir* (mouse):

<table>
<thead>
<tr>
<th>Case</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOM.SG</td>
<td>– hiir</td>
</tr>
<tr>
<td>PART.SG</td>
<td>– hiirt</td>
</tr>
<tr>
<td>GEN.SG</td>
<td>– hiire</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOM.PL</td>
<td>– hiired</td>
</tr>
<tr>
<td>PART.PL</td>
<td>– hiiri</td>
</tr>
<tr>
<td>GEN.PL</td>
<td>– hiirte</td>
</tr>
</tbody>
</table>

- Latin-square design: Each participant saw each specific noun (e.g. ‘mouse’) only once, but saw equal numbers of nouns in all 6 conditions
- 18 target nouns, 32 fillers (range of different cases and parts of speech)

---

**Example data:**

(12) *Jänest...* (rabbit-PART.SG)

_Jänest ajasid taga rebane, hunt ja karu._ [object, OVS order]

‘(A) fox, wolf and bear chased (the) rabbit *object.*’

(13) *Porgand...* (carrot-NOM.SG)

_Porgand on väga tervislik._ [subject]

‘(A) carrot is very healthy.’

(14) *Raamatute...* (books-GEN.PL)

_Raamatute lugemine avardab mõttemaailma._ [compx-obj]

‘Reading books broadens the mind.’

(15) *Rebase...* (fox-GEN.SG)

_Rebase saba on kohev ja ilus._ [poss-subj]

‘(A) fox’s tail is fluffy and beautiful.’
Coding labels: subj = subject; part subj = partitive subject, as in (2a); obj = object; compx-obj = patient embedded in complex structure (e.g. raamatu lugemine ‘reading [of a] book’); poss-sub = possessor of subject; pred-compl = predicative complement. (Other coding labels were also used but these were the most common.)

3.2 Results

3.2.1 Nominative nouns:

Figure 1. When participants were given a sentence-initial noun in nominative case, what grammatical role did the noun have in the sentences written by participants?

- Mostly interpreted as subject (approx. 90% of NOM prompts), regardless of animacy or number
- Frequency studies: sentence-initial position contributes to subject bias
- Processing accounts: cognitive preference for clause-initial S is possibly universal, over and above frequency, structural bias and animacy (Demiral et al. 2008, Bornkessel-Schlesewsky & Schlesewsky 2009 i.a.)

---

3 Raamatu lugemine avardab mõttemaailma.
book.GEN.PL reading.NOM broaden.PRS.3SG mind.PART.SG
‘Reading books broadens the mind.’
3.2.2 Partitive nouns

![Proportion of continuations](image)

*Figure 2.* When participants were given a sentence-initial noun in **partitive case**, what grammatical role did the noun have in the sentences written by participants?

- PART nouns tend to be interpreted as **objects** (average ~65%) or if plural, either objects or **partitive subjects** (average 25%)
- **Singular PART nouns:** More object interpretations with *animate* singular PART nouns than *inanimate* singular PART nouns (anim: 81% objects, inanim: 55.6% objects, \( p < .001 \)).
- **Plural PART nouns:** More partitive subject interpretations with *animate* than *inanimate* (\( p < .05 \)), no effect of animacy on rate of object interpretations.

- Speaker’s expectations regarding partitive nouns are sensitive to both number and animacy.
3.2.3 Genitive nouns

- **Singular animate GEN nouns** are likely to be interpreted as possessors of the subject.
- **Singular inanimate GEN nouns** are split between possessors, objects, and patients embedded in complex structures (e.g. *raamatu lugemine* ‘[the] reading [of a] book’).
- **Plural GEN nouns** are split between possessors and patients embedded in complex structures.

- Strong interaction with number and animacy
- Only 13% of GEN SG inanimate nouns interpreted as object of a transitive verb, i.e. as syntactically accusative objects ⇒ presumably related to word order effects?
- Although genitive nouns cannot signal the subject role, sentence-initially they are likely to be embedded in a subject NP.

3.3 Summary of Experiment 1:

*When nouns are in sentence-initial position, we see indications of subject biases in all three cases:*

- Nominative nouns were overwhelmingly interpreted as subject (independently of any other factors, e.g. animacy)
- 25% of partitive plural nouns were interpreted as subject (depending on whether the case-marked noun was animate or not)
- Even genitive nouns were frequently assigned a possessor role and embedded in a clause-initial subject noun phrase (depending on animacy)

**Question:** To what extent is this preference for subject interpretations due to the sentence-initial position of the nouns?
- We conducted Experiment 2 to find out what grammatical role participants assign to case-marked nouns when they are not constrained to sentence-initial position.
4. Experiment 2

4.1 Method, design, participants

- **Method**: Sentence completion. Used same prompt nouns as Experiment 1, but did not specify sentence-initial position.
- **Task**: Write a complete sentence using the given noun anywhere in the sentence.
- 42 native Estonian participants, none of whom had participated in Experiment 1. Web-based experiment.

4.2 Results

4.2.1 Linear order

![Figure 4](image.png)

*Figure 4. Linear order patterns of bare nouns: Out of the bare nouns in each case, what proportion were used in sentence-initial vs. non-initial position? (collapsing number and animacy)*

- Sentence-initial bias for NOM and GEN nouns, reverse pattern for PART
- For each case-marker, at least 25% of nouns were freely produced in sentence-initial position

4.2.2 Nominative nouns

- Experiment 2 confirmed the subject preference of nominative nouns.

---

4 We report data only for ‘bare nouns’, i.e. nouns that participants produced without preceding numbers (e.g. *two rabbits*) or other modifiers. This keeps the data set most comparable to Experiment 1, and also avoids complications resulting from the fact that in Estonian, numbers greater than ‘one’ as well as some quantifiers require partitive nouns.
Figure 5. Bare nominative nouns in sentence-initial position: Grammatical role as a function of animacy and number.

- **Nouns produced in initial position:** NOM nouns in initial position were nearly all subjects (over 89%), (Figure 5)
- **Nouns produced in non-initial position** (not shown on graph): NOM nouns used in non-initial position showed effects of both number and animacy => Animate plural nouns were still mostly assigned a subject role, whereas inanimates were divided between subject and object.

### 4.2.3 Partitive nouns

Experiment 2 confirmed the object preference of partitive nouns.

Figure 6. Bare partitive nouns in sentence-initial position: Grammatical role as a function of animacy and number.
PART nouns were assigned the object role, regardless of number, animacy or sentence position. See Figure 6.

The rate of PART subject continuations is numerically higher with animates (27.3%) than with inanimates (14.3%; almost twice as many). Experiment 1: animate nouns were used significantly more often as partitive subjects than inanimate nouns (35% vs. 14%)

4.2.4 Genitive nouns

Experiment 2 confirms tendency to assign clause-initial animate GEN nouns possessor of subject role.

Nouns produced in initial position: (i) animate nouns are mostly interpreted as possessor-of-subject (i.e., not as object) whereas (ii) inanimate nouns are split between different grammatical roles.

- Rate of possessor-of-subject interpretations: 78% with singular animate genitives, 57% with plural animate genitives

Nouns produced in non-initial position (not shown on graph): Higher rate of object interpretations than initial nouns!

- Rate of object interpretation: 66.7% with singular animate genitives, 88.2% with singular inanimate genitives

- With both animate and inanimate singular nouns, object interpretations are now the most frequent option

This contrasts strikingly with what we saw in Experiment 1 and in the sentence-initial usage in Experiment 2.

Figure 7. Genitive nouns in sentence-initial position: Grammatical role as a function of animacy and number.

- Nouns produced in initial position: (i) animate nouns are mostly interpreted as possessor-of-subject (i.e., not as object) whereas (ii) inanimate nouns are split between different grammatical roles.
  - Rate of possessor-of-subject interpretations: 78% with singular animate genitives, 57% with plural animate genitives

- Nouns produced in non-initial position (not shown on graph): Higher rate of object interpretations than initial nouns!
  - Rate of object interpretation: 66.7% with singular animate genitives, 88.2% with singular inanimate genitives
  - With both animate and inanimate singular nouns, object interpretations are now the most frequent option

This contrasts strikingly with what we saw in Experiment 1 and in the sentence-initial usage in Experiment 2.
• These patterns make sense, given that sentence-initial objects, although grammatical, are less frequent in Estonian than sentence-initial subjects (Lindström 2004, Tael 1988).

5. Discussion and conclusions

Based on the results of two sentence-continuation experiments, the interpretation of case-marked nouns in Estonian shows that each case form/marker interacts differently with the following factors:

• initial vs. non-initial position
• animacy (animate/inanimate)
• number (singular/plural)

• Nominative nouns – morphologically unmarked case forms – tend to be interpreted as subjects, regardless of animacy, as has been previously demonstrated in other languages (e.g. Demiral et al. 2008, Bornkessel-Schlesewsky & Schlesewsky 2006).
  
  o Is it the nominative case form, or the unmarked form in interaction with initial position, which is interpreted as subject?

• With partitive nouns, it is number rather than animacy which affects the likelihood of subject vs. object interpretations.
  
  o Is the variable interpretation of partitive case connected to the semantics of partitivity, whereby plural referents have different semantic interpretations available to them than singular referents? (e.g. Cann and Miljan 2012)
  
  o Is animacy an independent factor (e.g. Dahl 2008, Rosenbach 2005, 2008) or reducible to other factors (Silverstein 1976, Blake 2001) in the interpretation of grammatical relations, since it interacts differently with different case-markers?

• With genitive nouns, animacy is a significant factor, together with sentence position. These two factors conspire to enable nouns in genitive, a non-subject case, to occur nevertheless in the subject function – as a modifier of the subject!
  
  o Case-marking interacts with animacy and sentence position.
  
  o The more syntactically underspecified a case-form is, the stronger the bias to interpret sentence-initial nouns as subjects.

• More generally, the initial subject bias and the semantic animacy feature of the case-marked noun seem to interact closely with information provided by case-marking on the noun (if case-marking does not provide strong enough cue).
References
Bornkessel, I., & Schlesewsky, M. 2006. The extended argument dependency model: a
Bornkessel-Schlesewsky I, Schlesewsky M. 2009. *Processing syntax and morphology: a
neurocognitive perspective*. Oxford: Oxford University Press.
Cann, Ronnie & Miljan, Merilin. 2012. Differential case-marking: Syntactic descriptions and
141–150.
Demiral Şükri, Schlesewsky Matthias, Bornkessel-Schlesewsky, Ina. 2008. On the universality of
Hietam, Katrin. 2003. *Definiteness and Grammatical Relations in Estonian*. Ph.D dissertation,
University of Manchester, UK.
number: Estonian speakers’ interpretations of grammatical role.*
Kaiser, E.; Miljan, M. & Vihman, V.-A. (Under review) *Estonian speakers’ interpretations of
morphological case: Implications for Case/Agree.*
5 September 2015. Leiden University Centre for Linguistics (LUCL), Leiden.
Language Production and Comprehension*. Dordrecht: Springer.
distinguishing between core arguments in Estonian]. In Lindström, Liina (ed.), *Lauselikmeist eesti keeles* [Grammatical relations in Estonian]. Tartu Ülikooli eesti keele õppetooli
preprintid 1. Tartu, 40–49.
Miljan, Merilin & Ronnie Cann. 2013. Rethinking case marking and case alternation in Estonian.
Norris, Mark. 2015. *Case-marking in Estonian pseudopartitives*. Paper presented at the 41st Meeting
Prat-Sala M, Branigan H. 2000. Discourse constraints on syntactic processing in language production:
A cross-linguistic study in English and Spanish. *Journal of Memory and Language*, 42:168–
182.
*Lingua* 118, 151–171.
Tael, K. 1988. *Sõnjärjehallid eesti keeles* (võrrenduna soome keelega) [Word order patterns in
Estonian (in comparison with Finnish)]. Preprint, KKI-56. Tallinn: Institute of Estonian
Language and Literature.