Deriving agreement mismatches with Polish numerals
Heidi Klockmann (Utrecht University, UiL-OTS)

h.e.klockmann@uu.nl

What are agreement mismatches?

Mismatches between the features of the probe (e.g. verb) and would-be goal (e.g. subject), which in some languages are triggered by numerals:

**Polish:**

(1) a. Ptaki spały  
Bird.M.PL.NOM slept.NV.PL 
“Birds slept”

b. Pięć ptaków spało  
Five bird.M.PL.GEN slept.N.SG 
“Five birds slept”

**Inari Sami (Toivonen, 2007):**

(2) a. Almah kuá’lásteh onne  
Man.PL.NOM fish.3PL today 
“The men are fishing today”

b. Kyehti alma láin meecist  
Two man.SG.GEN were.3DU forest.LOC 
“Two men were in the forest”

Considering subject-verb number agreement, there are four logical possibilities, two involving agreement mismatches and two involving full agreement. All are found, the languages below reporting behavior with numerals:

**Table 1: Subject-Verb number agreement possibilities**

<table>
<thead>
<tr>
<th>Subject #</th>
<th>Verb #</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singular</td>
<td>Hungarian, Turkish, Finnish</td>
<td>Inari Sami, Skolt Sami, Russian (2-4)</td>
<td></td>
</tr>
<tr>
<td>Plural</td>
<td>Polish (5+), other Slavic languages (5+)</td>
<td>English, Dutch, Spanish, Polish (2-4)</td>
<td></td>
</tr>
</tbody>
</table>

In this talk, I focus on the Plural-Singular agreement mismatch as found in Polish. I claim:

- Polish agreement mismatches are the result of default agreement (Preminger, 2011)
- Default agreement is triggered by the probe’s inability to find an active phi-complete goal, due to a missing feature on 5+ numerals (phi-incomplete) and genitive case-marking on the noun (inactive)
- This analysis of numerals suggests an answer to the treatment of categories and semi-lexicality

---

1 In association with the VIDI Project “The Uniformity of Linguistic Variation”, http://linguisticvariation.com/en/

2 Abbreviations: ACC = accusative; ANIM = animate; DAT = dative; DEV = devirilized gender; DU = dual; F = feminine; GEN = genitive; INAM = inanimate; INST = instrumental; LOC = locative; M = masculine; MP = masculine personal gender; N = neuter; NOM = nominative; NV = non-virile gender; PART = particle; PL = plural; SG = singular; V = virile gender
1. 5+ induced agreement mismatches

Numerals 5+, but not numerals 1, 2, 3, or 4 trigger an agreement mismatch and genitive marking on the noun:

(3) Pięć ptaków spało
    Five bird.M.PL.GEN slept.N.SG
    “Five birds slept”

(4) a. Jeden ptak spał
    “One bird slept”

   b. Dwa / trzy / cztery ptaki spały
    Two.M.NOM / three.NOM / four.NOM bird.M.PL.NOM slept.NV.PL
    “Two / three / four birds slept”

Numerals 1, 2, 3, and 4, but not 5+, show gender agreement on the numeral itself (with accompanying agreement on the noun and NOM on the noun and numeral, c.f. (4)):

Table 2: Gender agreement in subject (nominative) position for numerals 1, 2, 3, 4, and 5

<table>
<thead>
<tr>
<th></th>
<th>Masculine</th>
<th>Neuter</th>
<th>Feminine</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>jedni.PL</td>
<td>jednego</td>
<td>jedno</td>
</tr>
<tr>
<td>2</td>
<td>dwaj</td>
<td>dwa</td>
<td>dwie</td>
</tr>
<tr>
<td>3</td>
<td>trzej</td>
<td>trzy</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>czterej</td>
<td>cztery</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>pięciu</td>
<td>pieć</td>
<td></td>
</tr>
</tbody>
</table>

Note: While it is clear that trzej-trzy and czterej-cztery is gender agreement (due to the presence of subject-verb agreement and nominative case marking), it is debatable whether the pięciu-pięć distinction is also gender agreement. I claim it is not; instead pięciu is a genitive form – I return to this when I discuss masculine personal gender in sections 3 and 4.

2. Analysis of 5+ induced agreement mismatches

Gist of the analysis:

- The neuter singular marking on the verb is default agreement (Preminger, 2011)
- Numerals 1, 2, 3, and 4 resemble adjectives and agree in gender (and number) with the noun; as they only agree and do not assign case, they do not provoke default agreement
- Numerals 5+ are semi-lexical, phi-incomplete nouns which assign genitive case to their complement (via a case-assigner); they provoke default agreement

2.1 DEFAULT AGREEMENT

Default Agreement (Preminger, 2011): Agree is obligatory, but its success is not. If agree fails, default features are inserted.

---

3 In this section, I focus on non-masculine personal gender, due to the complications associated with the masculine personal gender. I return to it in sections 3 and 4.
In Polish, default verbal features are neuter singular (Dziwirek, 1990):

**Infinitival subjects**

(5) 
Wyłysieć byłoby dla mnie tragedią
To.grow.bald would.be.N.SG for me.GEN tragedy.INST
“To grow bald would be a tragedy for me” (modeled after Swan 2002: 391)

**Impersonals**

(6) 
Nudziło mi się
Was.bored.N.SG me.DAT PART
“I was bored”

**Weather verbs**

(7) 
Padało
Rained.N.SG
“It rained”

---

**Claim: 5+ induced agreement mismatches are instances of default agreement**

2.2  **Featural Make-up of Numerals**

**Numeral 1:** Agrees in
- number (*jedne*NV.PL, *jedniv*PL – relevant with pluralia tantum words, such as *drzwi*PL ‘door’),
- and case (for masculine inanimate singular nouns: *jeden*NOM/ACC, *jednegogen*, *jednym*INST/LOC, *jednemu*DAT)

Therefore, it has unvalued gender, number, and case features

**Claim:** Numeral 1 is syntactically an adjective.

**Numerals 2,3,4:** Agree in
- but not in number as they have no number morphology and are restricted to modifying plural nouns (two bird(s): *dwa ptaki*PL, *dwa ptak*SG)

Therefore, they have unvalued gender and case, but valued number (assumed to be PL)

**Claim:** Numerals 2,3,4 syntactically resemble adjectives, but are a cross between adjective and noun (their nominal behavior to be highlighted in sections 3 and 4).

**Numerals 5+:**
- Have case morphology (*pięć*NOM/ACC, *pięciu*GEN/LOC/DAT, *pięcioma*INST)
- Do not agree in number, as they have no number morphology and are restricted to modifying plural nouns (five bird(s): *pięć ptaków*PL, *pięć ptaka*SG)
- Do not agree in gender (see table 2 and following note)
Seem to be missing a gender feature altogether (also claimed in Dziwirek, 1990):

*Historical argument:* 5+ numerals were once nominal heads that triggered feminine singular agreement on the verb (Miechowicz-Mathiason & Dziubala-Szurek, 2012; Rutkowski, 2006). Now they do not which could indicate a loss of gender.

(8) Ona siedm panien szła  
that.F.NOM seven.NOM maidens.GEN walked.F.SG  
"Those seven maidens were walking" (Old Polish – Rutkowski, 2006: 93)

*Coordination argument:* If 5+ numerals had a neuter gender feature (as they trigger N.SG on the verb), we would expect the coordination of two 5+ numeral phrases to lead to NV.PL on the verb, like neuter nouns do – however, it still leads to N.SG

(9) a. Krzesła i biurka rozbiły się  
Chair.N.PL.NOM and desk.N.PL.NOM broke.NV.PL PART  
"Chairs and desks broke"

b. Pięć krzeseł i sześć biurek rozbiło się  
Five chair.GEN and six desk.GEN broke.N.NOM PART  
"Five chairs and six desks broke"

Therefore, they have unvalued case, missing gender, and valued number (PL)

**Claim:** 5+ numerals are semi-lexical, phi-defective (in gender) nouns which assign genitive to their complements.

**Table 3: Feature Specifications**

<table>
<thead>
<tr>
<th>Gender</th>
<th>Number</th>
<th>Case</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>unvalued</td>
<td>unvalued</td>
</tr>
<tr>
<td>2,3,4</td>
<td>unvalued</td>
<td>PL</td>
</tr>
<tr>
<td>5+</td>
<td>—</td>
<td>PL</td>
</tr>
</tbody>
</table>

2.3 **DERIVING THE VERBAL AGREEMENT FACTS**

The interaction of the phi-features of these numerals with the operation Agree (Chomsky 2000, 2001) produces the patterns shown in section 1.

**Numeral 1:** Syntactic adjective. Agrees in gender/number/case with the noun (10a). Verb agrees with the construction (10b)
Numerals 2,3,4: Adjective-Nouns – agree like an adjective but carry the structure of a noun (with a case-assigner that is inactive for non-masculine personal gender). Numeral agrees in gender with the noun (11a). Verb agrees with noun (shared case feature between numeral and noun as in Pesetsky and Torrego (2004), or default nominative on numeral) (11b).

(11)  
\[
\begin{align*}
\text{a.} & \quad \text{b.} \\
& \quad \text{DP/NP} \\
& \quad \text{A/N} \\
& \quad 2,3,4 \\
& \quad \text{FP} \\
& \quad \text{[__Gen]} \\
& \quad \text{[[Pl-Num]}} \\
& \quad \text{[__Case]} \\
& \quad \text{AGREE} \\
\end{align*}
\]

Numerals 5+: Semi-lexical, phi-defective nouns. FP assigns genitive to the noun (12b). Verb cannot agree with the numeral because it is phi-incomplete, missing gender (12c). Verb cannot agree with the noun because it is already genitive and inactive (12c). Verb receives default agreement features (neuter singular) (12d). Numeral gets default nominative (12d).

(12)  
\[
\begin{align*}
\text{a.} & \quad \text{b.} \\
\text{FP} & \quad \text{TP} \\
\text{F} & \quad \text{AGREE} \\
\text{DP/NP} & \quad \text{T} \\
\text{N} & \quad \text{verb} \\
\text{[__Gen]} & \quad \text{[+Gen]} \\
\text{[[Pl-Num]} & \quad \text{[[Pl-Num]} \\
\text{[__Case]} & \quad \text{[nom-Case]} \\
\text{Noun} & \quad \text{Noun} \\
\text{[+Gen]} & \quad \text{[+Gen]} \\
\text{[[Pl-Num]} & \quad \text{[[Pl-Num]} \\
\text{[nom-Case]} & \quad \text{[nom-Case]} \\
\text{AGREE} & \quad \text{default} \\
\end{align*}
\]
Deriving agreement mismatches with Polish numerals
Heidi Klockmann

3. Gender induced agreement mismatches

Numerals 5+ trigger genitive marking\(^4\) on the numeral itself in the presence of masculine personal gender (in contrast to non-masculine personal gender, 13b).

\begin{equation}
\begin{aligned}
a. \textbf{Pięć} & \text{ chłopców} & \text{ spało} \\
& \text{ Five.M} & \text{ boy.PL \text{M} \text{GEN}} & \text{ slept.PL.N} \\
& \text{ “Five boys slept”} \\
b. \textbf{Pięć} & \text{ ptaków} & \text{ spało} \\
& \text{ Five} & \text{ bird.PL \text{M} \text{GEN}} & \text{ slept.PL.N} \\
& \text{ “Five birds slept”} \\
\end{aligned}
\end{equation}

Numerals 2,3,4 do one of two things in the presence of masculine personal gender: (a) they agree with the noun in gender and show full verbal agreement (as they did in section 1) or (b) they and the noun are marked as genitive accompanied by an agreement mismatch:\(^5\)

\begin{equation}
\begin{aligned}
a. \textbf{Dwaj} & \text{ chłopcy} & \text{ spali} \\
& \text{ Two.M \text{NOM}} & \text{ boy.PL \text{M} \text{NOM}} & \text{ slept.V.PL} \\
& \text{ “Two boys slept”} \\
b. \textbf{Dwóch} & \text{ chłopców} & \text{ spało} \\
& \text{ Two} & \text{ boy.PL \text{M} \text{GEN}} & \text{ slept.PL.N} \\
& \text{ “Two boys slept”} \\
\end{aligned}
\end{equation}

\(^4\) Contra Franks (1994, 2002), Przepiórkowski (2004) and Rutkowski (2002), I claim that this is genitive and not accusative case. See Klockmann (2012) for a discussion of and arguments for this choice.

\(^5\) The difference is stylistic, (a) being used more in written Polish and (b) more in spoken Polish (Swan, 2002).
Numerals 1 agrees with the noun accompanied by full verbal agreement:

(15) Jedni chłopcy spali
    One.MP.PL.NOM boy.MP.PL.NOM slept.V.PL
    “Some boys slept”

4. Analysis of gender induced agreement mismatches

Gist of the analysis
- The neuter singular marking found on the verb is again default agreement
- Featural specifications of the numerals is the same as above
- Genitive marking on the numeral is due to “case leaking” whereby the case assigned to the noun is also assigned to the numeral, triggered by masculine personal gender – this is claimed to be a case of Cyclic Agree (Rezac, 2003)

4.1 CYCLIC AGREE

*Cyclic Agree* (Rezac, 2003): If a probe does not deactivate after a search into its c-command domain, it can also search upwards when new material is Merged. This is a consequence of the *Earliness Principle* of Pestesky and Torrego (2001: 400) which states that “an uninterpretable feature must be marked for deletion as early in the derivation as possible”.

(16) a. Probed XP Goal

b. XP Probe Goal

4.2 DERIVING THE VERBAL AGREEMENT FACTS

*Numeral 1*: Identical to before. See (10).

*Numerals 2,3,4 & 5+:*
- FP selected by these numerals is defective; as a result, $f$ cannot deactivate after agreement with a masculine personal noun (17a), so it extends its search space upwards and also agrees with and assigns case to the numeral (17b). This is cyclic agree (Rezac, 2003).

(17) a. FP

b. DP/NP

```
    N
   /   \
[2,3,4]  [5+]
[Gen]   [Gen]
[Pl-Num] [Pl-Num]
[gen-Case] [gen-Case]

    F
   /   \  \
[Gen]   [Num]
[MP/DEVGen] [MP/DEVGen]
[Pl-Num] [Pl-Num]
[gen-Case] [gen-Case]

    N
   /   \  \
[Gen]   [Num]
[MP/DEVGen] [MP/DEVGen]
[Pl-Num] [Pl-Num]
[gen-Case] [gen-Case]
```
• (2,3,4 numerals agree in gender with the noun)
• The verb cannot agree with the numeral or the noun as both are inactive (18)

(18)

Verb gets default features instead (19)

(19)
6. Implications and Conclusions

This analysis defines numerals in terms of phi-features and derives their behaviors as a consequence of the interaction of those features with Agree.

If we consider numerals in the context of adjectives and nouns, we see that the numerals 2,3,4 and 5+ show properties of both adjectives and nouns:

Table 4: Adjectival, nominal, and semi-lexical properties of numerals

<table>
<thead>
<tr>
<th>Property Type</th>
<th>Property</th>
<th>Adj</th>
<th>2,3,4</th>
<th>5-10, 100</th>
<th>1000</th>
<th>Noun</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjectival</td>
<td>Case agreement (structural positions)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Case agreement (oblique positions)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Gender agreement</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Number agreement</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Nominal</td>
<td>Genitive assignment (structural positions)</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Genitive assignment (oblique positions)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Has plural morphology</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Semi-lexical</td>
<td>5+ mismatches</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Gender mismatches</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Case alternation⁶</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

Based on Baker (2003)’s theory of categories, one could say that nouns carry all valued phi-features and adjectives all unvalued phi-features. This would be a distinguishing factor between adjectives and nouns, and a defining feature of those categories.

My analysis of Polish numerals suggests that it is also possible to have combinations of valued, unvalued, and missing phi-features within a single element. These elements show intermediate (e.g. both nominal & adjectival) and/or unexpected (e.g. semi-lexical) behaviors, as in the table above. They constitute a class of semi-lexical elements.

One could imagine we have a discrete “continuum” between adjective and noun, whereby the possible number of in-between categories is constrained by the number of features involved and the different combinations of valued, unvalued, and missing features that are possible (where the total possibilities are \(3^k-2\), \(k\) being the number of phi-features involved), e.g.:

(20) Lexical | . . . . . . . . Semi-Lexical | . . . . . . . . | Lexical

Adjective (all unvalued) | . . . . . . 2,3,4 | . . . . . . 5+ | . . . . . . 1000 | Noun (all valued)

⁶ Not discussed here. See Klockmann (2012) for a full discussion of this phenomenon
Questions for future research:

- Why does masculine personal gender trigger case-leaking?
- What can we say about the nature of the case-assigning FP? Does it search for features? Can it also receive default features if it encounters only improper goals (as would be the case with masculine personal gender)?
- Can a feature based account be applied to the numeral systems of other Slavic languages? What of agreement mismatches in other languages, such as Inari Sami?
- Are certain features more adjectival or nominal, i.e. do they contribute in a different way to the nouniness or adjective-ness of a semi-lexical element?
- What other sorts of semi-lexical elements are there, and how do they fit into the proposed scheme?

7. References


