Grammaticalizing from N to Q in Polish and Slovak pseudopartitives: A corpus-based study of case marking patterns

Keywords: pseudopartitives; Slavic; quantifiers; grammaticalization; case

Introduction: We investigate a grammaticalization path from noun to quantifier through Polish and Slovak pseudopartitives, a binominal construction where one noun, N1, quantifies another noun, N2 (1a). In the Polish tradition, this process has been termed "numeralization" (Schabowska 1962; see also Rutkowski 2006), but it resembles patterns of change documented in English (Brems 2011) and Spanish (Verveckken 2015), where a reanalysis of the N1 from head to modifier has been proposed. There is evidence that some Polish N1s are shifting to allow a numeral-like subject-verb agreement pattern (e.g. Herda 2019), and we explore whether this change extends to case assignment. We find evidence of a numeral-like case pattern for some N1s, which appears to start in the locative, and argue that such N1s are undergoing a category change from N to Q.

Morphosyntactic patterns: In Polish and Slovak pseudopartitives, N2 typically occurs in the genitive, this resembling other types of binominals (1b); we term this the "N-pattern". This behavior contrasts with numeral constructions, where the quantified noun is marked genitive in structural case contexts (nominative, accusative), but otherwise appears in the case of the case context (Babby 1987's heterogenous/homogeneous distinction), the numeral being inflected (Slovak/Polish) or uninflected (Slovak) (2); we term this the "Q-pattern".

Method: We conducted a corpus study, investigating patterns of case assignment with select N1s. The list of N1s per language, as well as corpora can be found in Table 1/Appendix. A series of queries were constructed per case context per N1. These queries captured N-patterns, Q-patterns, and possible variations of these. Sample size was limited to 100 per query, and false positives were filtered out manually. Results in the locative, dative, and instrumental are discussed here.

N-pattern: All N1s except Slovak *pár*$_2$ and Polish *parę* 'a few' predominately showed the N-pattern, where N2 was genitive across different case contexts. These N1s seem to remain highly nominal.

Q-pattern: Several N1s additionally showed a Q-pattern, where N2 appeared in the case of the case context. For most N1s, this pattern was minimal, but some N1s showed it productively for the locative; see Table 1. The results suggest that these N1s have started grammaticalizing into quantifiers, with the locative leading the change. Slovak/Polish *pár*$_2$*/parę* predominantly showed a Q-pattern, suggesting they have already grammaticalized.

Analysis: Grammaticalization from N to Q interacts with case assignment. When N, an N1 absorbs the external case, but as Q, it fails to intervene, leading to the external case on N2. The availability of multiple patterns for a single N1 suggests a synchronic layering of grammaticalization steps (Brems 2011), which may eventually lead to divergence between the two patterns (as is likely for Polish *parę*/*para*) or a full shift to Q. That the Q-pattern infiltrates the case paradigm incrementally may further suggest a "climbing" of the case hierarchy (Caha 2009), where the locative, a smaller case than dative or instrumental, is affected first. Altogether, our findings provide evidence of ongoing change in the case patterns of Polish and Slovak N1s.

			N-pattern (N1$_X$ N2$_{GEN}$)			Q-pattern (N1$_X$ N2$_X$)		
Lang.	**N1**	**Gloss**	**LOC**	**DAT**	**INST**	**LOC**	**DAT**	**INST**
PL	*kupa*	‘heap’	✓	✓	✓	0	0	0
	fala	‘wave’	✓	✓	✓	<10	0	<10
	para	‘pair’	✓	✓	✓	<10	<10	<10
	szereg	‘row’	✓	✓	✓	✓	0	<10
	mnóstwo	‘multitude’	✓	✓	✓	✓	<10	<10
	parę	‘a few’	✓*	0	<10	✓	✓	✓
SK	*vlna*	‘wave’	✓	✓	✓	<10	<10	0
	hromada	‘pile’	✓	✓	✓	<10	<10	<10
	kopa	‘heap’	✓	✓	✓	=10	<10	<10
	rad	‘row’	✓	✓	✓	✓	<10	<10
	množstvo	‘multitude’	✓	✓	✓	✓	✓	✓
	pár$_1$	‘pair’	✓	✓	✓	0	<10	0
	pár$_2$	‘a few’	<10	<10	0	✓	✓	✓

Table 1: Prevalence of N-pattern and Q-pattern per N1 per case in Polish and Slovak, with 0 indicating no attestations, <10 or =10 a small number of attestations, and ✓ a productive pattern. The addition of an asterisk to the checkmark, i.e. ✓, indicates that despite its prevalence, the pattern may not be productive (to be elaborated on in the talk).*

(1) Noun-like morphosyntax

a. Z dziećmi jest **kupa** **roboty**.
S deťmi je **kopa** **roboty.**
With children.INST is heap.NOM work.GEN
‘It’s a lot of work with children.’

b. Przegląd oferty *PL*
Prehľad ponuky *SK*
overview offer.GEN
‘an overview of the offer’

(PL: Herda 2019; SK: author translation)

(2) Numeral-like morphosyntax

a. *Heterogenous case assignment*
Pięć książek było czytanych.
Päť kníh bolo čítaných.
Five books.GEN were.N.SG(DEF) read.GEN
‘Five books were read.’

b. *Homogenous case assignment*
z pięcioma dziećmi *PL*
s piatimi/päť deťmi *SK*
with five.INST/.Ø children.INST
‘with five children’

Appendix: Corpora

- Polish Web 19 (plTenTen19), Sketch Engine, Jakubíček et al (2013), 4 bil. words
- Slovak Web Corpus (SNK Web 7.0), Jazykovedný ústav Ľ. Štúra SAV (2022), 5.3 bil. words

References:

Babby, L. (1987). Case, prequantifiers, and discontinuous agreement in Russian. *Natural Language and Linguistic Theory 5*, 91-138. • **Brems, L. (2011).** *Layering of size and type noun constructions in English.* De Gruyter. • **Caha, P. (2009).** *The nanosyntax of case*. Doctoral dissertation. University of Tromsø. • **Herda, D. (2019).** On the effect of pluralization on the numeralization of nouns in English and Polish: A contrastive corpus-based study. *Linguistica Silesiana, 40*, 139-155. • **Jakubíček, M., Kilgarriff, A., Kovář, V., Rychlý, P., & Suchomel, V. (2013, July).** The TenTen corpus family. In *7th International Corpus Linguistics Conference CL* (pp. 125-127). • **Jazykovedný ústav Ľ. Štúra SAV. (2022).** *Slovenský národný korpus – web-7.0*. Bratislava. • **Schabowska, M. (1962).** O formalnej numeralizacji rzeczowników. *Język Polski 42*(2), 116–124**.** • **Rutowski, P. (2006, Feb. 17).** Grammaticalization in the nominal domain: The case of Polish cardinals. 4th Workshop in General Linguistics, University of Wisconsin, Madison. • **Verveckken, K. D. (2015).** *Binominal Quantifiers in Spanish*. De Gruyter.