The categorial status of numerals: Evidence from Polish and Serbian

Heidi Klockmann
h.e.klockmann@uu.nl
Utrecht University (UiL-OTS)

Anja Šarić
saric@em.uni-frankfurt.de
Goethe University

1. Introduction

Within theories of categories, numerals have remained problematic because (i) within a single language, they do not form a uniform category and (ii) they often cannot be subsumed under one of the more standard categories, such as adjective, noun, or verb.

In a survey of 60 languages, Corbett (1978b) observed that numerals tend to fall somewhere between adjectives and nouns, the higher numerals behaving more noun-like.

He further observed that in the Slavic languages, the numerals form natural classes which impose different case and agreement patterns (Corbett 1978a). For example, numeral 1 behaves differently from numerals 2, 3, and 4, which again behave differently from numerals 5 and above.

In this talk, we focus on the category of these different numeral classes in Serbian and Polish.

Main claim: There is no single category “numeral” across, or even necessarily within, languages which could be used to capture numeral behavior. Rather, numerals must be treated on a case-by-case basis.

2. Serbian and Polish numerals

2.1 A theory of categories

Baker (2003) distinguishes between the three lexical categories in the following way:

(1) Nouns: No specifier, Referential index
       Adjectives: No specifier, No referential index
       Verbs: Specifier, No referential index

If we equate referential indices with valued phi-features (e.g. Pereltsvaig 2006), the following system arises.

(2) Nouns: No specifier, Valued phi-features
       Adjectives: No specifier, Unvalued phi-features
       Verbs: Specifier, Unvalued phi-features

This formulation of the traditional categories opens up the possibility of having elements which do not have complete sets of valued or unvalued phi-features, and thereby allows for deviations from more prototypical adjectival or nominal behaviors. We utilize this in the analysis to come.
2.2 Numeral 1

In both Serbian and Polish, numeral 1 behaves like a true adjective, agreeing with the quantified noun in phi-features and case.

**Serbian:**

(3) Jedna \text{ptica} je spaval.a.
\begin{align*}
\text{One.NOM.F.SG} & \quad \text{bird.NOM.F.SG} \quad \text{slept.F.SG} \\
\text{‘One bird slept.’}
\end{align*}

(4) Mala \text{ptica} je spaval.a.
\begin{align*}
\text{Small.NOM.F.SG} & \quad \text{bird.NOM.F.SG} \quad \text{slept.F.SG} \\
\text{‘A small bird slept.’}
\end{align*}

**Polish:**

(5) Jeden \text{ptak} spał.
\begin{align*}
\text{One.NOM.M.SG} & \quad \text{bird.NOM.M.SG} \quad \text{slept.M.SG} \\
\text{‘One bird slept.’}
\end{align*}

(6) Mały \text{ptak} spał.
\begin{align*}
\text{Small.NOM.M.SG} & \quad \text{bird.NOM.M.SG} \quad \text{slept.M.SG} \\
\text{‘A small bird slept.’}
\end{align*}

We propose the following phi-composition for the numeral 1 in both languages.

\begin{equation}
1: \quad [ + \text{number}, + \text{gender}, + \text{case}]
\end{equation}

Numeral 1 is an adjective, having all features initially unvalued.

2.3 Numerals 5+

In Serbian, the class of 5+ numerals includes, numerals 5 to 20, complex numerals made out of these numerals (e.g. 25-30, 35-40, …), and higher numerals such as 100 and 1000.\(^1\)

In Polish, the class of 5+ numerals are the same as in Serbian, with the exception of numerals 1000 and higher, which form their own class.\(^2\)

**Serbian:**

In Serbian, numerals 5+ are invariant for gender, and obligatorily trigger neuter singular verbal agreement.

(8) Pet \text{satova} / \text{slika} / \text{ogledala} \quad \text{je visilo} \quad \text{na zidu.}
\begin{align*}
\text{Five clock.GEN.M.PL} / \quad \text{picture.GEN.F.PL} / \quad \text{mirror.GEN.N.PL} \quad & \quad \text{is} \quad \text{hung.N.SG} \quad \text{on wall} \\
\text{‘Five clocks/pictures/mirrors hung on the wall.’}
\end{align*}

Neuter singular on the noun cannot come from the numeral because (i) coordination of two numeral phrases does not trigger masculine plural, as is the case with coordination of two

\(^1\) The numerals 100 and 1000 (and above) have an additional form which does not show the properties of 5+ numerals but instead patterns with masculine collectives (see Šarić 2014).

\(^2\) Although, under certain circumstances, numeral 1000 patterns with the 5+ numerals in Polish.
neuter singular nouns; and (ii) modifying adjectives / demonstratives cannot bear neuter singular features, which we would expect if the numeral was neuter singular.

(9) Mači i pile su trčali/ *je trčalo.
Kitten.NOM.N.SG and chicken.NOM.N.SG are run.M.PL/ *is run.N.SG.
‘A kitten and a chicken ran.’

(10) Pet pevačica i pet glumica je pevalo/ *su pevali
Five singers.GEN and five actresses.GEN is sung.N.SG / *are sung.M.PL
‘Five singers and five actresses sang.’

(11) Ovih/ *Ovo pet glumaca je dobilo nagrade.
These.GEN.M.PL/ *this.NOM.N.SG five actors.GEN.M.PL is received.N.SG awards
‘These five actors received the awards.’

Even when an adjective modifies the numeral and not the numeral-noun complex, the adjective still agrees with the noun and not the numeral.

(12) Pretrčao je dobrih/ *dobra pet milja.
Run-he is good.GEN.F.PL/ *good.NOM.N.SG five mile.GEN.F.PL
‘He ran a good five miles.’

While it is clear that numerals 5+ are not specified with neuter singular features, it is still unclear what the features of these numerals are as we do not seem to find evidence for any of them.

- **Number:** Serbian 5+ numerals do not show number agreement. However, since the numeral triggers plural on the noun, we consider it quite unintuitive for the numeral to be entirely missing a number feature. We assume numerals 5+ have a plural number feature, valued via agree with the noun.

- **Gender:** Numerals 5+ do not show gender differences – there is only one morphological form for all of them, irrespective of the gender of the noun they quantify, and they always trigger default agreement. We propose numerals 5+ lack a gender feature.

- **Case:** 5+ numerals do not inflect for case. We suggest they always get default case in the process of default agreement.

---

3 With numerals 5+ adjectives agree with the head noun, whereas with numerals 2, 3 and 4 they agree with the numeral (see Šarić, 2014 for discussion).

(1) Dva visoka dečaka su igrala fudbal.
Two tall.NOM.N.PL boys.GEN.M.SG are played.N.PL football
‘Two tall boys played football.’

(2) Pet visokih dečaka je igralo fudbal.
Five tall.GEN.M.PL boys.GEN.M.PL is played.N.SG football
‘Five tall boys played football.’

4 With numerals 2, 3 and 4, the adjective modifying the numeral must agree with the numeral.

(1) Pojeo je dobrih/ *dobra/p tripeta pice.
Ate-he is good.N.PL/ *good.NOM.N.SG two.M.PL slice.GEN.N.SG pizza
‘He ate a good two slices of pizza.’
Polish:

- **Number:** Polish numerals 5+ have a **plural number feature**. We see this through demonstrative agreement: demonstratives can agree with either the numeral or the noun – they appear plural when agreeing with the numeral.

  (13) Te i tych pięć ptaków
      Those.PL / those.GEN five birds.GEN
      ‘Those five birds’

- **Gender:** 5+ numerals trigger neuter singular agreement, yet this does not seem to be due to neuter singular features on the numeral – demonstratives cannot be neuter singular, and likewise, the coordination of two numeral phrases does not produce a plural.

  (14) *To pięć ptaków
      This.N.SG five birds.GEN

  (15) a. Pięć krzeseł i sześć biurek rozbiło / *rozbiły się.
      Five chairs.GEN and six desks.GEN broke.N.SG / *broke.NV.PL SIE
      IE ‘Five chairs and six desks broke.’
    b. Krzesło i biurko *rozbiło / rozbiły się
      Chair.N.SG and desk.N.SG *broke.N.SG / broke.NV.PL SIE
      IE ‘A chair and a desk broke.’

I propose Polish numerals 5+ **lack a gender feature.** This can be used to derive the neuter singular verbal agreement, which seems to be default agreement.

- **Case:** Polish 5+ numerals inflect for case: pięcioma**INST,** pięciu**GEN/DAT/Loc,** pięć**NOM/ACC. Thus, they carry an **unvalued case feature.**

We propose the following phi-composition for the numeral 5 in both languages:

(16) 5: [ + number, + case ]

In both Serbian and Polish, default agreement is neuter singular (e.g. Dziwirek 1990 on Polish), as observed with impersonal constructions and weather verbs.

**Serbian**

(17) Dosadno mi je.
    Bored.N.SG me.DAT is
    ‘I am bored.’
Societas Linguistica Europaea
Numerals Workshop
September 2, 2015

(18) Grmelo je.
    Thundered.N.SG is
    ‘There was thunder.’

Polish
(19) Nudziło mi się.
    Bored.N.SG me.DAT się
    ‘I was bored.’
(20) Padało.
    Rained.N.SG
    ‘It rained.’

In impersonal and weather verb constructions, the noun is inactive as it has its case feature valued, and is thus an unsuitable goal.

We propose that the neuter singular verbal agreement features with numerals are also examples of default agreement.

Why should these numerals trigger default agreement? Defectivity!

In numeral-noun constructions, the noun is marked genitive and is also an unsuitable goal. Furthermore, the numeral is feature defective (lacking gender) and cannot enter into an Agree relation with the verb. Default agreement thus occurs as a last resort (Preminger 2011).

2.4 Numerals 2, 3, and 4

Serbian:

Under certain accounts (Zlatić, 1997; Belić, 2008) numerals 2, 3, and 4 in Serbian show true adjectival and verbal agreement. These accounts assume a third number value for Serbian – dual/minor paucal. However, these accounts lack empirical evidence (see Šarić, 2014) and we here propose a novel approach.

Modified nouns are in their singular forms and they cannot be the source for verbal agreement, which is plural.

Numerals 2, 3, and 4 can freely quantify singuläria tantum nouns:

(21) Luk je istrulio.
    Onion.NOM.M.SG is rotten.M.SG
    ‘Onion got rotten’
(22) *Luk su istrulili.
    Onion.NOM.M.SG are rotten.M.PL
(23) Dva / tri / četiri luka su istrulila.
    Two / three / four onion.GEN.M.SG were rotten.N.PL
    ‘Two / three / four onions got rotten.’
However, they cannot quantify *pluralia tantum* nouns:

(24) Makaze su otupele.
Scissors.NOM.F.PL are blunt.F.PL
‘The scissors got blunt.’

(25) *Makaze je otupela*
Scissors.NOM.F.PL is blunt.F.SG

(26) *Dve / tri / četiri makaze / makaza*
Two / three / four scissors.NOM.F.PL / scissors.GEN.F.PL

(27) Dva / tri / četiri para makaza su otupela.
Two / three / four pair.GEN.M.SG scissors.GEN.F.PL is blunt.N.PL
‘Two / three / four pairs of scissors got blunt.’

This shows two things: (i) the quantified noun must be singular and (ii) it is the numeral which is the source of the verbal agreement.

If we examine the agreement features on the verb, we observe two agreement patterns, one with feminine nouns and another with masculine and neuter nouns. Quantified feminine nouns trigger feminine plural agreement and quantified masculine and neuter nouns trigger neuter plural agreement.

(28) Dve devojčice su plivale.
Two girls.GEN.F.SG are swum.F.PL
‘Two girls swam.’

(29) Dva dečaka / deteta su plivala.
Two boy.GEN.M.SG / child.GEN.N.SG are swum.N.PL
‘Two boys / children swam.’

(30) Tri / četiri devojčice su plivale.
Three / four girls.GEN.F.SG are swum.F.PL
‘Three / four girls swam.’

(31) Tri / četiri dečaka / deteta su plivala.
Three / four boy.GEN.M.SG / child.GEN.N.SG are swum.N.PL
‘Three / four boys / children swam.’

Quantified feminine nouns trigger feminine agreement as expected, and quantified neuter nouns, neuter agreement, but curiously, quantified masculine nouns do not trigger masculine agreement. When no numeral is present, masculine agreement is possible, and thus, the neuter agreement is directly triggered by the numeral.

(32) Dečaci su plivali.
Boys.NOM.M.PL are swum.M.PL
‘Boys swam.’

Thus, there seems to be a gap for masculine gender. Why?

- Masculine gender is the only gender with a set of collective numerals, so the marking of masculine gender on the numeral is already available through another set of numerals.
- Native speakers prefer the use of a collective numeral over a cardinal numeral with masculine nouns.
Numerals 2, 3 and 4 used to be nouns, so it’s not all that surprising that they have restrictions on which features they can realize.

The question then is: why do we see the masculine being replaced by neuter and not by feminine? We propose that the neuter gender is less marked than feminine, thus, it fills in for the missing masculine.

This leads us to conclude that the features of numerals 2, 3, and 4 in Serbian are the following:

(33) 2, 3, 4: \[ + \text{number (pl)}, + \text{gender (f / n)} \]

Analysis: The numerals 2, 3, and 4 have an inherently valued plural feature and agree for gender with the noun they modify. In the case of masculine nouns, the numeral cannot “read” masculine gender from the noun, hence it ends up with neuter features.

**Polish:**

With numerals 2, 3, and 4 in Polish, the numeral appears to agree in gender with the noun, and the verb matches the gender and number features of the quantified noun.

(34) Dwie dziewczyny pływały.
Two.NOM.F girl.NOM.F.PL swam.NV.PL
‘Two girls swam.’

(35) Dwaj chłopcy pływali.
Two.NOM.V boy.NOM.V.PL swam.V.PL
‘Two boys swam.’

(36) Dwa krzesła / stol/ rozbiły się.
Two.NOM.M/N chair.NOM.N.PL / table.NOM.M.PL broke.NV.PL SIE
‘Two chairs / tables broke.’

(37) Trzy / cztery krzesła / stol/ / kuchenki rozbiły się.
‘Three / four chairs / tables / cookers broke.’

(38) Trzej / czterej chłopcy pływali.
Three.NOM.V / four.NOM.V boy.NOM.V.PL swam.V.PL
‘Three / four boys swam.’

This leads us to conclude that the features of the numerals 2, 3, and 4 in Polish are:

(39) 2, 3, 4: \[ + \text{number (pl)}, + \text{gender (matches the noun)} \]

At the moment, numerals 2, 3, and 4 seem indistinguishable in feature set from numeral 1. However, numeral 1 and numerals 2, 3, and 4 differ from each other in a number of important ways.

Firstly, while numeral 1 can freely appear with singular and plural nouns, numerals 2, 3, and 4 are restricted to quantifying plural nouns:
Furthermore, numerals 2, 3, and 4 (and additionally 5+) have an optional agreement paradigm with masculine human (termed “virile”) nouns – both the numeral and the noun appear in the genitive, with the verb marked default; numeral 1 does not share this pattern.

Analysis: The numerals 2, 3, and 4 have an inherently valued plural feature and agree for gender with the noun they modify. The verb can then agree with the numeral-noun complex.

2.5 Summary

- In both Serbian and Polish, numeral 1 is an adjective, agreeing in gender, number, and case with the quantified noun.

- In both Serbian and Polish, numerals 5+ are defective, triggering default agreement. In neither language is there evidence for a gender feature, and we treat the lack of gender as the source of this defectivity.

- In both Serbian and Polish, numerals 2, 3, and 4 are inherently plural, but acquire gender morphology through agreement with the noun. In Serbian, this is realized differently than for Polish, because there is a gap for masculine gender. These numerals are not fully parallel, however, because the form of the noun they select for differs between the languages – genitive singular nouns in Serbian and plural nouns in Polish which are case-marked in the external environment.

3. Conclusions

- We treat nouns as having inherent gender and number and adjectives as being dependent on some other element for their gender and number specification, this being represented by the nature of the phi-features, either valued or unvalued.

- According to this treatment of categories, Serbian and Polish numerals, particularly 2, 3, 4, and 5+, do not fully pattern with either adjectives or nouns, but form intermediate categories with their own independent properties.

- Our findings are in line with Corbett (1978a, 1978b), who shows that numerals fall between adjectives and nouns and furthermore, that in the Slavic languages, they form natural groups that trigger different agreement patterns.
Even in two closely related languages, like Serbian and Polish, which make similar distinctions in their numeral classes, we find that these numeral classes are not entirely identical. Furthermore, within a single language, we have strong evidence that there is no single category of numeral. We conclude that numerals must be considered on a case-by-case basis for each language.

Food for thought: We notice that numerals in these languages are parasitic on the categories of adjective and noun. We want to pose the question, not of why numerals are so special, but rather why are numerals not special enough to deserve their own category?

4. References


