The fine structure of Slavic pronominal copula construction: Empirical investigations in Polish

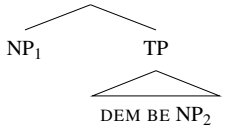
Sebastian Zawada (ICS Polish Academy of Sciences)

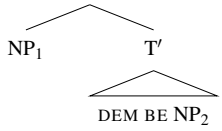
Despite extensive discussion in generative grammar, the syntax of the Slavic pronominal copula construction, exemplified in (1)–(2), remains a matter of debate. These constructions feature the demonstrative form *to* / *èto* (henceforth: DEM) positioned between the left-hand NP (NP_1) and the (sometimes optional) verbal copula, agreeing with the right-hand NP (NP_2).

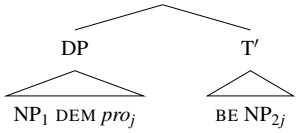
(1) [Leki]$_{NP_1}$ to była [ostateczność]$_{NP_2}$.
medicines.PL.M.NOM DEM was.SG.F last-resort.SG.F.NOM
'Medicines were a last resort.' (Pol., Corpus of Contemporary Polish)

(2) [Evropa]$_{NP_1}$ — èto byl [zemledel'českij materik]$_{NP_2}$...
Europe.SG.F.NOM DEM was.SG.M agricultural continent.SG.M.NOM
'Europe was an agricultural continent.' (Rus., Russian National Corpus)

We present two experiments demonstrating that NP_1 occupies a position in the left periphery, above TP, as illustrated in (3). Therefore, it cannot be analysed as a Spec,TP subject (contrary to Citko 2008 and Bondaruk 2013; see (4)) or as a part of a 'big DP' headed by DEM (Burukina et al. 2024; see (5)). (Although this abstract focuses on experiments – which are based on Polish – the talk also demonstrates (i) that the same phenomena, to varying degrees, occur in Russian and (ii) that known arguments *against* (3) do not go through.)

(3)


(4)


(5)


Experiment 1: Intervention NP_1 blocks *wh*-movement and topicalisation across it, in contrast to canonical copular clauses with instrumental predicates, see (6a)–(6b):

(6) a. *Kto$_i$ Janek [$_{TP}$ to był z zawodu ___$_i$]?
who.SG.M.NOM Janek TO was.SG.M by profession

b. Kim$_i$ [$_{TP}$ Janek był z zawodu ___$_i$]?
who.SG.M.INS Janek was.SG.M by profession
Intended: 'Who was Janek by profession?'

This effect would be unexpected if NP_1 were located within the TP: the Spec,TP position does not constitute barrier for A′-movement. By contrast, a left-peripheral analysis of NP_1 accounts for the data naturally: "Left-dislocated elements create barriers for A′-movement, precisely because they occupy A′-positions themselves" (Landau 2009: 96).

A 2×4 acceptability judgement experiment was conducted to confirm the existence of this phenomenon. 78 native speakers of Polish rated 16 stimuli from 16 token sets (2 stimuli per condition) on a 1–7 Likert scale (with 48 fillers). The first factor was the type of **structure** (the *to* structure vs. canonical copular clauses with instrumental predicates), and the second four-level factor was the (lack of) **movement**: (i) BASE declarative, (ii) *wh*-movement of the whole predicate (EXTtraction, illustrated in (6)), (iii) *wh*-movement from the predicate (SUBextraction) and (iv) TOPicalisation. The data were analysed using a linear mixed-effect model, which revealed a significant interaction between structure and all types of A′-movement ((ii)–(iv)): A′-movement reduced acceptability more in the *to* structure than the instrumental structure. The interaction effects were particularly strong for both types of *wh*-movement ($\beta_{EXT} = -1.74$; $\beta_{SUB} = -1.66$, $p < 0.001$) and weaker – but still significant – for topicalisation ($\beta_{TOP} = -0.79$, $p < 0.01$); see Table 1 for descriptive statistics.

	Mean$_I$	**Mdn$_I$**	**Mean$_T$**	**Mdn$_T$**
BASE	5.57	6	4.77	5
EXT	4.10	4	1.48	1
SUB	4.03	4	1.57	1
TOP	3.67	4	2.05	1

Table 1: Means and medians for Exp. 1 (I = instr. structure, T = *to* structure).

Experiment 2: Complex vs. bare *wh*-phrases and quantifiers NP_1 can host complex *wh*-phrases and quantifiers, but not bare ones, as illustrated in (7). Interestingly, in this regard, it patterns with quantified expressions placed in A′-positions in various languages (e.g., Cinque 1986, Rizzi 1997, van Craenenbroeck 2010).

(7) {[Żaden człowiek]$_{NP_1}$ / *[Nikt]$_{NP_1}$} to nie jest [zabawka]$_{NP_2}$.
no person nobody DEM NEG is toy
'{No person/ #Nobody} is a toy.' (Pol.)

The subtlety of this phenomenon motivated additional 2×2 experiments involving four quantifier types (see subplots in Fig. 1), with **structure** (like in Exp. 1) and **quantifier complexity** (bare vs. complex) as fixed factors. Fig. 1 presents the existence of the effect (the difference between *bare* and *complex* quantifiers is always bigger for the *to* structure than for the control setting, i.e., the instrumental structure), but also indicates the complexity of the dataset (large differences between subplots). Again, linear mixed-effect modelling revealed a significant interaction between structure and quantifier complexity for each quantifier type.

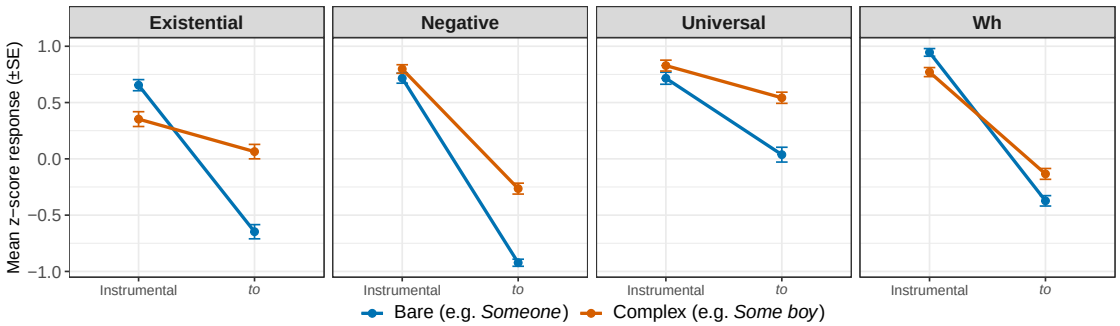


Figure 1: Mean z-score judgements for Experiment 2.

The explanation of this intriguing phenomenon recurring in the literature is that *complex* quantifiers are not operators (e.g., Cinque 1986, Dobrovie-Sorin 1990, van Craenenbroeck 2010). That is, they do not need to bind a variable and thus can be base-generated in a peripheral position, not related to a variable. In contrast, when a *bare* quantifier is base-generated in a dislocated position, spurious quantification arises: it functions as a genuine operator but lacks an associated variable in an A-position to bind. Crucially, regardless of the adequacy of this explanation, the empirical phenomenon itself – given that it is cross-linguistically typical for dislocated phrases – supports the TP-external analysis of NP_1.

Discussion Importantly, the TP-external status of NP_1 raises a question about its exact position. Existing TP-external analyses (e.g., Rutkowski 2006) assume that it occupies Spec,TopP. However, this is problematic, since in both Polish and Russian NP_1 is information-structurally inert and can be focused. We postulate, instead, that a dedicated peripheral head emerged through the integration of a hanging topic. Why did it emerge? The special semantic flavour of the construction at hand (analysed in Citko 2008, Bondaruk 2013: VI.2 for Polish) is tied to the presence of the demonstrative *to*, syntactically functioning as the subject (ending up in Spec,TP). The dislocated NP_1 is then required to provide lexical content for that demonstrative. In this respect, the pronominal copula construction strikingly resembles Pronominal Argument Languages, were *dislocated* NPs are needed to supply lexical content to pronominals that *obligatorily* occupy argument positions (see, e.g., Baker 1996: Part I).

References • **Baker, M. C.** (1996). *The Polysynthesis Parameter*. Oxford University Press, New York, NY. • **Bondaruk, A.** (2013). *Copular Clauses in English and Polish. Structure, Derivation and Interpretation*. Wydawnictwo KUL, Lublin. • **Burukina, I., Borise, L., and Den Dikken, M.** (2024). A 'big DP' analysis of Russian copular constructions with *èto*. *Glossa: A Journal of General Linguistics*, 9(1):1–37. • **Cinque, G.** (1986). Bare quantifiers, quantified NPs and the notion of operator at S-Structure. *Rivista di Grammatica Generativa*, 11:33–63. • **Citko, B.** (2008). Small clauses reconsidered: Not so small and not all alike. *Lingua*, 118(3):261–295. • **Dobrovie-Sorin, C.** (1990). Clitic doubling, *wh*-movement, and quantification in Romanian. *Linguistic Inquiry*, 21(3):351–397. • **Landau, I.** (2009). Against broad subjects in Hebrew. *Lingua*, 119(1):89–101. • **Rizzi, L.** (1997). The fine structure of the left periphery. In Haegeman, L., editor, *Elements of Grammar: Handbook in Generative Syntax*, pages 281–337. Springer Netherlands, Dordrecht. • **Rutkowski, P.** (2006). From demonstratives to copulas: A cross-linguistic perspective and the case of Polish. *Journal of Universal Language*, 7:147–175. • **van Craenenbroeck, J.** (2010). Complex *wh*-phrases don't move: On the interaction between the Split CP Hypothesis and the syntax of *wh*-movement. In Panagiotidis, E. P., editor, *The Complementizer Phase: Subjects and Operators*. Oxford University Press, Oxford.