Morphosyntactic variation in Latvian numeral phrases: a corpus-based study Natalia Perkova Stockholm University

Latvian cardinal numerals can be divided in two groups: declinable (1-9, e.g. *četri*NOM.PL.M *galdi* 'four tables (m)'- *četras*NOM.PL.F *aitas* 'four sheep (f)' - *četriem*DAT.PL.M *bērniem* 'to four boys (m)', etc.) and indeclinable (*vienpadsmit* 'eleven', *divarpus* 'two and a half', etc.). For some numerals, parallel forms exist, e.g., *desmit* (10, indeclinable) vs. *desmits*, *simt* (100) vs. *simts*; this fact conforms with cross-linguistically attested observation that numerals are generally disposed to form a continuum based on the principle "the higher the nounier" (Corbett 1978).

There are some cases of variation observed in Latvian numeral phrases. Indeclinable numerals tend to assign genitive to corresponding NPs, but in some contexts the wider syntactic environment plays the crucial role. The following factors defining the morposyntactic pattern of these numerals are usually mentioned:

1) in a NumP with complex numerals two consecutive genitives are not desirable: it is more preferrable to mark the whole numeral uniformly (Lagzdiņa 1980: 140), that is, [desmit tūkstošiNOM]NUM gaduGEN 'ten thousand years' is better than [desmit tūkstošuGEN]NUM gaduGEN;

2) the formal expression of agreement between the predicate and the subject correlates with a preferrable form of the NumP in the subject position: if the predicate is put into a "neutral" form (which is masculine singular for Latvian), the NumP will be likely to follow the most "neutral" pattern, that is without an overt nominative marking (1a); otherwise the nominative marking implies the occurrence of a proper agreement in gender and number, if it is possible (1b), see (ibid.: 141).

The same is true for the cases where number/gender features should be assigned to an attribute, cf. (2).

S. Lagzdiņa also mentions such examples where one deals with approximate quantity (*about NUM* + NPGEN, *more than NUM* + NPGEN). It seems that the choice of genitive is quite logical there, and some arguments could be presented in favour of complex prepositional quantifiers in such cases.

The corpus data can easily illustrate the correlation of morphosyntactic patterns with certain types of constructions. For example, for the collocation *desmit gad.*+ 'ten years' structural case marking is preferred in the following constructions: it is nominative in age constructions (3) and accusative (*desmit gadus*) in the contexts with *accusative temporis*. In most cases of genitive marking (*desmit gadu*) the same case is also licensed by the governing NP.

Another peculiarity of Latvian NumPs lies in the fact that in the positions of non-structural cases (dative or locative) genitive marking is prohibited: *no desmit gadiem*DAT / **gadu*GEN 'from ten years'. This is very similar to the distribution of genitive NPs in Russian NumPs ([*šest' korov*GEN] NOM/ACC 'six cows', but *šesti*DAT *korovam*DAT). Latvian quantifiers are even less strict in the assignment of case, and it seems that they tend to be "transparent" semantic units, so that the syntactic position of the whole NumP defines the case marking of the NP. This fact can be supported by the use of such indeclinable quantifiers as highly lexicalized *pāris* 'a couple' and a very specific lexeme *ik* 'every'.

Examples

(1a)	Pārvadā-t-s		5000	tonn-u kr		av-as.	
	transport-PP.PST-NOM.SG.M		5000	ton-GEN.PL	cargo	-GEN.SG	
	'5000 tons of cargo have been transported.' (Lagzdiņa 1980: 141)						
(1b)	Krājum-a	sakop	0-t-i		11	stāst-i.	
	collection-LOC.SG	put_tc	put_together-PP.PST-NOM.PL			11	story-NOM.PL
	'There are 11 stories in the collection.' (ibid.)						-

- (2) *četrpadsmit lab-āk-ie rezultāt-i* fourteen good-COMP-DEF.NOM.PL.M result-NOM.PL 'the best 14 results'
 (3) *Man ir desmit gad-i*
- (5) Man ir desmit gaa-i I-DAT be.PRS.3 ten year-NOM.PL 'I am 10 years old.'

References

Corbett, Greville G. 1978. Universals in the syntax of cardinal numerals. *Lingua* 46: 355–368.

S. Lagzdiņa. Lietvārda locījums pie nelokāmajiem skaitļa vārdiem. *Latviešu valodas kultūras jautājumi* - 16. laidiens. - Rīga, Avots, 1980. - 137. - 143. lpp.