

# Double Counting

The semantics of numerals and number in Estonian

Tom Roberts

Institute for Language Sciences, Utrecht University

SALT 34 @ UR, May 28, 2024



Universiteit  
Utrecht

# Variation in Noun-Numeral Constructions (NNCs)

Number-marking in languages with morphological plurals:

## Plural-marking

### (1) English

- a. one doctor/\*doctors
- b. three doctors/\*doctor

### (2) Dutch

- a. één arts/\*artsen
- b. drie artsen/\*arts

### (3) French

- a. un médecin/\*médecins
- b. trois médecins/\*médecin

## Singular (bare)-marking

### (4) Georgian

- a. erti ekimi/\*ekimebi
- b. sami ekimi/\*ekimebi

### (5) Turkish

- a. bir doktor/\*doktorlar
- b. üç doktor/\*doktorlar

### (6) Welsh

- a. un meddyg/\*meddygon
- b. tri meddyg/\*meddygon

# A fractured theoretical landscape

## A big question

What is the source of this cross-linguistic variation in NNCs?

- ❖ **SG operating on atoms vs minimal elements?** (Martí 2020; Scontras 2022)
- ❖ **Multiple kinds of PL?** (Mathieu 2014)
- ❖ **Strong vs weak singular?** (Bale et al. 2011; Alexiadou 2019)
- ❖ **Plural = agreement?** (Krifka 1989; Borer 2005; Deal 2017; Ionin & Matushansky 2018; Saĝ 2019)
- ❖ **Syntax-semantics interface output constraints?** (Farkas & de Swart 2010)
- ❖ ...

# General problems with NNC theories

Most accounts: strong (semantically marked) **singular** vs. unmarked plural

- ✦ Theoretically odd since plural is universally more morphologically marked (Farkas & de Swart 2010; Alexiadou 2019)

No accounts: Predict semantically potent intralinguistic number variation within NNCs

- ✦ 'Optional' plural in Western Armenian/Miya NNCs, but no semantic import (Bale et al. 2011)

A challenge for theories of NNCs: **Estonian**

- ❖ Productive meaning alternation depending on number-marking within NNCs
- ❖ Seems to suggest numerals can 'count' both individuals and groups, depending on nominal morphology

Proposal:

- ❖ Meaning of SG/PL are universally strong across languages (with the SG/PL distinction)
- ❖ Cross-linguistic variation in NNCs lives in the **numeral** rather than number
- ❖ Does away with theoretically undesirable dissociation of morphologically marked but semantically unmarked PL

# Estonian NNCs

# Estonian: A singular-marking language (...?)

## Singular NNCs

Num  $\neq$  1: singular, partitive N

- (7) a. kaks **kinnast**  
two glove.SG.PART  
'two gloves'  
b. kuus **nokklooma**  
six platypus.SG.PART  
'six platypi'  
c. tuhat **aastat**  
thousand year.SG.PART  
'one thousand years'

If Num = 1, singular, **nominative** N

- d. üks **laul**  
one song.SG.NOM  
'one song'

## Plural NNCs

All Num: **plural, nominative** N

- (8) a. kahe-d **kinda-d**  
two-PL glove-PL.NOM  
'two pairs of gloves'  
b. ühe-d **lapse-d**  
one-PL child-PL.NOM  
'one group of children'

## Pluralia tantum

- (9) a. ühe-d **kääri-d**  
one-PL scissors-PL.NOM  
'one pair of scissors'  
b. nelja-d **pulma-d**  
four-PL wedding-PL.NOM  
'four weddings'

# Interpretation of plural NNCs

**Groups:** conventional (e.g. gloves, shoes), or contextually-supplied:

- (10) For that reason it was very cool that did so well in this group: reaching the semifinals were three English-speaking teams, three Estonian teams, ...

ühe-d sakslase-d Bremenist ja ühe-d leeduka-d.

one-PL German-PL Bremen.ELA and one-PL Lithuanian-PL

‘...one German team from Bremen and one Lithuanian Team.’

(Norris 2018: 4)

**Kinds:**

- (11) On olemas kolme-d tähe-d: taeva-, kirja-, ja raha-  
is be-INE three-PL  $\sqrt{\text{TÄHT}}$ -PL: sky-, written-, and money-  
‘There are three kinds of *täht*: sky- (stars), written- (letters), and money-  
(bills).’

(Norris 2018: 5)

(The numeral is also marked for the plural; I follow Norris in assuming this to be a semantically vacuous reflex of concord.)



# Why Estonian is a problem

Estonian has systematic meaning differences in NNCs **depending on the number of the head noun**:

- ✦ Entirely unexpected for theories in which plural in NNCs is plain old agreement
- ✦ Intuitively, plural NNCs are counting larger units than single ones, requiring some flexibility in the numeral
- ✦ Complicates the picture of NNCs in general (not just plural- vs. singular-marking!)

# Analysis

# Proposal in a nutshell:

Estonian numerals count **singular and plural nouns differently**:

- ✦ numeral + SG N: counts atoms
- ✦ numeral + PL N: counts sums

This can be straightforwardly achieved given two assumptions:

- ✦ **Singular** nouns denote sets of atoms; **plural** nouns denote sets of sums
- ✦ Numerals are sensitive to this distinction

Cross-linguistic variation happens in the numerals themselves:

- ✦ Two parameters of variation: whether numerals count (only) atoms and whether they can generate sums

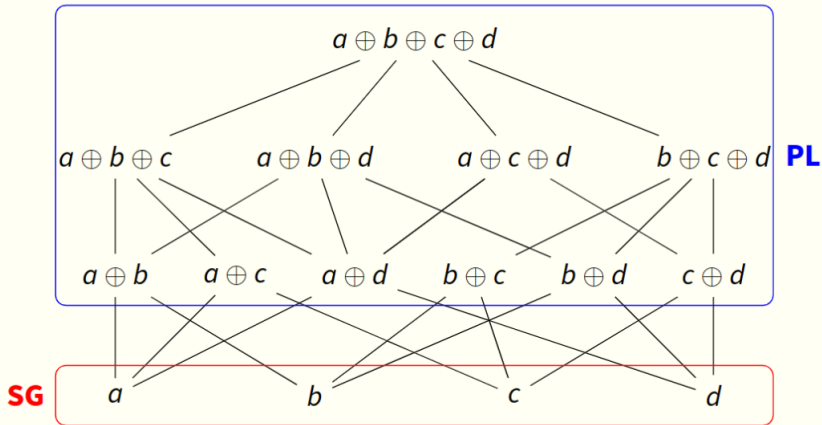
# Background: Assumptions

1. Numerals are modifiers of type  $\langle et, et \rangle$  (Ionin & Matushansky 2006)
2. Domain of entities forms Boolean semilattice upward-closed under sum  $\oplus$  (Link 1983)
3. Extension of bare nouns: entire semilattice (atoms + all sums)
  - Not crucial, but allows for SG/PL to be restrictive modifiers (Partee 2010; Bale et al. 2011)

SG and PL pick out atomic and non-atomic subsets of the semilattice, respectively:

- (12)    a.     $\llbracket \text{SG} \rrbracket = \lambda P_{et} \lambda x_e. P(x) \wedge \text{ATOM}(x)$   
          b.     $\llbracket \text{PL} \rrbracket = \lambda P_{et} \lambda x_e. P(x) \wedge \neg \text{ATOM}(x)$

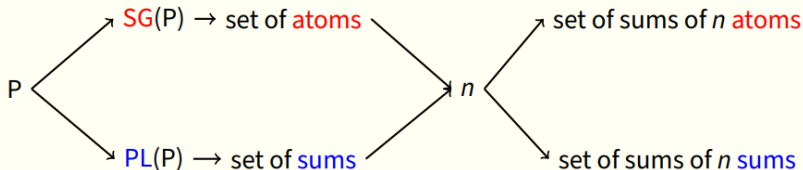
# Example lattice



# How Estonian numerals work

A numeral  $n$  combines with a (singular or plural) noun  $P$  and returns the (characteristic function of the) set of entities  $x$  which:

- a) Are (possibly sums of)  $P$ 's  $*P(x)$
- b) Can be divided into  $n$  subparts...  $\exists Q_{et}[\oplus Q = x \wedge |Q| = n$
- c) ...each of which is in  $P$ ...  $\forall y \in Q[P(y)$
- d) ...and are mutually disjoint  $\forall z \in Q[y \neq z \rightarrow y \cap z = \emptyset]]]$



## Application: Singular NNCs

Let  $\{a, b, c, d\}$  be the set of gloves.

$$(13) \quad \llbracket kaks \rrbracket = \lambda P_{et} \lambda x_e. *P(x) \wedge \\ \exists Q_{et} [\oplus Q = x \wedge |Q| = 2 \wedge \forall y \in Q [P(y) \wedge \forall z \in Q [y \neq z \rightarrow \\ y \sqcap z = \emptyset]]]$$

$$(14) \quad \begin{array}{l} \text{a. } \llbracket kinnas_{bare} \rrbracket = * \{a, b, c, d\} \\ \text{b. } \llbracket SG \rrbracket (\llbracket kinnas \rrbracket^M) = \{a, b, c, d\} \\ \text{c. } \llbracket kaks \rrbracket (\llbracket SG \rrbracket (\llbracket kinnas \rrbracket^M)) = \\ \quad \{a \oplus b, a \oplus c, a \oplus d, b \oplus c, b \oplus d, c \oplus d\} \end{array}$$

**‘The set of all sums of gloves which can be divided into two distinct individual gloves’**



## Application: Plural NNCs

- (15) a.  $\llbracket \text{kinnas}_{\text{bare}} \rrbracket = * \{a, b, c, d\}$   
b.  $\llbracket \text{PL} \rrbracket (\llbracket \text{kinnas} \rrbracket^M) = \{a \oplus b, a \oplus c, a \oplus d, b \oplus c, b \oplus d, c \oplus d, a \oplus b \oplus c, a \oplus b \oplus d, a \oplus c \oplus d, b \oplus c \oplus d, a \oplus b \oplus c \oplus d\}$   
c.  $\llbracket \text{kaks} \rrbracket (\llbracket \text{PL} \rrbracket (\llbracket \text{kinnas} \rrbracket^M)) = \{a \oplus b \oplus c \oplus d\}$

**‘The set of all sums of sums of gloves which can be divided into two disjoint **sums** of gloves.’**

# Pluralia tantum

Recall: Estonian pluralia tantum NNCs always plural-marked

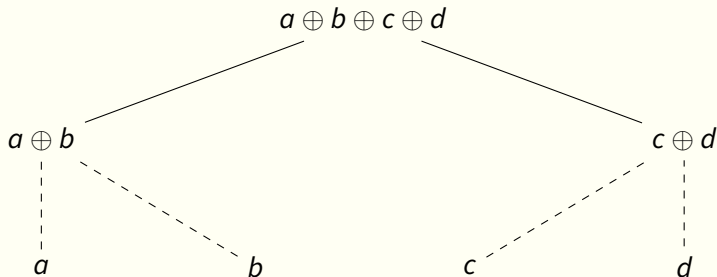
PT are at least partly arbitrary and lack a stable cross-linguistic characterization (Corbett 2019), but coherent categories exist:

- ✦ **Objects with two identical parts:** *käärid* ‘scissors’, *prillid* ‘glasses’, *püksid* ‘trousers’, *tangid* ‘tongs’, *pihid* ‘pliers’
- ✦ **Large events:** *pulmad* ‘wedding’, *matused* ‘funeral’, *peied* ‘wake’, *jõulud* ‘Christmas’
- ✦ **Complex objects:** *päitsed* ‘bridle’, *kangasteljed* ‘loom’
- ✦ **Other:** *leetrid* ‘measles’

My proposal: What is special about PT is that their mereology is deficient:

- ✦ Their minimal parts are **non-atomic**
- ✦ No closure under  $\oplus$  if ‘atoms’ are included

# The mereology of pluralia tantum



- ❖ A pair of scissors is *conceptually* (at some level) a sum of two blades, neither of which is individually ‘scissors’ (cf. gloves)
- ❖ The atoms can also not be freely summed: we cannot mix and match blades to make new pairs of scissors
- ❖ The extension of a hypothetical singular form is always empty, so they need to be plural in order to be non-vacuous

# Predictions and consequences

## Plurals and group size

**Prediction 1:** Plural NNCs can describe groups of arbitrary size  $>1$ .

Example: *two glove.PL* can describe two groups larger than pairs:

(16) *Context: I have 8 gloves in my closet, 4 yellow and 4 red.*

Mul on kahed kindad: kollased ja punased.

1SG.ADE is two.PL glove.PL yellow.PL and red.PL

'I have two kinds of gloves: yellow and red.'

And cannot easily describe two groups if one group has only one member:

(17) *Context: I have 5 gloves, 1 yellow and 4 red.*

??Mul on kahed kindad: kollased ja punased.

This prediction seems borne out.

## Plural NNCs require salient groupings

Odd feature of this semantics: equivalences like the following:

$$\llbracket \text{two glove.PL} \rrbracket = \llbracket \text{four or more glove.SG} \rrbracket$$

Plurals denote sums, but the ‘cover’ (in the sense of Sauerland 2003) of plural NNCs is non-recoverable

So we need to explain why they typically require this cover to be salient:

(18) *Context: A pile of 4 identical gloves is on the table.*

- a. #Kahed **kindad** on laual.  
two.PL glove.PL are table.ADE  
Intended: ‘Two pairs of gloves are on the table.’
- b. Neli **kinnast** on laual.  
four glove.SG are table.ADE  
‘Four gloves are on the table.’

# The pragmatics of plural NNCs

Informal sketch of how to account for this contrast:

- ✦ Plural NNCs are more morphologically marked than semantically equivalent singular NNCs
- ✦ Therefore, using a plural NNC must be done for some pragmatic effect
- ✦ Tentative suggestion: plural NNCs signal the desire to ‘unpack’ their referent in an atypical way:
  - ✦ Interpreting a plural NNC requires unpacking its elements into  $n$  groupings, rather than counting atoms
  - ✦ Analogous to differences in mental representations of universal quantifiers? (Knowlton et al. 2021)

# Outstanding issue 1: Case

Singular NNCs: partitive; plural NNCs: nominative. **Why?**

Classic syntactic story: some kind of competition between [PL] and [PART] (Borer 2005; Danon 2012; Mathieu 2014; Norris 2018)

But it's not clear that this competition story is independently motivated.

Besides: Estonian/Finnish partitive case is semantically contentful (Krifka 1992; Kiparsky 1998, 2001; Craioveanu 2014; Roberts 2020, a.m.o.)

→ What about the **meaning** of partitive case?



# Outstanding issue 1: Case

Finnish NNCs work similarly to Estonian.

Proposal for Finnish NNCs (Sutton & Little 2020): Numeral (type  $n$ ) + NOM = Type clash.

- ✦ Makes the faulty prediction that plural NNCs should not be nominative

But, S&L's proposal for partitive case in measure constructions might do the job:

- ✦ PART yields sets which exclude contextually maximal elements
  - ✦ In context of NNCs, maximal =  $\oplus P$ ? Then singular (indefinite) nouns will always be non-maximal.

## Outstanding issue 2: Plural reference

General problem for strong plural: plural N's in many languages in some contexts receive non-exclusive readings, including Estonian:

- (19) Kõik kellel on **lapsed**, teavad kindlasti kui hea asi on öölamp.  
all who.ADE is children know certainly how good thing is nightlight  
'Everyone who has children (1 or more) certainly knows what a good thing a nightlight is.'

Motivated many accounts of plural as including atoms (Sauerland et al 2005, Spector 2007, Ivlieva 2013, Martí 2020, Renans et al 2020, a.m.o.)

Alternative solutions are on the market!

- ✚ These nouns are number-neutral, not plural (Grimm 2013)

# The cross-linguistic picture

# What can('t) vary across languages

I propose that singular and plural are universal, and what varies is **numerals**, along two axes:

1. **Counting:** Atoms and sums or only atoms
2. **Elements in extension of NNC:** In extension of  $P$  or  $*P$ ?

	Count Atoms	Count Atoms & Sums
$[[\text{NNC}]] \subseteq P$	English	Implausible
$[[\text{NNC}]] \subseteq *P$	Turkish	Estonian

## Deriving English: plural-marking

**English:** Numerals count atoms; counted elements must match number morphology on noun

$$(20) \quad \llbracket two \rrbracket = \lambda P_{et} \lambda x_e. P(x) \wedge |\{y \sqsubseteq x : \text{ATOM}(y)\}| = 2$$

**If  $P$  is singular:**  $P$  is a set of atoms. There are no atoms which contain more than 1 atom, so only *one* is compatible with  $P$

**If  $P$  is plural:**  $P$  is a set of sums. All sums contain at least 2 atoms, so only and all numbers greater than 1 compatible with  $P$

## Deriving Turkish: singular-marking

**Turkish:** Numerals count atoms; counted elements **need not** match number morphology

$$(21) \quad \llbracket iki \rrbracket = \lambda P_{et} \lambda x_e. *P(x) \wedge |\{y \sqsubseteq x : \text{ATOM}(y)\}| = 2$$

**If  $P$  is singular:**  $*P$  contains all  $P$ -atoms, so any numeral compatible with  $P$

**If  $P$  is plural:**  $*P$  contains all  $P$ -sums, so numbers greater than 1 are compatible with  $P$  ...**erroneously expected to be good!**

Proposed solution: the plural is more marked than the singular, so in cases where both are licit, special inference assigned to plural

# PL NNCs in Turkish

NNCs in Turkish **not** categorically ruled out!

But: refer only to ‘well-known groups’ (Görgülü 2012; Sağ 2018, 2019; Alexiadou 2019, Pancheva & Cao 2024)

- (22) a. yedi **cüce-ler**  
seven dwarf-PL  
‘The Seven Dwarves (*Snow White*)’  
b. yedi **cüce**  
seven dwarf.sg  
‘seven dwarves (of no special kind)’

(cf. *kırk harami-ler* ‘forty thieves’ (lit. forty sinners), *yedi deniz-ler* ‘seven seas’, etc.)

Possible avenue: plural morphology emphasizes the status of the referent as a coherent group?

# Numerals as a locus of variation

Other languages supply evidence for numerals as a locus for NNC variation.

**Czech:** NNCs plural (with small *n*); dedicated numeral suffixes for group- and kind-counting (Dočekal 2012, see also Wągiel 2014 for Polish):

- (23)
- a. dva kliče  
two keys  
'two keys'
  - b. dv-oj-**e** kliče  
two-NONCARD-GRP keys  
'two bunches of keys'
  - c. dv-oj-**i** kliče  
two-NONCARD-KIND keys  
'two kinds of keys'



# Numerals as a locus of variation

**Syrian Arabic:** Noun paradigm with differential number marking dependent on numeral (Jochemsen 2020)

(24) Numerals 3-10: plural NNCs

tlatt kazzaab-en/\*kazzaab

three liar-PL/liar.SG

‘three liars’

(25) Numerals >10: singular NNCs

edaesh kazzaab/\*kazzaab-en

eleven liar.SG/liar-PL

‘eleven liars’

# Conclusion

# In a nutshell

The landscape of variation in NNCs is **more complicated** than commonly assumed.

Estonian: plural NNCs require counting groups, singular NNCs counting atoms–need for a semantically interpretable plural!

- Provides some empirical motivation to get away from ‘plural as semantically inert’
- ✦ Potentially good news for alignment of semantic/morphological markedness

Treating number meaning as universal means shifting the burden of variation onto numerals

This approach may generate interesting predictions, but is probably still too simplistic

# Future prospects

Number marking interacts in semantically interesting ways with other linguistic elements:

- ✦ (In)definiteness
- ✦ Case (in Finnish!) (Sutton & Little 2020)
- ✦ Verbal agreement (in Finnish!) (Kaiser 2023)

Future of this project: investigating these interactions in Estonian

Call to arms: We still need to test theories about number and numerals systematically against a broad typology of languages.

# Thanks!

*Aitäh* to Milica Denić, Jakub Dotlačil, Marju Kaps, Mark Norris, Rick Nouwen, Joost Zwarts, and anonymous SALT reviewers for helpful comments on various (some very early) stages of this work, as well as Marju, Einar Treimann, and Uku Visnapuu for judgments. All errors are in-house. his research was supported in part by Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) Grant #406.18.TW.009.

List of references available on request.

Get in touch: [t.d.h.roberts@uu.nl](mailto:t.d.h.roberts@uu.nl)