

Structure and Acquisition of Estonian Semantic Case

Evidence for discrete feature acquisition in Estonian case morphology

Brett Sutton, Georgetown University Department of Linguistics

Contact Information:

Department of Linguistics, Georgetown University
3700 O St. NW, Washington, DC 20057
Email: brs39@georgetown.edu

Abstract

Estonian has a complicated but well-structured system of grammatical and semantic case. This study proposes a feature-based description of the locative cases [2, 1], which is then used to analyze patterns of case acquisition from the CHILDES database. Results indicate that a grammatical account predicts the order of acquisition more accurately than a frequency or phonological approach and suggest the feature-based account of semantic case accurately describes the child's representations.

Background & Description

Estonian is widely known for its large inventory of case morphemes, which includes 3-4 structural cases and 7-11 semantic cases [8, 5]. The various Estonian case forms are shown below:

Case	Singular	Plural
Nominative	raamat	raamatud
Genitive	raamatu	raamatute
Partitive	raamatut	raamatuid
Illative	raamatusse	raamatutesse
Inessive	raamatus	raamatutes
Elicative	raamatust	raamatutes
Allative	raamatule	raamatutele
Adessive	raamatul	raamatutel
Ablative	raamatult	raamatutelt
Translative	raamatuks	raamatuteks
Terminative	raamatuni	raamatuteni
Essive	raamatuna	raamatutena
Abessive	raamatuta	raamatuteta
Comitative	raamatuga	raamatutega

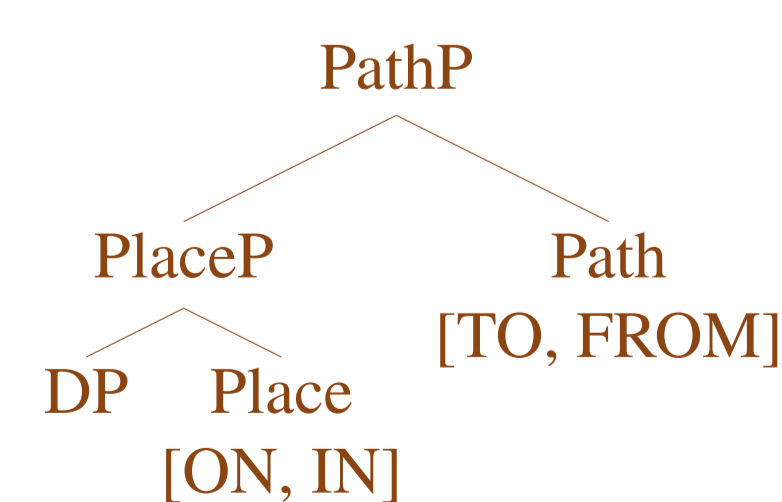
The nominative, genitive, and possibly partitive cases are all structural cases. The final four cases behave less like true case markers than postpositions. This can be seen by the way they do not participate in case concord:

(1)

suure.s maaja.s
big.INE house.INE
in a big house

suure.(*ga) keppi.ga
big(*.COM) stick.COM
with a big stick

Inspection of the morphology and semantics of the semantic case markers suggests they are polymorphemic and represent a variety of features. Focusing on the locatives, there appear to be four distinct morphemes. These morphemes parallel an analysis that divides locative PPs further into PathP and PlaceP [2, 1]:



The possibilities are summarized in the table below, with missing forms emphasizing the dependence of PathP on a lower PlaceP. It suggests the historical relationship between the common partitive case

marker *-t* and FROM, though their syntax and semantics are different.

PATH	PLACE	Meaning	Case
-l	-	ON	Adessive
	-e	TO + ON	Allative
	-t	FROM + ON	Ablative
-s	-	IN	Inessive
	-e	TO + IN	Illative
	-t	FROM + IN	Elicative
-	-	-	-
	-e	to?	-
	-t	from?	(Partitive)

The Acquisition Study

The appropriate description of the target grammar does not necessarily describe the representations a child develops while acquiring the forms. However, if the features suggested for Estonian semantic case are actively used in the acquisition of these morphemes, several predictions can be made regarding their relative order of acquisition.

- **Prediction 1:** Grammatical cases precede semantic cases
- **Prediction 2:** Adessive precedes Allative and Ablative
- **Prediction 3:** Inessive precedes Illative and Elative

Regarding the partitive case forms, children may use the FROM feature in their representation or they may analyze partitive as completely distinct from the FROM feature of Path. If the former is true, Partitive will follow the Place forms and coincide with ABL and ELA once the FROM feature is acquired. If the latter is true, partitive forms will be acquired earlier, along with GEN and NOM forms. Either options seems plausible in this account, but it is an answerable question.

- **Question 1:** Is partitive represented with the same FROM feature as ABL/ELA?

Alternatively, the child might develop representations that are not related to the analysis described here. Semantic cases might appear before grammatical cases, as found for English prepositions [4], or frequency might be the more important factor [6, 7] in determining acquisition order. Phonological complexity may also affect acquisition, with monophonemic forms appearing before more complex ones and after null forms. Finally, children may learn in a totally unconstrained way, with no child exhibiting a pattern similar to these possibilities or the others.

- **Alternative 1:** Children will acquire cases in an order predicted by frequency in their input.
- **Alternative 2:** Children will acquire cases in an order related to phonological ease: NOM, GEN > PRT, ADE, INE > ILL, ALL, ABL, ELA
- **Alternative 3:** Children follow no particular order in their acquisition of case morphemes.

To examine how the acquisition of case occurs, the transcripts for three children in the CHILDES database were analyzed over the period of time shown in the table below [3]:

Corpus	Speaker	Start	End	Sessions	Avg Utterances	MLU
Argus	Hendrik	1;8.13	2;5.30	17	87.9	2.5
Kohler	Martina	1;5.18	1;11.28	10	363	4.97
Vija	Andreas	1;07.24	3;01.3	8	400	3.67

Analysis

Each utterance in corpora was manually tagged with part of speech and all relevant morphology. The table below shows the age at which there was first evidence for a particular case form. Evidence of a case came in two ways: either a case was produced on two roots, or that case appeared on a root that was produced with another case.

	M	A	H
NOM			
GEN	M		A H
PRT	M		A H
ADE (ON)	M		A H
INE (ON+TO)	M		A H
ELA (ON+FROM)		M	A
ALL (IN)	M		A H
ILL (IN+TO)	M		H A
ABL (IN+FROM)		M	A
Age	1:03	1:04	1:05

The next step was to analyze the input the children received. The percentages are shown in the table below, showing a remarkably consistent order: relative rank is the same for all inputs, though Hendrik hears no examples of the ablative form in his sessions.

Case	Andreas		Hendrik		Martina	
	%	Order	%	Order	%	Order
NOM	41.5%	1	40.3%	1	41.0%	1
GEN	15.3%	2	18.9%	2	20.6%	3
PRT	14.2%	2	12.7%	3	14.7%	2
ADE	6.2%	4	5.8%	5	5.8%	3
INE	4.6%	4	5.2%	5	2.7%	3
ELA	2.2%	4	2.4%		2.6%	8
ALL	2.0%	4	1.1%	4	1.5%	3
ILL	0.4%	9	0.1%	5	0.4%	3
ABL	0.2%	4	0.0%		0.1%	8

The attested frequencies show the grammatical cases with much higher rates than the locative cases, meaning the **P1** and **A1** are essentially the same. Additionally, **P2** and **P3** are both consistent with **A1**; however, **A1** itself predicts that the [IN] cases will be acquired after the [ON] forms; this is not the case.

Results

The final table shows how the child's results compare to the predictions and alternatives outlined above.

Child	P1	P2	P3	Q1	A1	A2	A3
Andreas	☉	✓	✓	No	-	☉	☉
Hendrik	☉	✓	✓	No	-	-	☉
Martina	✓	✓	✓	No	-	☉	☉

☉: Confirm Prediction

✓: Consistent with prediction

-: Partially consistent with prediction

☉: Prediction is wrong.

- **P1:** Partially confirmed. Grammatical cases were produced before locatives for Andreas and Hendrik and mostly before for Martina.

- **P2:** Partially confirmed. In no cases did ADE follow ALL and ABL; though neither was there conclusive evidence that ADE was acquired much before the other two.

- **P3:** Confirmed. Like **P2**, the results were consistent with the prediction but did not totally conform—the forms combining IN with a path feature never came before IN by itself, though often they appeared at the same time.

- **Q1:** No. There was not a relationship between acquiring partitive case with neither ELA nor ABL. Andreas and Martina acquired PRT well before FROM cases, and Hendrik acquired PRT but never used FROM forms. This confirms the analysis that PRT is a grammatical case and is not related to the locatives despite its morphemic similarity.

- **A1:** Partially confirmed. The most common cases were in fact acquired earliest, as the frequency would predict. After this, though, the results are less strong. Rare locatives are acquired at similar times as relatively common ones, and some common ones in the input show up late.

- **A2:** Partially confirmed. Hendrik's production roughly follows what would be expected if more phonetic complexity was a limiting factor. For Martina and Andreas, on the other hand, PRT was acquired before and at the same time as GEN, which were followed by the locative cases later.

- **A3:** Not confirmed. Each of the possibilities discussed above can capture some of the data, but case forms were not acquired randomly for any child or for any subset.

Conclusions

A feature-based approach to Estonian case acquisition is supported by the evidence. Compared to the frequency and the phonological alternatives, the grammatical approach makes more consistent and accurate predictions. Gaps in the data could be hiding results that would contradict the analysis, though would likely not lend more support to another approach.

References

- [1] Marcel Den Dikken. On the functional structure of locative and directional pps. *The cartography of syntactic structure*, 6:74–126, 2010.
- [2] Ray S Jackendoff. *Semantics and cognition*, volume 8. MIT press, 1983.
- [3] Brian MacWhinney. *The CHILDES project: The database*, volume 2. Psychology Press, 2000.
- [4] Aliyah Morgenstern and Martine Sekali. What can child language tell us about prepositions? *Studies in language and cognition*, pages 261–275, 2009.
- [5] Mark Norris. *A theory of nominal concord*. PhD thesis, University of California Santa Cruz, 2014.
- [6] Sally Rice, Dominiek Sandra, and Mia Vanrespaille. Prepositional semantics and the fragile link between space and time. *Amsterdam studies in the theory and history of linguistic science*, 4:107–128, 1999.
- [7] Maja Savic and D Anelkovic. The role of input frequency in early language production: Children's usage of serbian prepositions. *Frequency Effects in Language Acquisition: Defining the Limits of Frequency as an Explanatory Concept*, 32:145, 2007.
- [8] Valter Tauli. *Standard Estonian Grammar*. Almqvist och Wiksell i distr, 1973.