

## Implicational Structure of the Estonian Nominal Paradigm

Heiki-Jaan Kaalep (Heiki-Jaan.Kaalep@ut.ee)  
Institute of Computer Science, University of Tartu

Speakers of an inflectional language must possess a common rule system for deducing the inflectional forms of a lexeme, applicable for words they have never met before. This rule system is fully automatic, not lexically driven. It operates on extramorphological properties of lexemes, e.g. phonological shape, semantic word class etc. Inflectional classes that embody the rules which operate on extramorphological properties, are called stable (Wurzel 1987) or active (EKG 1995). Inflectional classes which are lexically driven are called unstable (Wurzel 1987) or passive (EKG 1995). The stable, active classes determine the morphological system of the language. Unstable, passive classes are fossils from the past.

Word forms of a paradigm are connected with each other via a system of implicational relations, meaning that it is possible to deduce some form(s) from other(s), called principal part(s). When describing the morphological system of a language, one has to explicate which forms should be considered to be the principal parts and what are the rules for deducing the other forms.

The paper seeks answers to the following questions: what are the principal parts of the Estonian nominal paradigm; which forms are connected with implicational relations, and which are not? These questions have been asked (and answered) before, recently by e.g. (Kross 1984), (Viks 1992), (EKG 1995), (Ehala 1997), (EKK 2007), (Blevins 2008), resulting in different implicational rules and different ways of splitting the vocabulary into inflectional classes. The crucial question is: what are the criteria for judging that one solution is better than the others? The current paper differs from the previous attempts in that it substantiates its claims with frequency data from a text corpus. The paper assumes that the corpus is a model of language use that the speaker (and learner) is exposed to, and that the morphological system has to be consistent with it, in order to be plausible.

Word forms of a lexeme differ in their token frequency in real life speech and texts, and it is impossible for the less common ones to serve as the basis for inferring more common ones. J. Bybee (1995, p. 237) notes that the different prominence of members of paradigms in language usage leads to a hierarchical structure for paradigms in which more frequent members dominate the less frequent ones.

Table 1 shows the type frequency of the members of Estonian nominal paradigms, computed from a morphologically tagged corpus (500,000 thousand tokens), freely available at <http://www.cl.ut.ee/korpused/morfkorpus/>. It indicates the only plausible hierarchy of principal parts for stable, active morphology: singular nominative, singular genitive, singular partitive, etc, in a descending order of frequencies.

Normally, the graph of the hierarchy of principal parts should be a tree with a single root, and the implicational rules should be regular. Conversely, in the absence of regular implicational rules, the paradigm has to contain more than one hierarchy; and this in turn is possible only if all the wordforms that act as the heads of these hierarchies are common enough to get memorised by the speakers. It is the lexically-driven (i.e. unstable, passive) inflectional classes that represent items with a high token frequency.

Estonian inflectional classes differ in their regularity, and consequently, also in their number of principal parts and nature of implicational rules. The fact that Estonian grammar books present all the inflectional classes uniformly, using a constant number of principal forms, reflects a pedagogical stance, not the morphological system.

case	number	
	singular	plural (incl. vocal plural)
nominative	10686	3502
genitive	7654	2744
partitive	4711	2587
illative + additive	506 + 654	248 (6)
inessive	2161	637 (30)
elative	2365	984 (51)
allative	1952	723 (35)
adessive	1788	611 (34)
ablative	490	134 (10)
translative	2001	287 (10)
terminative	273	44 (4)
essive	500	103 (3)
abessive	167	49 (0)
komitative	1825	683 (13)

Table 1. The type frequency of Estonian nominal paradigm members

A description that assumes that it is actually necessary (not just for convenience of presentation, or pedagogy) to postulate the existence of more than one implicational hierarchy for most of the lexemes, is seriously faulted. The considerations and data, presented above, undermine the claims of e.g. (Ehala 1997) and (Blevins 2008) that Estonian nominal paradigms are best described via multiple implicational hierarchies, or that something else than the singular nominative could be considered the root of the hierarchy.

## References

- Blevins, J. P. 2008. Declension Classes in Estonian. *Linguistica Uralica*, vol. 44, no. 4, 241-267
- Bybee, J. L. 1995. Diachronic and typological properties of morphology and their implications for representation. In L. Feldman (ed.) *Morphological aspects of language processing*. Hillsdale, N.J.: Lawrence Erlbaum Associates. 225-246.
- Ehala, M. 1997. *Eesti keele struktuur. Õpik gümnaasiumile*. Kännimees.
- EKG 1995 - Ereht, M., Kasik, R., Metslang, H., Rajandi, H., Ross, K., Saari, H., Tael, K., Vare, S. 1995, *Eesti keele grammatika I. Morfoloogia. Sõnamoodustus*, Tallinn.
- EKK 2007 - Ereht, M., Ereht, T., Ross, K. 2007, *Eesti keele käsiraamat*, Tallinn
- Kross, K. 1984. *Eesti noomeni muutmistüüpide produktiivsus*. ENSV TA, Tallinn.
- Viks, Ü. 1992, *Väike vormisõnastik I. Sissejuhatus & grammatika. II. Sõnastik & lisad*, Tallinn.
- Wurzel, W.U. 1987. System-dependent Morphological Naturalness in Inflection. – in: Dressler, W. U., Mayerthaler, W., Panagl, O., Wurzel, W. U. (eds.). 1987. *Leitmotifs in Natural Morphology. Studies in Language Companion Series*. Vol 10. Amsterdam/Philadelphia. John Benjamins Publishing Company, 59-95.