

Case Concord as Subsumption: Mixed Agreement vs. Case Assignment in Slavic

Introduction Agreement and case assignment seem to be two very different phenomena. When two constituents agree, they usually agree in all appropriate morphosyntactic features. For example, in many languages adjectives and nouns agree in all the features they share, e.g., in various Slavic languages, in number, gender, and case. By contrast, case assignment results in the exponence of a specific grammatical case, but it does not depend on the presence of another constituent bearing the same case value. In some syntactic constructions, both phenomena may be observed, as in many nominative–accusative languages, where subject–verb agreement in ϕ features (i.e., number, gender, and person) is accompanied by the nominative case assignment to the subject.

MACA However, there are constructions that display a more complex, mixed agreement vs. case assignment (MACA) pattern. One such construction, not studied in any depth within LFG, is the numeral–noun construction (NNC) in North (i.e., East and West) Slavic languages, illustrated with Polish in (1).^{1,2}

(1)

position	‘five boys’		‘five girls’	
	‘five.M’	‘boys.M’	‘five.F’	‘girls.F’
NOM	<i>pięciu</i> .PL.M.NOM	<i>chłopców</i> .PL.M.GEN	<i>pięć</i> .PL.F.NOM	<i>dziewczyn</i> .PL.F.GEN
ACC	<i>pięciu</i> .PL.M.ACC	<i>chłopców</i> .PL.M.GEN	<i>pięć</i> .PL.F.ACC	<i>dziewczyn</i> .PL.F.GEN
GEN	<i>pięciu</i> .PL.M.GEN	<i>chłopców</i> .PL.M.GEN	<i>pięciu</i> .PL.F.GEN	<i>dziewczyn</i> .PL.F.GEN
DAT	<i>pięciu</i> .PL.M.DAT	<i>chłopcom</i> .PL.M.DAT	<i>pięciu</i> .PL.F.DAT	<i>dziewczynom</i> .PL.F.DAT
LOC	<i>pięciu</i> .PL.M.LOC	<i>chłopcach</i> .PL.M.LOC	<i>pięciu</i> .PL.F.LOC	<i>dziewczynach</i> .PL.F.LOC
INS	<i>pięcioma</i> .PL.M.INS	<i>chłopcami</i> .PL.M.INS	<i>pięcioma</i> .PL.F.INS	<i>dziewczynami</i> .PL.F.INS

In NNC, the numeral is commonly assumed to be the head – one argument being that, in ACCUSATIVE positions, it is the numeral that is the locus of ACC, while the noun is in the GENITIVE – and the noun in (1) may be replaced by a noun phrase of any complexity. The MACA pattern consists here in the noun being assigned GEN in NOM, ACC, and perhaps GEN positions, while agreeing with the numeral in case in other positions.

Virtually all derivational analyses bind MACA to some special properties of numerals. However, there is a little discussed idiosyncratic construction in Polish that displays virtually the same pattern, namely the “something interesting” construction illustrated in (2).

(2)

position	‘something interesting’		‘somebody interesting’	
	‘something.N’	‘interesting.N’	‘somebody.M’	‘interesting.M’
NOM	<i>coś</i> .NOM	<i>ciekawego</i> .GEN	<i>ktoś</i> .NOM	<i>ciekawym</i> .NOM
ACC	<i>coś</i> .ACC	<i>ciekawego</i> .GEN	<i>kogoś</i> .ACC	<i>ciekawego</i> .ACC
GEN	<i>czegoś</i> .GEN	<i>ciekawego</i> .GEN	<i>kogoś</i> .GEN	<i>ciekawego</i> .GEN
DAT	<i>czemuś</i> .DAT	<i>ciekawemu</i> .DAT	<i>komuś</i> .DAT	<i>ciekawemu</i> .DAT
LOC	<i>czymś</i> .LOC	<i>ciekawym</i> .LOC	<i>kimś</i> .LOC	<i>ciekawym</i> .LOC
INS	<i>czymś</i> .INS	<i>ciekawym</i> .INS	<i>kimś</i> .INS	<i>ciekawym</i> .INS

In this case, the MACA pattern is triggered by a handful of pronouns, all containing the root *c(o)*, including: *co* ‘what’, *coś* ‘something’, *cokolwiek* ‘whatever’, and *nic* ‘nothing’; they all combine with adjectival phrases in GEN when in NOM and ACC (and GEN) positions, and agree with them in all other positions, just like NNCs. This behaviour should be contrasted with the usual situation, exemplified in the final columns of (2), which illustrate that pronouns based on the root *kt(o)*, e.g., *ktoś* ‘somebody’, do not display the MACA pattern; here (and elsewhere in Polish) the adjective always agrees with the (pro)nominal head. A claim that *coś* ‘something’ has some numeral properties that *ktoś* ‘somebody’ lacks would be clearly *ad hoc*; additionally, Lithuanian pronouns such as *kažkas* ‘something’ also display MACA, even though Lithuanian NNCs do not (Rutkowski and Szczegot 2001). Hence, this MACA pattern cannot be explained with recourse to special properties of numerals.

¹Within LFG, some preliminary thoughts about Slavic NNC may be found in Neidle 1988 and Przepiórkowski and Patejuk 2012, but they amount to stipulating which numeral forms assign genitive and which agree, i.e., they miss the generalization that this is a very regular pattern, common to all North Slavic languages. The analysis of NNC below is free from such lexical stipulations.

²In the table below, nominative on the numeral is in grey as there are good arguments that it is actually accusative (e.g., Przepiórkowski 1999) – an issue that is orthogonal to this paper.

Analysis – Basics The analysis of Slavic MACA patterns offered below is based on some independently motivated theoretical constructs (presented on this page) and on some novel ideas (on the next page).

One piece of the analysis is the case markedness hierarchy in (3), proposed in Caha 2009 on the basis of systematic syncretisms and morphological containments in Slavic and other languages. On various accounts, cases higher in the hierarchy have larger structures or more features than cases lower in the hierarchy.

$$(3) \text{ NOM} < \text{ACC} < \text{GEN} < \text{LOC} < \text{DAT} < \text{INS} \quad (\text{Caha 2009})$$

The second independently motivated piece of analysis is the particular LFG encoding of case markedness proposed in Asudeh *et al.* 2024 and Belyaev *et al.* 2025. There, cases are represented as feature matrices, with features such as *NOM*, *ACC*, or *DAT* understood as privative: either present (with the + value) or absent.³ Case hierarchies such as (3) – but also more complex, non-linear hierarchies – are encoded via macros in such a way that macros corresponding to more marked cases in the hierarchy call macros corresponding to immediately less marked cases. The six macros defined in (4)–(9) are needed to encode the hierarchy in (3).⁴

$$(4) \text{ @NOM} := (\downarrow \text{CASE NOM}) = +$$

$$(5) \text{ @ACC} := \text{ @NOM} \\ (\downarrow \text{CASE ACC}) = +$$

$$(6) \text{ @GEN} := \text{ @ACC} \\ (\downarrow \text{CASE GEN}) = +$$

$$(7) \text{ @LOC} := \text{ @GEN} \\ (\downarrow \text{CASE LOC}) = +$$

$$(8) \text{ @DAT} := \text{ @LOC} \\ (\downarrow \text{CASE DAT}) = +$$

$$(9) \text{ @INS} := \text{ @DAT} \\ (\downarrow \text{CASE INS}) = +$$

$$(10) \left[\begin{array}{l} \text{PRED 'GIRL'} \\ \text{CASE} \left[\begin{array}{l} \text{NOM +} \\ \text{ACC +} \\ \text{GEN +} \end{array} \right] \\ \text{PHI} \left[\begin{array}{l} \text{NUMBER PL} \\ \text{GENDER F} \end{array} \right] \end{array} \right]$$

$$(11) \left[\begin{array}{l} \text{PRED 'AMBITIOUS'} \\ \text{CASE} \left[\begin{array}{l} \text{NOM +} \\ \text{ACC +} \\ \text{GEN +} \end{array} \right] \\ \text{PHI} \left[\begin{array}{l} \text{NUMBER PL} \\ \text{GENDER F} \end{array} \right] \end{array} \right]$$

Additionally, ϕ features are grouped into a separate feature matrix (cf. *INDEX* in HPSG), so that the basic f-structure of the form *dziewczyn 'girl.GEN.PL.F'* is as in (10) above.

The third independently motivated facet of the analysis is the “feature sharing” approach to agreement, with agreement features represented directly on all agreeing constituents, as proposed for LFG in Haug and Nikitina 2015. For example, the syncretic adjective *ambitnych* ‘ambitious’, when interpreted as *GEN.PL.F*, is represented at f-structure as plural not just by virtue of this adjective modifying a plural noun, as in standard LFG, but by virtue of expressing this feature directly, as in (11).

Given these assumptions, fully symmetrical concord (i.e., agreement within the nominal domain) may be encoded via the macros in (12)–(14), used in rules as shown in (15), and leading to functional representations as in (16) (with variables ① and ② indicating structure sharing).

$$(12) \text{ @CONCORD} := \text{ @CASECONCORD} \\ \text{ @PHIAGREEMENT}$$

$$(13) \text{ (preliminary version, to be replaced with (17))} \\ \text{ @CASECONCORD} := (\uparrow \text{CASE}) = (\downarrow \text{CASE})$$

$$(14) \text{ @PHIAGREEMENT} := (\uparrow \text{PHI}) = (\downarrow \text{PHI})$$

$$(15) \text{ NP} \rightarrow \text{ AP}^* \text{ N} \\ \downarrow \in (\uparrow \text{ADJ}) \quad \uparrow = \downarrow \\ \text{ @CONCORD}$$

$$(16) \text{ ambitnych dziewczyn} \\ \text{ 'ambitious.GEN.PL.F girls.GEN.PL.F'}$$

$$\left[\begin{array}{l} \text{PRED 'GIRL'} \\ \text{CASE } \textcircled{1} \left[\begin{array}{l} \text{NOM +} \\ \text{ACC +} \\ \text{GEN +} \end{array} \right] \\ \text{PHI } \textcircled{2} \left[\begin{array}{l} \text{NUMBER PL} \\ \text{GENDER F} \end{array} \right] \\ \text{ADJ} \left\{ \left[\begin{array}{l} \text{PRED 'AMBITIOUS'} \\ \text{CASE } \textcircled{1} \\ \text{PHI } \textcircled{2} \end{array} \right] \right\} \end{array} \right]$$

³See Dalrymple *et al.* 2009 for a similar but distinct idea: there, such attributes are binary (valued + or –, or absent) and they are used to model syntactic effects of syncretism.

⁴These macros are defined in such a way that they can be used in syntactic rules, e.g., to mark nominal dependents of numerals as genitive, as in (18) in the main text below. More general – parameterized – macros are needed to be also useful in lexical entries, e.g., to specify a given argument of a verb as genitive, as in the analysis of “something interesting” alluded to at the end of this abstract.

Analysis – MACA The gist of the proposed analysis of Slavic NNCs is that two constituents may agree in ϕ features and case, but one of them may be additionally assigned a specific case, which sometimes results in case inequality – lack of case concord – between the two constituents. This seemingly paradoxical “case agreement but perhaps no agreement in the end” idea makes sense if case concord is modelled via subsumption rather than equality. That is, the definition of @CASECONCORD in (13) above should be replaced with (17) below.

(17) @CASECONCORD := (\uparrow CASE) \sqsubseteq (\downarrow CASE) (final version, replaces (13))

Nothing substantial changes in the analysis of simple adjective–noun concord illustrated with (15)–(16) above: the values of the two PHI features are still identical, and also CASE features are the same, but without the constraint that they must remain identical, so that some other part of the grammar may in principle further specify the value of the adjective’s CASE. As this does not usually happen in adjective–noun concord, the two CASE values remain identical. However, this symmetry is broken in NNCs, which are licensed by syntactic rules such as (18).

(18) NumP \rightarrow Num NP
 $\uparrow=\downarrow$ $\downarrow = (\uparrow$ OBJ)
 @CONCORD
 @GEN

Such rules are directly responsible for MACA patterns. @GEN makes sure that the NP is always at least NOM+, ACC+, GEN+. When the numeral is in a NOM, ACC, or GEN position (i.e., bears one of these cases itself), the subsumption in @CONCORD is true and it does not have any additional effect, but when the numeral bears a more marked case, than @CONCORD results in the noun also receiving these additional case features.

This analysis rests on the assumption that lexical entries cannot *define* case, they can only *expose* case defined in the grammar (in the spirit of L_RFG). For example, imagine that the LOC.PL.F form of *pięciu* ‘five’ had the lexical entry in (19). Then such a lexical entry would be compatible with the grammatical specification of case as accusative, i.e., as [NOM+, ACC+], which would result in a constituent assigned the accusative in the grammar being realizable by a form lexically defined as bearing the locative case.

(19) (bad lexical entry)

pięciu Num (\uparrow PRED) = ‘5’
 (\uparrow PHI NUMBER) = PL
 (\uparrow PHI GENDER) = F
 (\uparrow CASE NOM) = +
 (\uparrow CASE ACC) = +
 (\uparrow CASE GEN) = +
 (\uparrow CASE LOC) = +

(20) (good lexical entry)

pięciu Num (\uparrow PRED) = ‘5’
 (\uparrow PHI NUMBER) = PL
 (\uparrow PHI GENDER) = F
 (\uparrow CASE GEN)
 $\neg(\uparrow$ CASE INS)

Instead, lexical entries should specify case via constraining statements, as in (20). This lexical entry is also more general than (19) was meant to be, as it defines a feminine form that exposes genitive, locative, and dative (see (1c) for a fuller paradigm). This is done by ensuring the presence of CASE GEN, which excludes nominative and accusative, and denying the presence of CASE INS, which excludes instrumental.

The full paper demonstrates that this analysis easily carries over to MACA in the “something interesting” construction, although in this case the analogous interplay of @CONCORD and @GEN is caused by the relevant pronouns – or the pronominal root *c(o)* – assigning the genitive to relevant adjectival dependents on top of the usual concord between (pro)nouns and such dependents.

References (abbreviated) • Asudeh, A., Bjorkman, B., Dowle, F., Myler, N., Siddiqi, D., and Sullivan, L. (2024). Fusional morphology, metasyncretism, and secondary exponence: A morphemic, realizational approach to Latin declension. LFG ’24. • Belyaev, O., Alekseev, D., Asudeh, A., Bjorkman, B., Dowle, F., Siddiqi, N., and Sullivan, L. (2025). Ossetic nominal inflection: Between morphology and syntax. “Arguments and Challenges for Lexicalism” at LFG ’25. • Caha, P. (2009). *The Nanosyntax of Case*. Ph.D. thesis, Tromsø. • Dalrymple, M., ed. (2023). *Handbook of Lexical Functional Grammar*. Language Science Press. • Dalrymple, M., King, T. H., and Sadler, L. (2009). Indeterminacy by underspecification. *JoL*, 45, 31–68. • Haug, D. and Nikitina, T. (2015). Feature sharing in agreement. *NLLT*, 34, 865–910. • Neidle, C. (1988). *The role of case in Russian syntax*. Kluwer. • Przepiórkowski, A. (1999). *Case Assignment and the Complement-Adjunct Dichotomy: A Non-Configurational Constraint-Based Approach*. Ph.D. thesis, Tübingen. • Przepiórkowski, A. and Patejuk, A. (2012). The puzzle of case agreement between numeral phrases and predicative adjectives in Polish. LFG ’12. • Rutkowski, P. and Szczegot, K. (2001). On the syntax of functional elements. In A. Przepiórkowski and P. Bański, eds., *Generative Linguistics in Poland: Syntax and Morphosyntax*, pp. 187–196. ICS PAS.